NEW MEDIT

CIHEAM BARI

MEDITERRANEAN JOURNAL OF ECONOMICS, AGRICULTURE AND ENVIRONMENT

Poste Italiane Spa Spedizione in Abbonamento Postale Periodico ROC Centro Nord aut. N° 0029 - € 15,00.

4
QUARTERLY
VOL. XXIII - N. 4
DECEMBER
2024

Stakeholder's willingness to pay for Tunisian forest services
Islem Saadaoui

Exploring the determinants of consumers' adoption of online grocery shopping in Cyprus
Marianthi Giannakopoulou, Stavroula Ioannou, Andreas Stylianou, Georgios Adamides

The Impacts of climate change on agriculture sector in Mediterranean Countries
Yasmine Abd El-Nasser, Ehab Abdelaziz Ibrahim

On Farm non-agricultural activities: Recent evolution and dynamics in Portugal
António Xavier, Maria do Socorro Rosário, Maria Leonor Silva Carvalho, Maria de Belém Costa Freitas

Timely repayment of agricultural loans in Serbia
Dejan Mirč, Marina Ignjatović, Dragan Milić, Vladimir Crnojević, Nedeljko Tica

Bologna University Press

ISSN: 1594-5685
www.newmedit.iamb.it


CIHEAM
BARI

NEW MEDIT

MEDITERRANEAN JOURNAL OF ECONOMICS, AGRICULTURE AND ENVIRONMENT

Vol. XXIII - n. 4/2024

Editor-in-chief
Maurizio Raeli

Managing Editor
Giulio Malorgio

Institutional Relations Manager
Debora Degl'Innocenti

Editorial office
Bologna University Press
Via Saragozza, 10
40123 Bologna (Italy)
tel.: +39 051 232882
fax: +39 051 221019
email: newmedit@iamb.it

Paper submission
http://www.newmedit.iamb.it

Publisher
Bologna University Press
Via Saragozza, 10
40123 Bologna (Italy)
tel.: +39 051 232882
fax: +39 051 221019
email: comunicazione@buponline.com

Subscription rate
Print: Italy: € 40; Foreign: € 90.

Subscription office
ordini@buponline.com

Abstract and Index citation
NEW MEDIT is indexed in: SCOPUS, EBSCO, ISI Web Science, CAB Abstracts, EconLit, AGRIS/FAO database

Web page
http://www.newmedit.iamb.it

ISBN: 979-12-5477-545-5

ISSN: 1594-5685

ISSN online: 2611-1128

DOI: 10.30682/nm2404

Graphic Layout
DoppioClickArt – San Lazzaro (BO)

Cover design
Debora Degl'Innocenti

Registrazione
Tribunale Ordinario di Bari, n. 1546 del 4/1/2002

Direttore Responsabile
Giulio Malorgio

NEW MEDIT è associato alla

Contents

Foreword pag. 1

Islem Saadaoui
Investing in nature: Stakeholder's willingness to pay for Tunisian forest services pag. 3

Marianthi Giannakopoulou, Stavroula Ioannou, Andreas Stylianou, Georgios Adamides
Exploring the determinants of consumers' adoption of online grocery shopping using the Theory of Planned Behavior: An empirical study pag. 17

Yasmine Abd El-Nasser, Ehab Abdelaziz Ibrahim
The impacts of climate change on agriculture sector - Case study: Mediterranean countries pag. 31

António Xavier, Maria do Socorro Rosário, Maria Leonor Silva Carvalho, Maria de Belém Costa Freitas
On farm non-agricultural activities: Recent evolution and dynamics in Portugal pag. 45

Dejan Mirč, Marina Ignjatović, Dragan Milić, Vladimir Crnojević, Nedeljko Tica
Timely repayment of agricultural loans: Evidence from Serbian farmers pag. 67

Ons Kammarti, Marielle Montginoul, Hamadi Habaieb
Household willingness to contribute financially to improve rural water access in Tunisia pag. 85

A. Malek Hammami, John Beghin
The determinants of US olive oil imports pag. 103

Gabriel Chico Viegas, Ana Marta-Costa, Rui Fragoso, Edgar Cambaza
The sustainability of sugarcane production: A systematic review of sustainability indicators pag. 121

Gülçin Koç, Cennet Oğuz
Cooperative branded food purchasing behaviour: Understand the role of consumer social responsibility, trust, and agri-rural ties through Theory of Planned Behaviour pag. 135

Foreword

Saadaoui's paper explores the economic value of Aleppo pine forests, in the border region of central Tunisia. The objectives is to assign a monetary value to forest conservation, considering the region's specific socio-economic context. Employing the Contingent Valuation Method the author highlights the feasibility and importance of reconciling economic development with ecological sustainability in a critical region.

Giannakopoulou et al. examine the behavioral model of consumers and the factors that influence adoption of online grocery shopping of the typical adopter profile in Cyprus. The authors show that perceived behavioral control carries more weight than intention in explaining online grocery shopping behavior, and that two antecedent constructs — attitude and subjective norms— have direct influence on intention for online grocery shopping.

El-Nasser and Ibrahim deal with the impact of climate change on food security in twelve Mediterranean countries. The results show that climate change has a negative impact on the agriculture sector for many Mediterranean countries Both temperature and CO2 emissions have a significant negative effect on food production and inflation in the selected countries. Moreover, combining climate change with a rapid increase in population growth rate will put more pressure on the agriculture sector.

On-farm non-agricultural activities in Portugal analyzed by *Xavier et al*. The results provide important insights of the on-farm non-agricultural activities dynamics and diversification as a factor for the development of farms, enhancing the endogenous resources of the territories and contributing to the multifunctionality of rural areas.

Mirč et al. faced the factors influencing timely loan repayment of smallholder farmers in Serbia. 1735 liquidated loans are used, collecting a set of 36 feasible determinant variables. The study's findings provide insight into the critical factors in substantially achieving a high repayment rate on borrowed funds.

Kammarti, Montginoul and Habaieb examine the willingness of households individually connected to locally managed water networks to contribute financially towards the rehabilitation of an existing water supply system and the implementation of a new water supply system managed by the national operator in Tunisia. Findings indicate that households are willing to pay 22% of the per-household costs of rehabilitation of an existing locally man-

aged network. Findings also indicate that the willingness to contribute financially towards the implementation of a given project is influenced by household characteristics and the state of the water supply within a household.

Hammami and Beghin examine determinants of U.S. olive oil imports and their dynamics in response to foreign supply shocks and changes in U.S. demand. On the supply side, determinants include exporters' capacity to export, multilateral trade resistance, and immigrants' networks to the U.S. On the consumer side, robust determinants encompass U.S. GDP, import unit value, and immigrant network effects. Migrants' stock, exporters' GDP and population, and total export revenues increase the probability of an exporter entering the U.S. market.

Chico Viegas et al. aim to identify indicators to assess the sustainability of sugarcane production. The results show that the most suggested indicators were Greenhouse gas emissions, Water use, Water quality, Employment generation, and Initiatives to promote the Local community's welfare, Profit, and Distance to sugar mills. Further, research combining the three dimensions of sustainability and those which separately evaluate the sustainability of sugarcane in the production stage on the field and in the factory is recommended.

Koç and Oğuz analyzed by the theory of planed behavior whether consumers' Social Responsibility, trust and agri-rural interest play role in purchasing the collective brand. The results confirmed that Social Responsibility, trust and agri-rural interest are important variables in consumers' collective brand purchase. The results imply that consumers' purchasing for the collective brand can be improved by increasing consumers' Social Responsibility levels. Thus, the programs focused on rising consumer Social Responsibility towards agri-rural development can also support consumers' purchase of the collective brand.

Investing in nature: Stakeholder—s willingness to pay for Tunisian forest services

Islem Saadaoui*

DOI: 10.30682/nm2404a
JEL codes: O13, Q57

Abstract

This study explores the economic value of Aleppo pine forests, a unique and threatened ecosystem in the border region of central Tunisia. These forests play a vital role in supporting small rural communities, but face increasing pressures and restrictions on their use. This research aims to assign a monetary value to forest conservation, considering the region's specific socio-economic context. Strategies for empowering local residents as key actors in developing sustainable cross-border initiatives are further investigated. Employing the Contingent Valuation Method, a survey of 350 local residents and international users was conducted to assess their Willingness to Pay for forest conservation efforts. Logistic regression analysis revealed that sociodemographic factors, such as monthly income and preferred payment method, significantly influence both and the likelihood of participation. These findings highlight the feasibility and importance of reconciling economic development with ecological sustainability in this critical region.

Keywords*: Economic assessment, Ecosystem service, Regional planning, Cross-border development initiative, Contingent valuation method.*

1. Introduction

Natural ecosystems generate goods and services that are essential to society's well-being (Haines-Young and Potschin, 2010). However, most of these products are exploited outside the market, without user fees for the population (Cenamor and Frishammar, 2021; Spangenberg and Settele, 2010; Farley and Costanza, 2010). Consequently, the absence of monetary value does not reflect consumers' willingness to pay for ecosystem and landscape preservation (Tian *et al*., 2020; Martín-López *et al*., 2007). This lack of economic value has led to the allocation of zero value to ecological services in decision-making (Mandle *et al*., 2021; Ouyang *et al*., 2020), resulting in the over-exploitation of ecosystems worldwide.

The study region, the Aleppo pine forests of central Tunisia, has suffered from overexploitation of its natural and environmental resources (Saadaoui *et al*.,2018), sometimes associated with irreversible degradation of vegetation cover caused by human and/or natural actions (Islem *et al*., 2014; Saadaoui, 2016).

These actions seriously threaten the ecological balance necessary for the sustainability of ecosystems and natural landscapes in Tunisia (Hammami *et al*., 2023; Jdaidi *et al*., 2023; Taghouti *et al*., 2021; Khalfaoui *et al*., 2020; Hasnaoui and Krott, 2019; Achour *et al*., 2018; Daly-Hassen *et al*., 2017; Daly-Hassen and Croitoru, 2013; Campos *et al*., 2008). The characteristics and components of these landscapes, as well as socio-economic evolution and the rapid urban

* Université du Littoral Côte d'Opale - ULR 4477 - TVES - Territoires Villes Environnement & Société, Dunkerque, France.
Corresponding author: islem.saadaoui@univ-littoral.fr

sprawl of recent decades, explain these phenomena. To overcome these problems, it is essential to develop appropriate tools to preserve the vital resources to sustainable development in the study region. Appropriate procedures for valuing environmental assets as "gifts of nature" are needed to raise awareness of the need to preserve them.

In general, there is no conventional market for this type of environmental good. Consequently, it is imperative to set up procedures that take into account the specific characteristics of forest and mountain landscapes. Forests represent a first-rate natural and environmental resources (Brockerhoff *et al*., 2017; Mori *et al*., 2017; Dobbs *et al*., 2011), offering a variety of goods and services that can be grouped into three main categories:

Production of traditional goods with a value derived directly from the market: This category includes everyday forest products such as wood, as well as non-wood products such as mushrooms, honey and medicinal plants.

Environmental services with a use value that can be estimated using methods based on market-revealed preferences: these include soil protection, erosion control, reservoir silting prevention and recreational activities.

The total economic value (TEV) of an ecosystem encompasses all the benefits people derive from it, both directly (use values) and indirectly (passive values) (Davidson, 2013). Environmental services with no direct use value, but characterized by non-use values (option values, existence values, bequest values) assessed using non-conventional approaches based on stated values (Contingent Valuation Methods): Examples of these services include biodiversity, pollution reduction, etc. (Mori *et al*., 2017; Chazdon *et al*., 2016).

This is the background to this study. Local authorities frequently express the need to better understand user preferences. What landscape attributes are they looking for? What is the users willingness to pay for the various attributes of a natural or recreational landscape? When a territorial project can have an impact on the environment, what are the criteria for valuing it, especially in a context of budgetary constraints, especially when the project does not involve marketable goods?

2. Methodology

Assessing the value of landscape attributes in natural recreational areas is a challenge due to the absence of a market with defined prices. Consequently, it is necessary to devise a method for assigning values to the landscape attributes of selected resorts, values that could be used in public decision-making. With this in mind, various preference assessment methods have been devised and developed to compensate for the absence of a market.

Two categories of methods can be distinguished: on the one hand, direct methods that simulate a market for non-market goods using one or more hypothetical scenarios (such as contingent valuation or choice experiments); on the other hand, those that reveal individuals' preferences by observing their behavior on complementary (such as hedonic prices, travel costs) or substitute (such as protection costs) markets (Oueslati *et al.*, 2008).

The contingent valuation method (CVM) has long been a popular tool for valuing non-market environmental assets. Recent studies continue to confirm its effectiveness (Manero *et al*., 2024; Grazhdani, 2024; Blasi *et al*., 2023; Baymuminova *et al*., 2023; Raihan, 2023; Manero *et al*., 2022; Viti *et al*., 2022; Perni *et al*., 2021; Cuccia, 2020; Guijarro and Tsinaslanidis, 2020; Rogers *et al*., 2019; Dupras *et al*., 2018; Krause *et al*., 2017; del Saz-Salazar and Guaita, 2013), building upon a strong foundation established in earlier research (Randall *et al*., 1983; Durden and Shogren, 1988; d'Arge and Shogren, 1989; Milne, 1991; Bateman and Turner, 1992; Jones, 1997; Oglethorpe and Miliadou, 2000)

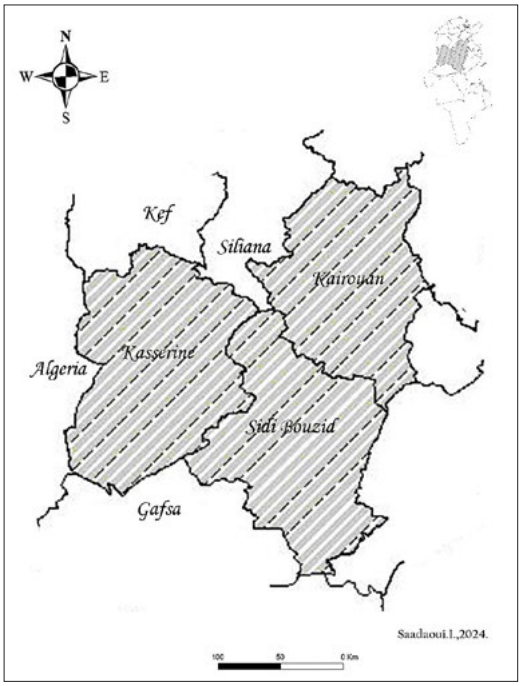
It involves a direct survey of park users, asking them about their willingness to pay (WTP) based on hypothetical scenarios. Each scenario represents a change in one or more attributes, with an associated means of payment.

The integration of biodiversity conservation into development plans for cross-border mountain areas calls on a variety of concepts and processes, most often combining considerations relating to the ecosystem, nature, heritage-identity and economic interest of the environments (Saadaoui *et al*., 2018).

This study proposes applying the contingent valuation method based on carrying out a field survey to determine the price that each person would be prepared to pay, in other words their willingness to pay for the conservation of an environmental asset. This study aims to conserve the nature of a small forest that has been unused for around a century, and to develop a natural recreational area in the cross-border mountains. The contingent valuation method is being tested at two sites: Dernaya forest (unexploited forest), an area of fallow land and forest clearings (Bouchebka). The research aims to identify two dinstinct willingness-to-pay values: the first is a conservation value (CV) and the second a management value (MV).

The criteria defined for CV and MV are determined using data collected from inhabitants of cross-border areas. The contingent valuation method is applied to a sample of 300 households from three regions in the mountain areas and 50 Algerian passenger respondents from the Bouchebka region (in Feriana).

Figure 1 - Location map of the study site.



2.1. *Contingent Valuation Method*

Contingent Valuation Method (CVM) is a survey-based technique used to estimate the value people place on environmental resources that lack a traditional market. People are asked their willingness to pay (WTP) to avoid degradation or improve these services.

How it Works: CVM creates a hypothetical market scenario where respondents are presented with a specific change in the ecosystem's condition and asked their WTP for that change. The researcher and respondent engage in a dialogue (through a well-designed questionnaire) to understand the respondent's true value for the ecosystem.

Consumer theory suggests that people make choices to maximize their satisfaction (utility). This study leverages this theory by using CVM to assign an economic value to environmental assets based on people's WTP. Essentially, the survey creates a hypothetical market to capture subjective values and estimate an objective value that could inform pricing decisions for environmental conservation or development projects. This section will detail the survey design and questionnaire structure used for data collection.

2.1.1. *Site selection*

Location: This study investigates two sites in the Kasserine governorate on the Tunisian-Algerian border: Allepo pine Forest of Dernaya and the Bouchebka Allepo pine forest clearings (Figure 1).

Selection Criteria: Several factors influenced site selection. Firstly, we chose the central Tunisia due to its proximity to the Algerian market (Bouchebka being the main border crossing). Secondly, we aimed for consistency in influencing factors by selecting sites with similar natural features (leisure parks, hiking trails, unexploited forests). To ensure variety in landscapes, the study sites were chosen for their contrasting aesthetic appeal (family-friendly atmosphere, diverse leisure areas, cultural heritage, etc.). Both locations are situated in the Feriana region, which is easily accessible to the target population for the proposed project.

2.1.2. *Assumptions and scenario design*

With the study scope defined, we propose two hypotheses to guide the research. Understanding people's payment behavior requires assumptions

about the explanatory power of certain variables. One hypothesis aligns with the theory of citizen well-being:

Hypothesis 1: Users recognize the need to contribute financially to support a sustainable development project in the study area.

Respondents' express payment intentions based on the specific scenario presented. Therefore, a clear and well-defined scenario is crucial. Offering the option to propose fixed prices or payment ranges significantly reduced the number of refusal-to-pay responses.

Hypothesis adopted during the work carried out: The analysis considers a sample of N indexed individuals i = 1, ..., N. Analysis focuses on whether one of the two projects was implemented for each individual in the sample, and note yi the coded variable associated with the project in question. Let ∀i ∈ [1, N]:
yi = (1.0) if the project has been carried out for individual i, and if the project has not been carried out for individual i; yi = (1.1).

Note here the choice of the (0, 1) coding commonly used for dichotomous models.

This allows us to define the probability of the event occurring as the expectation of the coded variable yi, since:

$$E(yi) = Prob(yi = 1) \times 1 + Prob(yi = 0) \times 0 = Prob(yi = 1) = pi$$

The aim of the dichotomous models is to explain the occurrence of the envisaged event as a function of a certain number of characteristics observed for the respondents (individuals in the sample). By applying these models, we seek to specify the probability of achieving the objective of our study (the conservation and/or development project).

Contingent valuation scenarios

The researcher plays a key role in creating a believable scenario for the CVM survey. This scenario should clearly describe:

The good (environmental resource): What is it? What benefits does it provide? What actions are needed to improve or conserve it?

Payment method: How will respondents' contributions be collected (mandatory fees, entrance fees, voluntary donations, etc.)? Who will benefit from these contributions ?

2.2. *Questionnaire design and administration*

2.2.1. *Developing the Questionnaire*

The CVM method requires a well-designed questionnaire to collect data on respondents' WTP. Our questionnaire addressed response formats, potential for non-responses, and potential biases. It included 15 questions.

2.2.2. *Targeted Sampling*

To enhance the credibility of our results and ensure participation, we employed targeted sampling. Since 60% of Feriana's population is employed (6,162 households), we focused on this group to target individuals likely to contribute financially. The survey was administered to 300 households and 50 Algerian tourists, with each household including at least one employed person.

2.2.3. *Scenario Presentation*

After socio-economic questions, respondents were presented with two options: Contribute annually to a fund via a tax or making voluntary quarterly donations. Before choosing a payment method, respondents selected one of two projects: *Maintain a forested area and develop a natural leisure park (with images provided). *Preserve nature in the Bouchebka mountains without development.

2.2.4. *Information collected*

The questionnaire gathered four types of important information for understanding factors influencing WTP decisions:

Site Use: Respondents' frequency of visiting or using the site.

Resource Assessment: Their perception of the natural resources' value.

Project Opinions: Their views on the proposed programs.

Sociodemographic Data: Information like age, education, and income.

This data helped us propose payment amounts relevant to respondents' perceptions.

Tourists: The same questionnaire was presented to Algerian tourists to gauge their interest in contributing to a Bouchebka forest leisure park project.

Payment Preferences: For residents, the goal

was to determine their preference between the two scenarios (conservation vs. development) and calculate an average "Consent to Pay" (CTP) value. For tourists, the CTP represented a potential park entrance fee.

Questionnaire Approach: A closed-ended survey design with dichotomous response options was employed (accept/reject) regarding offered payment amounts.

2.2.5. *Additional Information*

The questionnaire aimed to capture the following:

Resident Behavior: Understanding how residents interact with the ecological services provided by the natural environment.

Willingness to Pay: Assessing residents' WTP for preserving or enhancing the region's economic/environmental value.

Human-Environment Connection: Exploring the link between respondents and their natural surroundings.

Socio-economic data (gender, age, marital status, education, income) was also collected to support further analysis.

2.3. *Econometric models: Probit and Logit models*

The econometric treatment of willingness-to-pay is an important step in contingent valuation studies. Probit and Logit models are used to explain the values of Y through X, i.e. to estimate the probability that Yi = 1 knowing Xi. Note that:

$$\Pr (Yi = 1/Xi) = \Pr (Xi\,\theta + \varepsilon i \geq 0 / Xi) = \\ = \Pr (Xi\,\theta \geq - \varepsilon i / Xi) = F\text{-}\varepsilon (Xi\,\theta).$$

The only difference between the Probit and Logit models is the specification of F. In both cases, the distribution of residuals is symmetrical, so F-ε can be replaced by Fε.

The Probit model corresponds to the Gaussian specification introduced in the previous section. F is therefore the distribution function of a center-reduced Gaussian, usually denoted Φ:

$$F(Xi\theta) = \Phi (Xi\theta) = \int Xi\ \theta\ [(e\text{-}t^{2}\ /2) / (\sqrt{2\pi})]\ x\ 2t$$

The corresponding density, usually denoted ϕ, is:

$$f (Xi\theta) = \phi (Xi\theta) = e - (xi\ \theta)^{2}/2/\sqrt{2\pi}$$

There is virtually no difference between these two laws, the introduction of the logistic law being simply motivated by its simplicity in this framework (Crépet *et al*., 2009).

Hanemann's (1984) framework for discrete choice models based on a random utility function is adopted to implement the Logit model in this research, with the aim of converting responses (yes/no) to the question of whether to pay consent.

The respondent's utility function is given by u (j, y, x), which breaks down into two parts:

$$u\ (j, y, x) = n\ (j, y, x) + \text{ej avec } n\ (j, y, x)$$

It is important to consider that:

- ej: is the random part of the utility function, representing an unobservable component.
- y: represents the respondent's income.
- x: represents the individual characteristics on which the utility function depends.
- j: indicator variable equal to 1 if the respondent agrees to pay the proposed value and 0 if not.

The respondent agrees to pay an amount M, to conserve and/or improve a natural environment and preserve the region's natural heritage if:

$$[n\ (1, y - M, x) + e1] > [n\ (0, y, x) + e0].$$

Consequently, the respondent will agree to pay if the well-being is achieved or the condition of the good is improved and will refuse to pay the amount if the opposite is the case.

The formula for the probability of agreeing to pay is as follows:

$$\text{Prob} = \text{Prob}\ [n\ (1, y - M, x) + e1 > n\ (0, y, x) + e0]$$

The following form can be used:

$$\text{Prob} = \text{Prob}\ [n\ (1, y - M, x) - n\ (0, y, x) > e0 - e1]$$

Where:

$\eta = (e0 - e1)$ and $\Delta n = n\ (1, y - M, x) - n\ (0, y, x)$

The answer is then a random variable composed of a probability density.

The cumulative distribution function ($F\eta\ (\Delta n)$) can be thought of as the probability of accepting the amount M either:

$$\text{Prob (accept M)} = Pi = F\eta\ (\Delta n) = 1 - G\ \eta\ (\Delta n)$$

- Pi: the probability that respondent i agrees to pay the proposed amount and

- G η (Δ n) the complement: the probability of refusing this amount.

At this point, the question arises of the functional form of the individual utility function and the distribution function, for which specific models are used: the logit model and the linear regression model. In the logit model, it is a distribution function of a logistic variable, while in the linear regression model, Fη (Δ n) represents the distribution function of a normal distribution (Maddala et Lahiri, 1992).

According to Abichou and Zaibet (2008), two features make the logistic function interesting for modeling discrete choices: its interval is limited to 0 - 1. The logit function can be used as a probabilistic function, offering the possibility of being linearized by logarithmic transformation. If Δ n follows a logistic distribution, then the cumulative distribution function is given by (Maddala, 1983):

$$Pi=F\eta\ (\Delta n) = \exp \Delta n / (1 + \exp \Delta n)$$

This leads to the following equation:

$$\log [F\eta\ (\Delta n) / (1- F\eta(\Delta n)]= \Delta n$$

The logit model is thus written as follows:

$$\text{logit}= \log (Pi / (1- Pi)) = \Delta n = WTP$$

3. Results

This section presents the findings from our survey of 300 households in the Feriana region and 50 Algerian tourists.

3.1. *Socioeconomic characteristics*

The majority of Feriana households were male-headed (92%) with young adults under 50 years old comprising 71.3% of respondents. Most heads of household had a high level of education (26% with higher education) and the average household size was 5 people (58% with more than 5 members). Dominant professions included agriculture, construction, and other manual labor (82%), with 34% earning between 200-400 €. Regarding the environment, 100% of respondents were aware of the surrounding forests. The primary activities reported were collecting forest products (56%) and grazing (37%). Despite proximity, 40% of respondents had not visited the mountainous areas.

Among the 50 Algerian passengers surveyed, 84% were men, 60% were under 35, and 34% had higher education. Most traveled in groups of two (58%) and stayed less than five days in the region (64%). The main reason for short stays was the lack of high-quality leisure facilities and services, with 75% viewing Feriana as a transit area.

3.2. *Willingness to pay propensity*

This study investigates the relationship between socioeconomic characteristics and residents' payment preferences towards a potential project, hypothesizing that these characteristics influence financial support decisions. Our results showed that 42% of respondents favored a leisure project, while 24% preferred a nature conservation project. Overall, 67% of respondents agreed to contribute financially to their chosen project.

Analyzing the independent variables revealed that income level, education level, household size, and prior forest experience influenced willingness to pay. Individuals with higher incomes and education were more likely to contribute, while willingness to pay decreased with household size. Those who had visited the forest environment were also more inclined to contribute financially. Interestingly, the activity undertaken in the forest mattered, with walkers demonstrating a greater willingness to pay compared to hunters. Finally, indecisiveness regarding project choice before payment requests was linked to a higher refusal rate (55%). These findings support the hypothesis that socio-economic characteristics play a significant role in financial contribution decisions, validating the initial hypothesis. Furthermore, the analysis of factors influencing WTP aligns with respondents' declarations, contributing to the economic evaluation's theoretical validity.

3.3. *Revealed Willingness to Pay*

The survey revealed a total annual WTP of € 6,595 for Feriana residents, with an average of € 32.7 per person. Participants favored both the leisure park development and natural environment conservation projects. While the development

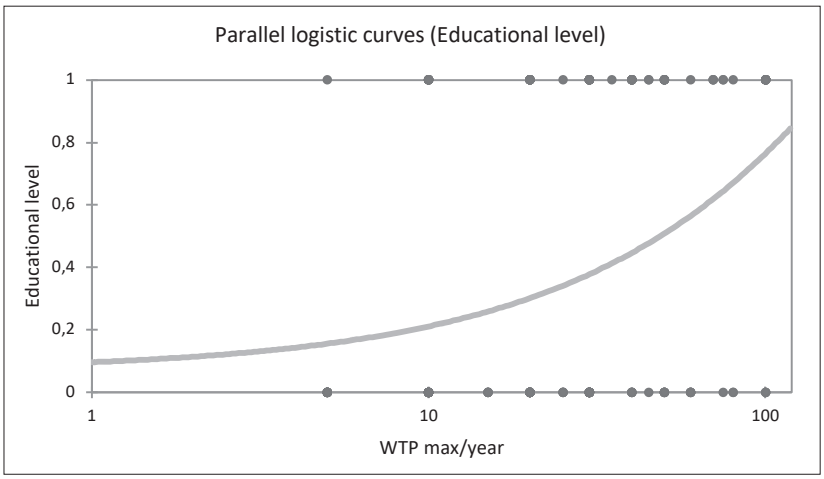


Figure 2 - Effect of the 'educational level' variable on the value of WTP.

value was higher than the conservation value, the difference was not statistically significant. Residents were offered two payment methods (tax or voluntary contribution), which may influence the final average CTP amount. Those opting for voluntary payments represented 63.1% of the sample, but their individual contributions were less than half those who chose the annual tax. Income level significantly influenced the amount paid, with a clear increase observed across income brackets. Individuals in the highest income class (≥€ 800,000/month) contributed more than double the average value of other income classes.

For Algerian tourists, WTP for park entrance fees totaled € 272,500 for 50 visitors, translating to an average of € 5,450 per person. This value may vary depending on socio-economic characteristics. Our findings suggest a trend where older populations are less likely to respond positively, while younger individuals (under 50) are more inclined to contribute. The influence of education and age on WTP is also evident. Educated respondents were more likely to accept the WTP, with those having university degrees willing to pay higher amounts to visit a nature park. Conversely, individuals without higher education were less likely to accept the WTP, with some refusing a 25 € contribution entirely, particularly those with primary or secondary schooling. University-educated respondents demonstrated a strong willingness to pay, with 93% prepared to contribute € 10 or more, including 45% willing to pay € 25 per visit. Only 7% in this category refused to participate. Age was negatively correlated with WTP value, indicating that older individuals were less willing to pay higher amounts. In fact, 60% of those willing to pay were under 35 years old.

3.4. *Logistic regression*

3.4.1. *Residents' Willingness to Pay*

Building on the initial descriptive analysis, this study aims to identify the variables influencing residents' WTP for the upkeep of their natural environment, particularly the surrounding forest in the Feriana region. Simple descriptive analysis proved to be insufficient for this purpose. Therefore, we employed logistic regression to define a Logit model and gain a deeper understanding of these variables.

The results revealed an average WTP of € 32.6 per household, with an average household composition of 4.8 people and 2.1 children. Notably, residents showed a preference for voluntary bi-monthly payments. Most household heads fell into the low- to medium-income category, with over 60% earning less than € 600 per month. Interestingly, the logistic regression analysis indicated that most respondents favored the development of a leisure park in the surrounding natural areas, with an average preference score of 1.9.

Educational Level: Figure 2 reveals a clear trend: individuals with higher education levels (approaching a value of 1) tend to have a higher maximum annual willingness to pay (WTP). This positive correlation is statistically significant ($R^2 = 0.458$).

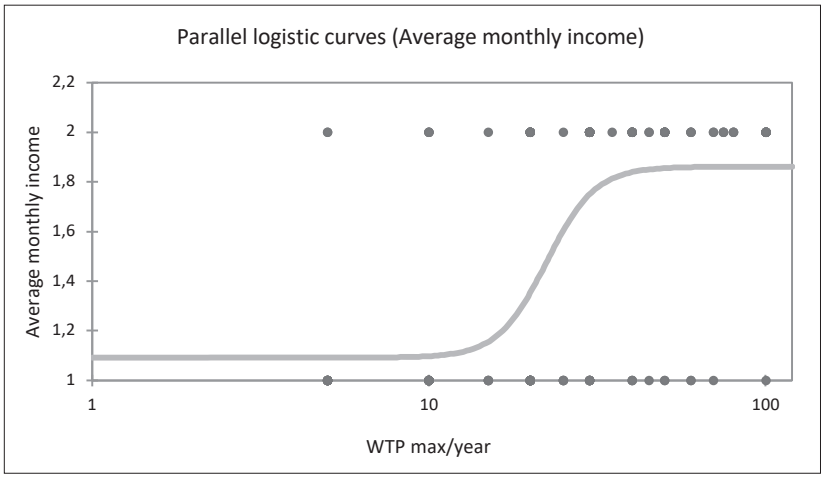


Figure 3 - Effect of the 'average monthly income' variable on the value of WTP.

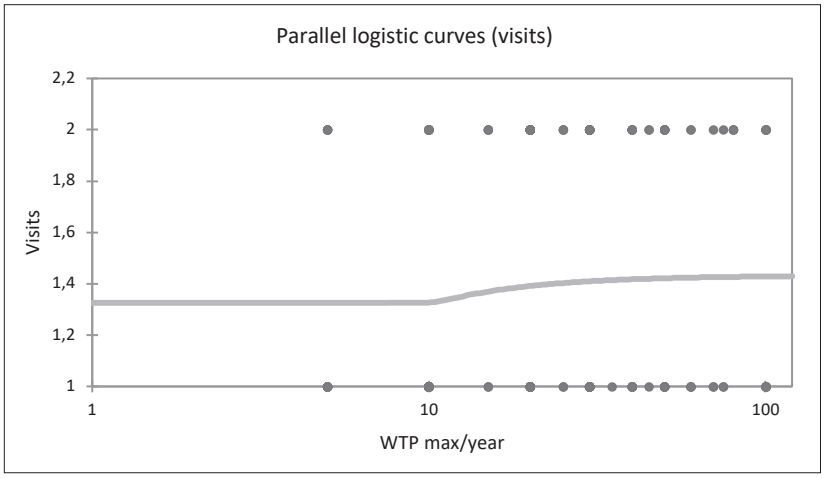


Figure 4 - Effect of the visit variable on the WTP value.

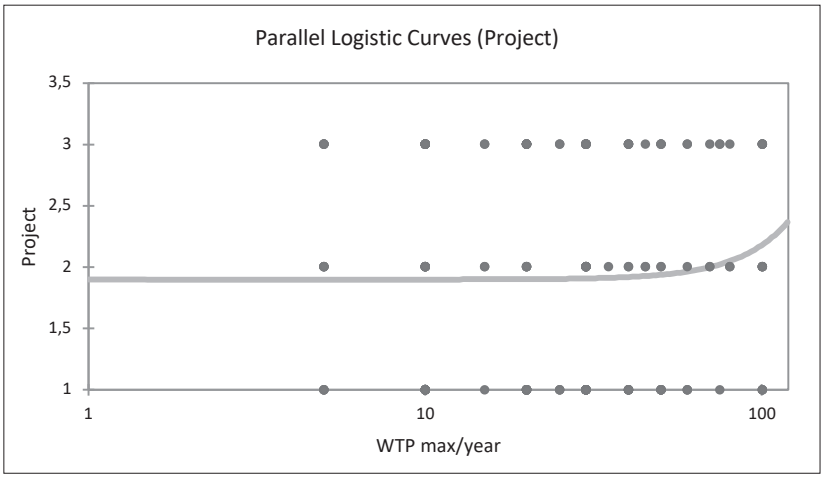


Figure 5 - Effect of the variable 'project type' on the value of WTP.

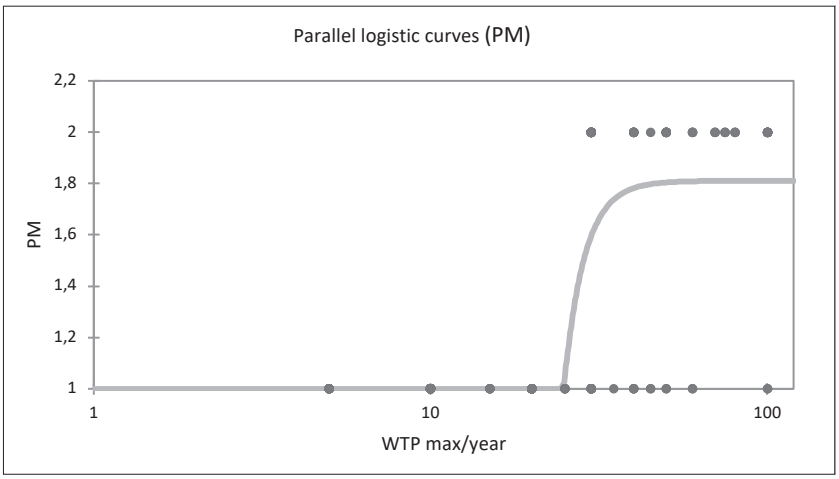


Figure 6 - Effect of the 'Payment method' variable on the WTP value.

Income: Figure 3 suggests a positive but weak relationship between WTP and average monthly income. As income increases towards a value of 2 (indicating income above €600), there's a slight increase in WTP. However, the R^2 value of 0.028 indicates this correlation is not very strong.

Frequency of Visits: Similar to education level, Figure 4 shows a positive correlation between visiting the natural areas and WTP. People who visit more frequently (visit variable closer to 1) are more likely to express a higher WTP. However, the R^2 of 0.005 suggests this effect is weakly significant.

Project Type: Development appears to be a stronger incentive for higher WTP compared to conservation (Figure 5). This relationship is also weakly significant ($R^2 = 0.005$).

Payment Method: The payment method has the most significant impact on WTP ($R^2 = 0.639$). Figure 6 demonstrates that WTP is consistently above €10 and reaches its maximum (€100) when the payment mode tends towards 2 (annual tax).

Other Variables: While most variables showed positive correlations with WTP, Figure 7 indicates that educational level and household size might be negatively correlated. In logistic regression, a positive coefficient for a dichotomous variable means those with modality 1 (e.g., visitors) are more likely to pay than those with modality 0 (non-visitors). The opposite applies to negative coefficients. For continuous variables (like income), a positive coefficient suggests that WTP increases as the variable increases (Saadaoui, 2013).

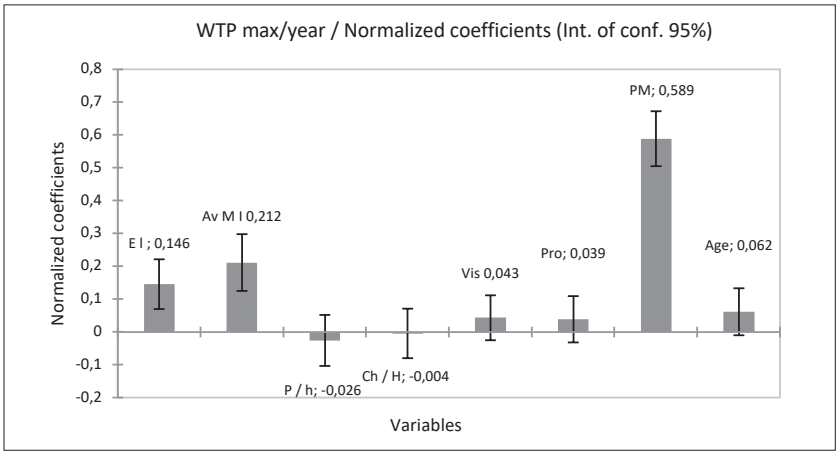


Figure 7 - WTP/Normalized coefficient.

Table 1 - Correlation between descriptive variables and WTP.

Variables	*Educ Level*	*Monthly Income*	*Person / House*	*Child / House*	*Visits*	*Project*	*Payment Mode*	*Age*
WTP max/ Year	0,439	0,612	-0,128	-0,100	0,055	0,063	0,769	0,101

Source: XLSTAT 2023 statistical analyzes based on survey data.

Table 2 - Correlation matrix between the explanatory variables and WTP.

Variable	*Gender*	*Age*	*Edu Level*	*Groupe size*	*Length stay*
WTP	-0,115	-0,137	0,571	0,099	0,577

Source: XLSTAT 2023 statistical analyzes based on survey data.

Analysis of the correlation between descriptive variables and WTP revealed positive correlations for factors including education level, average monthly income, frequency of visits to the natural area, project type (development vs. conservation), preferred payment method, and respondent age. The payment method variable displayed a particularly strong positive correlation with WTP.

Furthermore, Table 1 indicates that average monthly income and preferred payment method have the most significant effects on the monetary value of WTP among the explanatory variables examined.

3.4.2. *Tourists' Willingness to Pay*

The second part of the analysis examined tourists' and foreign passengers' WTP for access to a potential natural leisure park. The average WTP estimated by the logistic regression model was € 5450, but this masked a wider range of individual responses. While around 5% of respondents refused to pay entirely, the majority (46%) were willing to pay a more modest entrance fee of € 2.50. An additional 22% and 10% were prepared to pay € 5 and € 20 respectively.

Table 2 (in the referenced table) details the correlations between explanatory variables and the dependent variable, WTP Interestingly, gender and age were negatively correlated with WTP, suggesting a potential preference for the park among younger visitors and those traveling in larger groups. The analysis (Figure 8) also revealed positive correlations between:

Group Size: Larger groups were more likely to pay to visit the park.

Education Level: Higher education was associated with a greater WTP.

Length of Stay: Tourists staying longer in the region were more likely to pay for park access.

These findings suggest that targeting park marketing and development towards younger visitors, larger travel groups, and those with higher education levels could be an effective strategy.

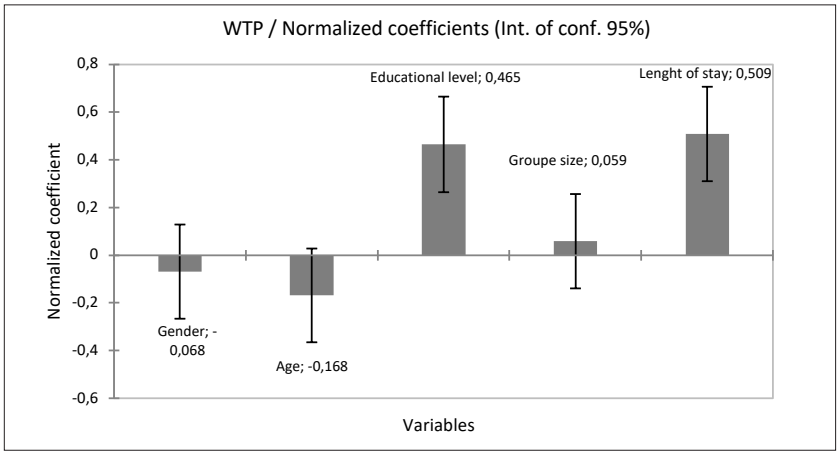


Figure 8 - Descriptive variable correlation /WTP.

Table 3 - Normalized Coefficients (WTP).

Logit estimates			*Observations = 50*		
Source	*Value*	*t*	*Pr > |t|*	*Lower Bound (95%)*	*Upper Bound (95%)*
Gender	-0,068	-0,698	0,489	-0,266	0,129
Age	-0,168	-1,723	0,092	-0,365	0,029
Educational level	0,465	4,673	*< 0,0001*	0,264	0,665
Family members	0,059	0,602	0,550	-0,139	0,257
Length of stay	0,509	5,176	*< 0,0001*	0,311	0,707

Educational level and length of stay drive variability in WTP

Table 3 (refer to table for details) highlights that educational level and length of stay are the most significant explanatory variables influencing the variability of tourists' WTP for park access. In other words, these factors have the strongest impact on how much tourists are willing to pay to visit the park.

4. Discussion

The study investigated stakeholder's preferences for the conservation and development of the Tunisian Aleppo pine forests. Both proposed projects, creating a natural leisure park or maintaining a forested area, aimed to revitalize and enhance the landscape. By comparing these options, the trade-offs between development and preservation were explored, and the economic value residents placed on each approach.

The findings revealed a clear preference for the development project, with residents expressing a higher WTP for a park compared to pure conservation efforts. This suggests that residents value the potential economic and social benefits associated with a leisure park, such as increased tourism and recreational opportunities.

Furthermore, the analysis of payment methods yielded interesting insights. Residents showed a stronger preference for voluntary contributions over annual taxes. This might indicate a desire for more control over their financial contributions or a perception that a tax wouldn't directly benefit the park's development.

The significant influence of education level on WTP is noteworthy. Individuals with higher education levels displayed a greater propensity to pay for both conservation and development projects. This potentially reflects a heightened awareness of the environmental benefits associated with protecting natural areas.

Length of stay also emerged as a significant factor for tourists' WTP. Tourists planning longer stays were more willing to pay for park access, suggesting a greater perceived value from the park experience. These findings can inform marketing strategies by targeting park development and promotion towards younger visitors, larger travel groups, and those with higher education levels who plan extended stays in the region.

Overall, this study sheds light on resident preferences and the potential economic viability of a natural leisure park in the Feriana region. The results highlight the importance of considering both conservation and development objectives, along with resident and tourist perspectives, when formulating sustainable environmental management plans.

Future research could explore the specific features and activities most desired by residents and tourists within a potential park to further optimize its development and economic potential.

A similar study was conducted by Daly-Hassan et *al*. (2015) to evaluate cork oak forests in Tunisia for the purpose of selecting adaptation options for these forests to climate change. Their findings revealed comparable results regarding the importance of the values of goods and services provided for different uses.

Daly-Hassan and Ben Mansoura (2005) recognized the need for changes in forest legislation to improve management flexibility, product trade, and overall forest health. This included

incorporating stakeholders, whose active participation is crucial for sustainable forest development. Even today, after two decades, these changes remain a priority, almost an obligation, to conserve this national treasure and protect the rights of future generations.

5. Conclusion

This study employed the CVM to assess the value of the Aleppo pine forests of central Tunisia, considering both conservation and development. Residents were presented with scenarios involving either the preservation of non-harvestable forests or the development of a natural leisure park. The CVM approach allowed us to estimate the economic value that both residents and tourists placed on these natural areas.

Logistic regression analysis provided valuable insights into the factors influencing residents' WTP for environmental improvements. Two key variables emerged as significantly positive influences: monthly income and preferred payment method. Residents with higher incomes demonstrated a greater capacity to contribute financially, while voluntary payments were generally preferred over annual taxes. This highlights the importance of considering affordability and payment flexibility when designing conservation or development initiatives.

The findings demonstrate the potential for reconciliation between economic development and ecological sustainability. Residents expressed a clear preference for the development project, suggesting a perceived value associated with the potential economic and social benefits of a leisure park. However, the positive correlations between WTP and educational level for both conservation and development projects underscore the importance of environmental protection for residents. This highlights the need for future projects to strike a balance between development and conservation, ensuring economic growth while maintaining the ecological value of the natural areas.

This study contributes to a limited body of research within our region, particularly in comparison to the dominance of North American studies in the EVRI database. These findings offer valuable insights applicable to other developing countries grappling with similar challenges of balancing economic development with environmental protection. Active resident engagement in decision-making through valuation methods like CVM can lead to sustainable solutions that benefit both the environment and local communities. Future research could delve deeper into specific park features and activities desired by residents and tourists to further optimize park development and its economic potential, ensuring a future where a thriving natural environment coexists with a prosperous local economy.

References

Abichou H., Zaibet L., 2008. Evaluation de la valeur récréative du parc Ennahli (Tunis). *New Medit*, 7(4): 54.

Achour H., Toujani A., Rzigui T., Faïz S., 2018. Forest cover in Tunisia before and after the 2011 Tunisian revolution: A spatial analysis approach. *Journal of Geovisualization and Spatial Analysis*, 2: 1-14.

Bateman I.J., Turner R.K., 1992. *Evaluation of the environment: the contingent valuation method* (Vol. 92). Norwich, UK: Centre for Social and Economic Research on the Global Environment.

Baymuminova N., Shermukhammedova G., Choi J.G., 2023. Estimating the Economic Value of Ichan Kala Using the Contingent Valuation Method (CVM). *Sustainability*, 15(3): 2631.

Blasi E., Rossi E.S., Zabala J.Á., Fosci L., Sorrentino A., 2023. Are citizens willing to pay for the ecosystem services supported by Common Agricultural Policy? A non-market valuation by choice experiment. *Science of The Total Environment*, 893: 164783.

Brockerhoff E.G., Barbaro L., Castagneyrol B., Forrester D.I., Gardiner B., González-Olabarria J.R., Lyver P.O'B., Meurisse N., Oxbrough A., Taki H., Thompson I.D., van der Plas F., Jactel H., 2017. Forest biodiversity, ecosystem functioning and the provision of ecosystem services. *Biodiversity and Conservation*, 26: 3005-3035.

Campos P., Daly-Hassen H., Oviedo J.L., Ovando P., Chebil A., 2008. Accounting for single and aggregated forest incomes: Application to public cork oak forests in Jerez (Spain) and Iteimia (Tunisia). *Ecological Economics*, 65(1): 76-86.

Cenamor J., Frishammar J., 2021. Openness in platform ecosystems: Innovation strategies for complementary products. *Research Policy*, 50(1): 104148.

Chazdon R.L., Brancalion P.H., Laestadius L., Bennett-Curry A., Buckingham K., Kumar C., Moll-Rocek J., Guimarães Vieira I.C., Wilson S.J., 2016. When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration. *Ambio*, 45(5): 538-550.

Crépet A., Harari-Kermadec H., Tressou J., 2009. Using empirical likelihood to combine data: application to food risk assessment. *Biometrics*, 65(1): 257-266.

Cuccia T., 2020. Contingent valuation. In: *Handbook of Cultural Economics*, *3rd ed.* Cheltenham: Edward Elgar Publishing, pp. 95-105.

Daly-Hassen H., Croitoru L., 2013. Évaluation économique des biens et services des forêts tunisiennes. *Forêt méditerranéenne*, 34(4): 299-304.

Daly-Hassen H., Gader G., Potthast M.C., 2015. *Une approche économique pour choisir les options d'adaptation des forêts au changement climatique: Application à la forêt de chêne liège en Tunisie*. Proceedings of the XIVème Congrès Forestier Mondial, Durban, South Africa, pp. 7-11.

Daly-Hassen H., Mansoura A.B., 2005. Tunisia. In *Valuing mediterranean forests: towards total economic value*. Wallingford, UK: CABI Publishing, pp. 105-122.

Daly-Hassen H., Riera P., Mavsar R., Gammoudi A., Garcia D., 2017. Valuing trade-offs between local forest uses and environmental services in Tunisia. *Journal of Environmental Economics and Policy*, 6(3): 268-282.

d'Arge R.C., Shogren J.F., 1989. *Non-market asset prices: A comparison of three valuation approaches*. Studies in Environmental Science, Vol. 36. Amsterdam: Elsevier, pp. 15-36.

Davidson M.D., 2013. On the relation between ecosystem services, intrinsic value, existence value and economic valuation. *Ecological economics*, 95: 171-177.

del Saz-Salazar S., Guaita-Pradas I., 2013. On the value of drovers' routes as environmental assets: A contingent valuation approach. *Land Use Policy*, 32: 78-88.

Dobbs C., Escobedo F.J., Zipperer W.C., 2011. A framework for developing urban forest ecosystem services and goods indicators. *Landscape and urban planning*, 99(3-4): 196-206.

Dupras J., Laurent-Lucchetti J., Revéret J.P., DaSilva L., 2018. Using contingent valuation and choice experiment to value the impacts of agri-environmental practices on landscapes aesthetics. *Landscape Research*, 43(5): 679-695.

Durden G., Shogren J.F., 1988. Valuing Non-market Recreation Goods: An evaluative survey of the literature on the travel cost and contingent valuation methods. *Review of Regional Studies*, 18(3): 1-15.

Farley J., Costanza R., 2010. Payments for ecosystem services: from local to global. *Ecological economics*, 69(11): 2060-2068.

Grazhdani D., 2024. Results of Two Non-Market Valuation Methods Used to Estimate Recreational Fishing in the Lakes Prespa Watershed. *Journal of Environmental Management and Tourism*, 15(1): 52-68.

Guijarro F., Tsinaslanidis P., 2020. Analysis of academic literature on environmental valuation. *International journal of environmental research and public health*, 17(7): 2386.

Haines-Young R., Potschin M., 2010. The links between biodiversity, ecosystem services and human well-being. *Ecosystem Ecology: a new synthesis*, 1: 110-139.

Hammami S.E., Mimoun A.B., Haha M.N.B., 2023. Evaluation économique des services écosystémiques et du coût de non action en Tunisie: Cas du bassin versant O. Lachbal. *New Medit*, 22(4): 109-128.

Hanemann W.M., 1984. Welfare evaluations in contingent valuation experiments with discrete responses. *American journal of agricultural economics*, 66(3): 332-341.

Hasnaoui A., Krott M., 2019. Forest governance and the Arab spring: A case study of state forests in Tunisia. *Forest policy and economics*, 105: 99-111.

Islem S., Hayet I., Christopher R.B., Hichem R., 2014. Contribution to the study of the flora in the central-west of Tunisia: landscape dynamics and evaluation of plant biodiversity of mountain Bouchebka. *International Journal of Innovation and Applied Studies*, 6(2): 257.

Jdaidi N., Selmi H., Aloui F., Jedidi S., Chaabane A., 2023. Évaluation des facteurs de menace et de vulnérabilité potentielles des plantes médicinales et aromatiques au nord-ouest tunisien. *Revue Marocaine des Sciences Agronomiques et Vétérinaires*, 11(1): 14-21.

Jones C.A., 1997. Use of non-market valuation methods in the courtroom: Recent affirmative precedents in natural resource damage assessments. *Water Resources Update*, 109(1): 10-18.

Khalfaoui M., Daly-Hassen H., Stiti B., Jebari S., 2020. Toward decision-making support: Valuation and mapping of new management scenarios for Tunisian Cork oak forests. *Forests*, 11(2): 197.

Krause M.S., Nkonya E., Griess V.C., 2017. An economic valuation of ecosystem services based on perceptions of rural Ethiopian communities. *Ecosystem services*, 26: 37-44.

Maddala G.S., 1983. Methods of estimation for models of markets with bounded price variation. *International Economic Review*, 361-378.

Maddala G.S., Lahiri K., 1992. *Introduction to econometrics* (Vol. 2, p. 525). New York: Macmillan.

Maddala G.S., Wu S., 2000. Cross-country growth regressions: problems of heterogeneity, stability and interpretation. *Applied Economics*, 32(5): 635-642.

Mandle L., Shields-Estrada A., Chaplin-Kramer R., Mitchell M.G., Bremer L.L., Hawthorne P., Johnson J.A., Robinson B.E., Smith J.R., Sonter L.J., Verutes G.M., Vogl A.L., Daily G.C., Ricketts T.H., 2021. Increasing decision relevance of ecosystem service science. *Nature Sustainability*, 4(2): 161-169.

Manero A., Nikolakis W., Woods K., Grafton R.Q., 2024. Non-market valuation and Indigenous Peoples' values: researcher constraints and policy impacts. *Environmental Science & Policy*, 153: 103679.

Manero A., Taylor K., Nikolakis W., Adamowicz W., Marshall V., Spencer-Cotton A., Nguyen M., 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.

Martín-López B., Montes C., Benayas J., 2007. The non-economic motives behind the willingness to pay for biodiversity conservation. *Biological conservation*, 139(1-2): 67-82.

Milne M.J., 1991. Accounting, environmental resource values, and non-market valuation techniques for environmental resources: A review. *Accounting, Auditing & Accountability Journal*, 4(3): 81-109.

Mori A.S., Lertzman K.P., Gustafsson L., 2017. Biodiversity and ecosystem services in forest ecosystems: a research agenda for applied forest ecology. *Journal of Applied Ecology*, 54(1): 12-27.

Oglethorpe D.R., Miliadou D., 2000. Economic valuation of the non-use attributes of a wetland: A case-study for Lake Kerkini. *Journal of Environmental Planning and Management*, 43(6): 755-767.

Oueslati W., Madariaga N., Salanié J., 2008. Évaluation contingente d'aménités paysagères liées à un espace vert urbain. Une application au cas du parc Balzac de la ville d'Angers. *Revue d'Etudes en Agriculture et Environnement-Review of agricultural and environmental studies*, 87: 77-99.

Ouyang Z., Song C., Zheng H., Polasky S., Xiao Y., Bateman I.J., Liu J., Ruckelshaus M., Shi F., Xiao Y., Xu W., Zou Z., Daily G.C., 2020. Using gross ecosystem product (GEP) to value nature in decision making. *Proceedings of the National Academy of Sciences*, 117(25): 14593-14601.

Perni Á., Barreiro-Hurlé J., Martínez-Paz J.M., 2021. Contingent valuation estimates for environmental goods: Validity and reliability. *Ecological Economics*, 189: 107144.

Raihan A., 2023. A review on the integrative approach for economic valuation of forest ecosystem services. *Journal of Environmental Science and Economics*, 2(3): 1-18.

Randall A., Hoehn J.P., Brookshire D.S., 1983. Contingent valuation surveys for evaluating environmental assets. *Natural Resources Journal*, 23(3): 635-648.

Rogers A.A., Dempster F.L., Hawkins J.I., Johnston R.J., Boxall P.C., Rolfe J., Kragt M.E., Burton M.P., Pannell D.J., 2019. Valuing non-market economic impacts from natural hazards. *Natural Hazards*, 99: 1131-1161.

Saadaoui I., 2016. Biosurveillance de la dynamique de la flore dans la forêt transfrontalière du Centre-Ouest tunisien: le cas des monts Bouchebka. In: *I International Conference on Research for Sustainable Development in Mountain Regions/Mountains 2016*.

Saadaoui I., Bryant C.R., Rejeb H., Petrisor A.I., 2018. Biodiversity conservation and strategies of public awareness. Case study: the natural landscapes of central Tunisia. *Present Environment and Sustainable Development*, 12: 263-278.

Spangenberg J.H., Settele J., 2010. Precisely incorrect? Monetising the value of ecosystem services. *Ecological complexity*, 7(3): 327-337.

Taghouti I., Ouertani E., Guesmi B., 2021. The contribution of non-wood forest products to rural livelihoods in Tunisia: the case of Aleppo pine. *Forests*, 12(12): 1793.

Tian Y., Wu H., Zhang G., Wang L., Zheng D., Li S., 2020. Perceptions of ecosystem services, disservices and willingness-to-pay for urban green space conservation. *Journal of Environmental Management*, 260: 110140.

Viti M., Löwe R., Sørup H.J., Rasmussen M., Arnbjerg-Nielsen K., McKnight U.S., 2022. Knowledge gaps and future research needs for assessing the non-market benefits of Nature-Based Solutions and Nature-Based Solution-like strategies. *Science of the Total Environment*, 841: 156636.

Exploring the determinants of consumers—adoption of online grocery shopping using the Theory of Planned Behavior: An empirical study

Marianthi Giannakopoulou*, Stavroula Ioannou*, Andreas Stylianou*, Georgios Adamides*

DOI: 10.30682/nm2404b
JEL codes: C92, Q13

Abstract

Consumer adoption of online grocery shopping has been gradually increasing over the past two decades. This article examines the behavioral model of consumers and the factors that influence adoption of online grocery shopping of the typical adopter profile in Cyprus. Using data from 302 respondents of an online survey, it was found that 45.7% have purchased groceries online. Our research has shown that perceived behavioral control carries more weight than intention in explaining online grocery shopping behavior, and that two antecedent constructs —attitude and subjective norms— have direct influence on intention for online grocery shopping. From a managerial perspective, these findings provide insights to retailers, marketers and policymakers in enhancing the knowledge of this emerging online service. From a theoretical point of view, this study contributes to the limited existing research in the Euro-Mediterranean region, linking consumer behavior to the adoption of online grocery shopping.

Keywords*: Online grocery shopping, Consumer adoption, Consumer behavior, Theory of Planned Behavior.*

1. Introduction

Information technology has changed the way people communicate, socialize, learn, and consequently even how they do shopping (Alaimo *et al*., 2022). The global food and grocery retail market was valued approximately at USD 12.29 trillion in 2020, and is projected to be worth USD 17.29 trillion by 2027 (Research and Markets, 2022). In a survey of 30,000 respondents from 63 countries, carried out in 2016 by the Nielsen Group, showed that 24% have purchased groceries online, specifically packaged grocery food, and 21% have purchased fresh groceries (Nielsen, 2017).

Technological advancement has become an indispensable part of life. According to recent reports, globally 5.3 billion people use the internet, that is the equivalent of 65.4% of the world's total population (Shewale, 2023). Moreover, there are 7.4 billion mobile-phone subscriptions (Nielsen, 2017), hence one might describe today's consumers as more "connected" than ever before. According to the Cyprus

* Agricultural Research Institute, Nicosia, Cyprus.
Corresponding author: gadamides@ari.moa.gov.cy

Statistical Service, in 2021 Internet access in households reached 93.4% compared to 92.8% in 2020 (CYStat, 2021). The Cyprus Statistical Service also reports that 80% of households have access to a personal computer (desktop, laptop or tablet) and a computer is found in almost all (95.4%) households with dependent children. Furthermore, 22.7% of all enterprises stated that due to the Covid-19 pandemic the enterprises initiated or increased their efforts to sell goods or services via the Internet (websites, apps).

Online grocery shopping (OGS) is an alternative way of buying food and other household necessities using a web-based shopping service, instead of the traditional in-store grocery shopping. Online grocery shopping is been promoted as an alternative, additional way to alleviate issues to food access faced by the elderly, disabled, residents in areas with limited access to grocery shops (Rogus *et al*., 2020). In a previous study, carried out in Cyprus in 2006, it was found that only 14% of the respondents had purchased groceries online, indicating that Cypriot consumers were quite unfamiliar with OGS (Adamides *et al*., 2006). However, findings from the current survey have shown that 45.7% of the respondents have bought groceries online and in fact 72.6% of them agree that purchasing groceries online is an easy/straightforward process. Hence, Cypriot consumers are now much more familiar with online grocery shopping than 15 years ago.

The aim of this study is to investigate consumers' adoption of online grocery shopping and the behavioral model of Cypriot consumers, along with their demographic and household characteristics. This study started in 2021 (Summer-Autumn), during the Covid-19 pandemic, where several measures were in place (since 2020) to limit the spread of the virus (e.g., restriction of free movement). In the context of a pandemic, online grocery shopping can offer interesting benefits to consumers, including social distancing, home delivery and time savings. The associated measures introduced in Cyprus, and worldwide, caused several severe economic and societal disturbances, leading to an expansion of e-commerce in general, and accelerating its adoption. Subsequently, the online grocery retail expanded during the pandemic (Dannenberg *et al*., 2020; Jensen *et al*., 2021; Jribi *et al*., 2020).

The Theory of Planned Behavior (TPB), developed by Ajzen (1991), argues that human actions and therefore human behavior are directly influenced by the individual's intentions to adopt a behavior as well as the actual control of the behavior. Essakkat *et al*. (2021) states that intentions are assumed to capture the motivational factors that affect their behavior; they show how hard people are willing to try and how much effort they are planning to exert to perform a behavior. As a general rule, the stronger the intention to engage in a behavior, the more likely should be its performance (Ajzen, 1991). According to the TPB, three conceptually independent factors determine Behavioral Intentions, viz. a) the Attitude towards the behavior, b) Subjective norm and c) Perceived Behavioral Control (PBC). More specifically, attitude towards a behavior refers to individual's positive or negative evaluation of performing a behavior. Subjective norms refer to the social pressure that an individual perceives from "important others", for instance his/her parents, friends, to perform a certain behavior. Perceived behavioral control, as described above, represents the perceived ease or difficulty of adopting a behavior and it is assumed to reflect individual's experience and the obstacles he/she encountered in respective previous behavior. In general, a person's intention to perform a behavior becomes stronger when: a) he/she positively evaluates the behavior in question, b) he/she believes that others important to him approve of his decision, and c) he/she feels in control of the factors influencing it. Essakkat *et al*. (2021), notes that the TPB distinguishes between behavioral intention and actual behavior. The TPB is one of the most utilized and discussed theoretical models for predicting intentions and behavior (Jensen *et al*., 2021; Sentosa & Mat, 2012; Strydom, 2018; Topa & Moriano, 2010). Each of the latent variables (Behavioral Intention, Attitude towards a Behavior, Subjective Norms and Perceived Control of Behavior) reveals a different aspect of an individual's behavior and can be a target for intervention. Therefore, the collection of the above information is deemed necessary and constitutes a good basis for investigating the perceptions of

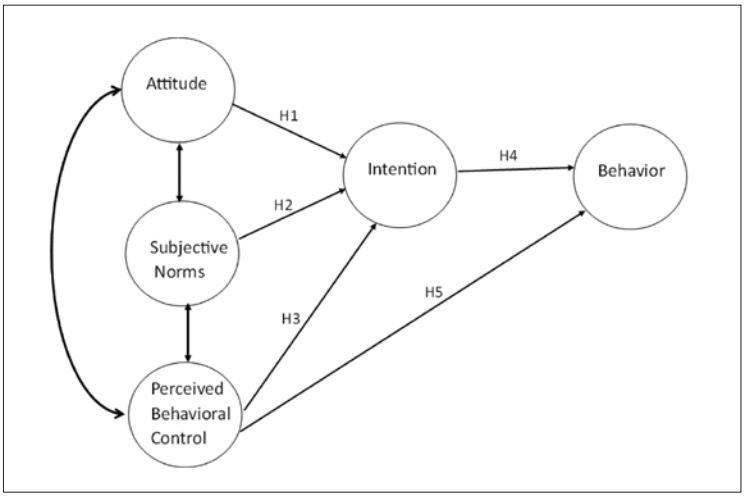


Figure 1 - The Theory of Planned Behavior (adapted from Ajzen, 1991), along with research hypotheses.

Cypriot consumers towards online grocery shopping. Figure 1 shows the theoretical framework of the TPB model as described above and adopted in this study. The hypotheses established in this study are as follows:

- Hypothesis 1 (H1). A positive attitude towards online grocery shopping has a positive and direct effect on the intention to OGS.
- Hypothesis 2 (H2). Subjective Norms, i.e., beliefs and expectations from immediate family and close friends towards OGS has a positive and direct effect on the intention to OGS.
- Hypothesis 3 (H3). OGS perceived behavioral control has a positive and direct effect on the intention to OGS.
- Hypothesis 4 (H4). Intention to OGS has a positive and direct effect on OGS behavior.
- Hypothesis 5 (H5). Perceived behavioral control has a positive and direct effect on OGS behavior.

The remainder of this paper is structured as follows: In the next section, a literature review on online grocery shopping and consumer characteristics is provided. Then, the methodological steps, data collection and analysis are described in detail. Next, the main results of the study are provided and discussed. The paper provides new empirical evidence on the understanding of the determinants of online grocery shopping in Cyprus and concludes with suggestions for future work.

2. Literature review

A bibliographical survey in Scopus for the period 2000 until 2023 resulted merely in 21 articles related to the query ("theory of planned behavior" OR "TPB" AND "online grocery"). The analysis of the search results shows a steady increase in the number of published studies on the subject of online grocery shopping, especially since 2018. Most studies were carried out in Denmark, followed by Brazil, Canada, Germany, the United Kingdom and the United States. In this sense, this study provides additional, more recent information for the Euro-Mediterranean region, using as a case study the Republic of Cyprus.

Etumnu *et al*. (2019), examined the drivers of online grocery shopping in the United States and explored linkages between consumer demographics and their perceptions towards OGS. They hypothesized that specific socio-demographics (e.g., age, children in the household) affect the tendency to shop groceries online, and that consumers who do OGS have favorable perceptions about OGS compared to in-store grocery shopping (e.g., convenience). They concluded that younger people, irrespective of their gender, residency, and their previous online shopping experience, are more likely to shop groceries online.

Van Droogenbroeck and Van Hove (2017), using logistic regression analysis, examined the

impact of both personal and sociodemographic variables of consumers in Belgium and the effects on the adoption of online grocery shopping. They found that the higher the educational level of the respondents, the more likely were to adopt OGS, while gender was non-significant. They concluded with suggestions, for example that supermarkets should target households with young children and highly educated full-time working parents, as for those groups of consumers the probability to adopt OGS is significantly higher. In addition, supermarkets should focus on providing solutions (e.g., develop websites with a clear layout, sufficient but concise product information, as well as a logical and fast ordering process) for poor-time consumers, because time saving is an important reason for adopting OGS.

Driediger and Bhatiasevi (2019), used the Technology Acceptance Model to identify differences in the acceptance of online grocery shopping in the context of developing countries (e.g., Thailand). They found that perceived usefulness is very likely to develop intentions to use OGS. They also found a positive relationship between perceived ease of use and perceived usefulness, meaning that consumers who find OGS easy to use, are more likely to perceive it as useful. Furthermore, they found a positive relationship between perceived ease of use and intention to use. In terms of gender, perceived enjoyment and perceived ease of use was higher for females. With regard to perceived usefulness towards intention to use, they found that OGS effectiveness and efficiency were more appealing to the higher income respondents. There were no significant differences between married and single respondents, and likewise, the size of the household did not seem to have any significant differences in their acceptance and behavior.

A recent study, conducted in the Middle East North Africa (MENA) region, including Mediterranean countries such as Egypt, Algeria, Tunisia and Morocco, showed that during the Covid-19 pandemic there was an increase of 33% of online grocery shopping (Minawi, 2020). Likewise, in Lebanon, Ben Hassen *et al*. (2022), also found a change in the grocery shopping behavior with a rise in online shopping, as consumers were trying to avoid in-store shopping, however not as much as in the MENA region. This may be attributed to the country's (Lebanon) weak digital infrastructure, which as noted, the Internet infrastructure was sorely inadequate.

In Italy, Dominici *et al*. (2021), explored the effects of socio-demographics and situational factors that influence individuals' likelihood to buy food online. The results, from respondents who participated in the Italian National Institute of Statistics multipurpose survey "Aspects of Italian Daily Life", indicated that young (age range 35-54), well-educated, female, living in a small family, with good or adequate overall economic condition were more likely to adopt online food purchasing. Another study (Maltese *et al*., 2021), examined the consumers' preferences for e-grocery in two major cities in Italy (Rome and Milan) in 2018. Results from this study showed that pricing or timing strategies (e.g., reduction in the delivery fees), can steer consumers towards ordering grocery online, preferably having them delivered at home. Alaimo *et al*. (2022) used Partially Ordered Set (POSET) to measure customer satisfaction in Italy, for online food shopping, during the Covid-19 pandemic. They found that the ease of use of online tools and websites, with which consumers searched and obtained product information, the attractiveness of the user interface and the security perception all had a strong impact on the overall user satisfaction. They concluded that a high level of consumer satisfaction, positively influences the intention to purchase food products (e.g., groceries) online.

Goethals *et al*. (2012), examined the perceptions of French consumers towards e-grocery retailing. They found no statistical significance between men (16%) and women (18%); however, they did find significant evidence ($p=0.001$) depending on consumer's age. Specifically, 32% of consumers age 30-45 years and 20% of those between 45 and 60 years have bought groceries online, while only 10% of consumers age 20-30 and none of consumers older than 60 did their grocery shopping online.

Shih-Chih and Shing-Han (2010), attempted to develop an integrated model, combining the concepts of Technology Readiness (TR) and

TPB, in order to predict and explain consumers' adoption and continued use of e-services. The study took place in Taiwan and the sample consisted of 405 undergraduate students. The findings of the study showed that attitude towards the behavior and perceived behavioral control were two important factors that influenced consumers' intention to adopt e-services, while subjective norms did not influence consumers' intention to continue using e-services. Moreover, consumers' technological readiness had a strong influence on attitude, subjective norm and perceived behavioral control. Therefore, personality traits, as measured in the TR, had a significant impact on technology adoption and users seemed to view the adoption of technology services more openly and positively, and were less likely to focus on its negative aspects.

In a more recent study conducted among older adults in the United States, Wu and Song (2021), proposed an integrated model based on the Technology Acceptance Model (TAM) and the TPB to predict older adults' attitudes and intentions towards online shopping. Two main characteristics of this demographic population, namely perceived lack of mobility and perceived social isolation were used as external variables to better predict their online shopping behavior. The sample consisted of 366 US adults born in or before 1965. According to the results of the study, perceived lack of mobility in shopping was positively related to perceived usefulness, but not to perceived ease of use of online shopping. In other words, older adults who have difficulty with shopping mobility did not find it easy to make online purchases. Moreover, perceived ease of use influenced perceived usefulness of online shopping and these two variables influenced attitudes towards online shopping. Perceived usefulness, perceived behavioral control and attitudes were positively related to the intention to continue shopping online. In addition, perceived social isolation of the target group was negatively related to perceived behavioral control, but positively related to subjective norms. Older adults who lack social connections may believe that they did not have enough resources to do a particular task, but may be more concerned about the views of significant others about online shopping and may have felt more pressure to cope with those views.

Blomqvist *et al*. (2015), examined consumer attitudes towards online grocery shopping. The purpose of that study was to investigate whether consumers' positive attitudes towards OGS influence their intention to purchase groceries online. The study was conducted in Sweden, which is considered one of the most developed e-commerce markets in Europe. In order to test the relationship between consumers' attitudes and intentions, the TPB was used as the underlying theoretical model. The results of the research confirmed the appropriateness of the TPB as a framework for describing the relationship between consumers' attitudes and behavior towards OGS. Moreover, the subjects in the sample appear to have favorable attitudes towards OGS. Finally, the results showed that there was a strong positive correlation between consumers' positive attitudes towards online grocery shopping and their intention to adopt such behavior.

In Spain, Elghannam and Mesias (2019) investigated consumers' perceptions of the potential use of social media, to create online, short food supply chains. The results shed light on what consumers perceive as opportunities or barriers to such short food supply chains, that are found on social media. Foremost, 51% of respondents found the idea of using social media networks for online food shopping to be good, useful, practical and innovative, while a smaller proportion (49%) found the idea strange and risky. As for the incentives that motivate consumers to buy food online, firstly, it is trust in the brand and the company or even a good previous experience with the same company. Another important issue highlighted by participants was their demand for products with quality guarantees and that the products comply with regulations and standards. Furthermore, it was found that a competitive price (lower than that of a conventional chain) or the availability of products not available on the local market, would also motivate consumers to buy food online. Another interesting result was that of delivery service, as consumers consider that secure platforms as well as efficient delivery systems are a key motivation to accept this kind of online transactions.

Based on the findings of the literature review, it appears that relevant research in the Mediterranean region is scarce. Thus, this study aims to fill this gap. The application of TPB in this research will contribute to the collection of useful information for understanding the behavior of Cypriot consumers towards online grocery shopping.

3. Methodology

This study was based on an online survey in Cyprus using a structured, self-administered questionnaire, adapted from previous studies (Ajzen, 1991; Symeou, 2017). The questionnaire was posted online using Google Forms between July and September 2021, during the Covid-19 pandemic and the related 'stay at home' lockdown and other related measures to prevent infection in Cyprus. The choice of online survey was selected because it is cheap, easy to set up, convenient to the participant (in terms of time and place), and do not require a physical contact between the interviewers and the respondents, that was a concern (and restriction) during the Covid-19 outbreak.

The survey population consisted of primary household grocery shoppers (i.e., person primarily responsible for most of the grocery shopping in their household) aged 18 and over, living in Cyprus. The sampling method used was the proportionate stratified random sampling and the required sample size was determined to 284 persons with a 5% margin of error and a 95% confidence level. The stratification variable was the respondents' district of residence; specifically, five strata were created based on the five districts of the Republic of Cyprus as follows: Nicosia, which represented 39% of the sample and total population, Limassol, 28%, Larnaka, 17%, Paphos, 11%, and Famagusta, 5%. Prior to the survey being administered to the general public, a pretest (pilot study) of 18 respondents was conducted with the goal to increase the validity and reliability of our testimonial survey evidence. Measurement scales were constructed on a five-point Likert scale (1 = strongly disagree; 5 = strongly agree).

Attitude towards OGS was measured by three items representing respondents' overall evaluation of the attractiveness towards carrying out OGS. Specifically, it was measured by respondents' level of agreement to the following statements: (1) online grocery shopping makes me feel safer because of the coronavirus pandemic; (2) by buying grocery products online, I avoid commuting and the hassle of shopping at peak times; and (3) by buying grocery products online, I save valuable time in my transactions, making quick and easy comparisons of prices and information about the various products.

Subjective Norms was measured by obtaining the respondents' level of agreement to the following three statements: (1) most people who are important to me and my family would appreciate the fact that I buy grocery products online; (2) most families that are important to my family buy grocery products online on a regular basis; and (3) generally, my wider environment expects me to buy grocery products online.

Three variables that represented respondents' opinions of the convenience of online grocery shopping as well as potential difficulties associated with it were used to gauge Perceived Behavioral Control. A 5-point Likert scale measured respondents' level of agreement to the following statements: (1) if I wanted to buy grocery products online, it would be an easy task for me; (2) it is easy to find all the grocery products I need online; and (3) I believe I have the money needed for delivery of the products.

Intention to OGS was measured by obtaining the respondent's response to the following two items: (1) the next time I will buy grocery products, it is very likely that I will do it online, and (2) I intend to start/continue buying grocery products online on a regular basis.

As objective measurement of behavior can be difficult (Fishbein & Ajzen, 2011), previous research mainly used subjective self-reported data, apart from assessment scales (Ajzen, 2019). In this context, researchers have used other subjective measures such as frequency data for doing a given behavior (Ajzen, 2019; Mirkarimi *et al*., 2016), time spent engaging in a particular behavior (Ajzen, 2020; Prapavessis *et al*., 2015), and simple binary assessments (use/non-use) (Lai & Chen, 2011; Tzeng *et al*., 2022). In this study, behavior was measured by two items rep-

resenting respondents' actual behavior towards OGS: (1) have you ever purchased grocery products online (binary, yes or no), and (2) how often do you purchase grocery products online on a five-point Likert scale (frequency, 1 = never; 5 = very often).

Data analyses were conducted using the open-source statistics program JASP (version 0.17.2.1). JASP was used for descriptive analysis to find out the demographic characteristics of the sample. Cronbach's α was adopted to test the unidimensional internal consistency (reliability) of the measurement items scales (Attitude, Subjective Norms, Perceived Behavioral Control, Intention, and Behavior). Values of α above 0.60 are considered satisfactory (Field & Miles, 2009). Confirmatory factor analysis (CFA) fit indices e.g., Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA), were used to evaluate the goodness-of-fit of a specified model to the observed data, followed by using Structural Equation Modelling (SEM) to test the hypothesized relationships among study constructs. The SEM module in JASP is based on Yves Rosseel's R package lavaan (Rosseel, 2012).

Structural equation modelling is a powerful multivariate data analysis research technique that is widely used in many fields of study to verify complex phenomena (Kang & Ahn, 2021; Sentosa & Mat, 2012; Xia & Yang, 2019). It is the combination of factor analysis and multiple regression analysis and it is used to analyze structural relationships between measured variables (indicators) that can describe the corresponding factor (latent variables). In the case of the TPB model, latent variables are: Attitude towards the behavior, Subjective Norms, Perceived Behavioral Control, Intention and Behavior.

4. Results

4.1. *Descriptive statistics*

Table 1 displays sample averages for several demographic and household variables. A total of 364 questionnaires were finally collected, with 302 of the respondents indicating that they were the primary household grocery shopper, while regional (district of residence) quotas were met. These 302 questionnaires were taken into consideration for further analysis, since when respondents answered "No" to the first question, whether they were the primary household grocery shopper, the interview was terminated. The respondents' median age was 41, which is in line with that of Cyprus population, viz., 37.7 (Eurostat, 2021). At 67.4%, women are dominating, which seems to be in accordance with various studies (Driediger & Bhatiasevi, 2019; Frank & Peschel, 2020; Jribi *et al*., 2020), implying that the grocery shopping is a task performed mainly by women.

Table 1 - Survey sample demographics (n = 302).

Variable	*Mean ± SD or Percentage*
Age	41.71 ± 9.59
Gender	
Male	32.56
Female	67.44
Educational level	
Primary education	0.66
Secondary education	89.07
Tertiary education	10.27
Marital status	
Married	73.18
Not Married	19.87
Other	6.95
Occupation	
Private sector employee	36.75
Public sector employee	47.02
Other	16.23
Monthly income (€)	
Up to 1,000	8.94
1,001 – 2,000	24.50
2,001 – 3,000	29.47
3.001 – 4,000	22.51
More than 4,000	14.57

4.2. *Theory of Planned Behavior constructs measurements*

Descriptive statistics of the questionnaire items used as measurement variables are shown in Table 2, including the mean values of Attitude, Subjective Norms, Perceived Behavioral Control, Intention, and Behavior constructs.

Table 2 - Descriptive statistics of the questionnaire items (latent variables).

Construct / Questionnaire item	*Mean*	*Standard Deviation*
Attitude		
AT1: Online grocery shopping makes me feel safer because of the coronavirus pandemic.	3.49	1.34
AT2: By buying grocery products online, I avoid commuting and the hassle of shopping at peak times.	2.98	1.26
AT3: By buying grocery products online, I save valuable time in my transactions, making quick and easy comparisons of prices and information about the various products.	3.63	1.22
Cronbach's α = 0.87		
Subjective Norms		
SN1: Most people who are important to me and my family would appreciate the fact that I buy grocery products online.	3.01	1.19
SN2: Most families that are important to my family buy grocery products online on a regular basis.	2.25	1.11
SN3: Generally, my wider environment expects me to buy grocery products online.	2.12	1.23
Cronbach's α = 0.82		
Perceived Behavioral Control		
PBC1: If I wanted to buy grocery products online, it would be an easy task for me.	3.98	1.14
PBC2: It is easy to find all the grocery products I need online.	3.13	1.19
PBC3: I believe I have the money needed for delivery of the products.	3.63	1.17
Cronbach's α = 0.64		
Intention to use		
ITU1: The next time I will buy grocery products, it is very likely that I will do it online.	2.57	1.24
ITU2: I intend to start/continue buying grocery products online on a regular basis.	2.65	1.36
Cronbach's α = 0.85		
Behavior		
BEH1: Have you ever bought grocery products online?	0.46	0.50
BEH2: How often do you buy grocery products online?	2.21	1.18
Cronbach's α = 0.62		

Our sample appears to have favorable attitude towards online grocery shopping (Mean >2.5). In terms of subjective norms, the social environment of the sample seems to be weakly supportive of online grocery shopping. Regarding perceived behavioral control, the sample does not appear to face any significant barrier to online grocery shopping and does not consider the cost of product delivery as being high. Similarly, the sample's intentions for online grocery shopping are supportive. Concerning the behavior of the sample towards the OGS nearly half of the participants (45.7%) indicated that they have made grocery purchases online. Of those, 24.17% have purchased fruits and vegetables, 21.52% dairy products, 44.7% canned food, and 23.84% frozen food. It is interesting to mention that none of the respondents has purchased meat products online. Moreover, the great majority of the respondents (75.75%) stated that they use the grocery store's website to purchase groceries online.

In addition to the theory-related information, respondents were asked several other questions (Table 3) in order to better understand Cypriot consumers' behavior towards online grocery shopping. In this sense, participants believe that additional information about products, such as expiry date, origin, and ingredients would increase their willingness to online grocery shopping. Furthermore, it is essential to have a daily picture, particularly for fresh products, so that consumers know what they are purchasing. Moreover, lower prices of products available online compared to the corresponding prices of

Table 3 - Other related questions towards OGS.

Questionnaire item	*Mean*	*Standard Deviation*
It is more likely to buy groceries online from services that are located in Cyprus than abroad	3.64	1.35
I would buy groceries online, provided that the delivery time was met	3.32	1.27
I would buy groceries online, provided that additional information was provided (e.g. expiry date, ingredients, country of origin)	3.54	1.30
I would buy groceries online, provided I had a daily picture of the products that are in the store (especially fresh produce)	3.67	1.29
I would buy groceries online, if accessing and using the respective website was an easy process for me	3.38	1.38
I would buy groceries online, if I was offered lower product prices compared to the corresponding product prices in retail stores	3.84	1.26
I would buy groceries online, if I found product offers	3.82	1.25
I would buy groceries online, if I felt there was security in the transaction	3.51	1.35
I would buy products online, which are not available in the local market	3.79	1.28
I would buy products online if there was a home delivery service in my area	3.57	1.33
I would buy products online knowing that there is a refund if you don't receive the correct product	3.87	1.29

products in the retail store, as well as offers on online products seem to influence consumers' decision to buy groceries online. In addition, home delivery time, as well as the ease of use of the website and transaction security appear to play an important role in online grocery shopping. Among the additional questions, getting a refund option in case consumers don't receive the right product ranked first.

4.3. *Structural equation modelling analysis*

We use structural equation modelling analysis to determine the relative contributions of Attitudes, Subjective Norms, and Perceived behavioral Control to the prediction of Intentions and the relative contributions of Intention and Perceived Behavioral Control to the prediction of Behavior.

As suggested by Anderson and Gerbing (1988), this study uses a two-step approach. Firstly, the reliability of the scales was tested and evaluated. Secondly the model was analyzed in order to test the structure and to evaluate the model's ability to predict a certain outcome.

Measurement Reliability and Validity

As mentioned in the Methodology part, Cronbach's alpha was calculated to evaluate the internal reliability of the questionnaire. The value of *a* was found to be 0.88 and thus exceeded the recommended threshold of 0.7, as suggested by Nunnally (1967).

Confirmatory Factor Analysis (CFA) has been administered in pursuance of testing the measurement model. The results of the CFA indicate that each item loaded on its respective underlying concept (Table 4). As recommended by Kang and Ahn (2021), the model fit indices also suggest that the measurement model was a good fit to the data (χ^2 = 152.48, df=55, p<0.001, Comparative Fit Index [CFI]=0.94, Tucker-Lewis Index [TLI]=0.92 Root Mean Square Error of Approximation [RMSEA] =0.08).

Testing of the Structural Equation Model

The SEM model (Figure 2) describes well our data as suggested the model fit statistics [(χ^2(302) =203.19, df=70, p<0.001); CFI=0.92; TLI=0.91; RMSEA=0.08]. Our model explains 45% of the variance in consumers' Intention

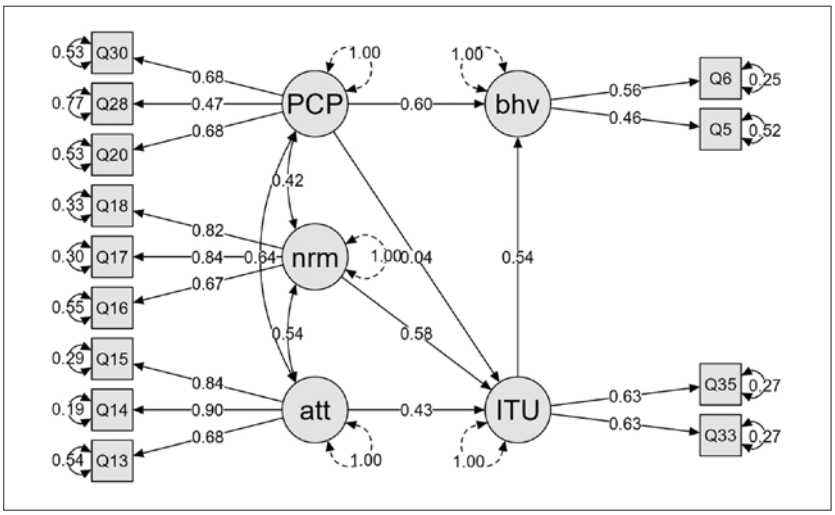


Figure 2 - Path diagram of the TPB fitted to the raw latent variables.

to OGS (R^2=0.45) and 56% of the variance in consumers' Behavior to OGS (R^2=0.56). This is expected, because as explained by Chandon *et al*. (2005), consumers' self-reported purchase intentions do not perfectly predict their future purchase behavior. Similarly, Sheeran (2002), noted that the gap between intentions and behavior is not negligible, as a meta-analysis indicated that on average, intentions explain 28% of the variance in future behavior. In addition, Sheppard *et al*. (1988), in a meta-analysis of 87 behavior studies found a frequency-weighted average of 0.53 with wide variations across measures of intentions and types of behavior. The regression coefficients and results of hypotheses tests are shown in Table 4.

5. Discussion

Our data confirm that the TPB model is appropriate for examining the behavioral model of the Cypriot consumer in terms of online grocery shopping. Thus, the information collected through this model effectively describes the various aspects and perceptions of Cypriot consumers regarding online grocery shopping.

Cypriot consumers' Subjective Norms and Attitude towards OGS influence their Intentions to buy groceries online, but this is not the case for the Perceived Behavior Control. However, Perceived Behavioral Control affects their actual Behavior. Likewise, Behavioral Intentions to OGS are likely to translate into actual Behavior.

Table 4 - Parameter estimates for applying the TPB.

Path (hypothesis)	*Path Coefficients*	*Std Error*	*Pr(>|z|*	*Result of Hypothesis test*
Attitude → Intention (H_1)	0.43	0.12	<0.001	Statistically significant
Subjective Norms → Intention (H_2)	0.58	0.10	<0.001	Statistically significant
Perceived Behavioral Control → Intention (H_3)	0.04	0.12	<0.77	Not statistically significant
Intention → Behavior (H_4)	0.54	0.10	<0.001	Statistically significant
Perceived Behavioral Control→ Behavior (H_5)	0.60	0.14	<0.001	Statistically significant

Fitting the TPB model to the survey data shows that Subjective Norms (β=0.58) have greater influence than Attitude (β=0.43) or Perceived Behavioral Control (β=0.04) on Intention to OGS. The results suggest that Attitude (z-value=3.471, $p<0.001$) and Subjective Norms (z-value 5.49, $p<0.001$) have a relatively strong and significant relationship and support H1 and H2, Intention for OGS.

Perceived Behavioral Control (β=0.04, p=0.77) shows insignificant effect on Intention and thus rejects H3. With $R^2 = 0.45$ the three variables Attitude, Subjective Norms, and Perceived Behavioral Control together account for 45.9% of the variance in Intention to online grocery shopping. Both Intention to online grocery shopping (β=0.54, $p<0.001$) and Perceived Behavioral Control influence OGS behavior (β=0.60, $p<0.001$) and have a significant effect on OGS; thus, supporting H4 and H5.

The results of this study also suggest that Perceived Behavioural Control carries more weight (β=0.60) than Intention (β=0.54), in explaining online grocery shopping behavior. This is in line with other studies (Blitstein *et al*., 2020; Driediger & Bhatiasevi, 2019), where the behavior is more likely to be affected by other functional factors such as the availability of resources (e.g., internet access) and access to services (e.g., delivery in remote areas). The significance of Perceived Behavioral Control as the construct with the largest effect on online grocery shopping behavior in the TPB model, confirms that people should feel in control of their ability to online grocery shopping. More specifically, similar to other studies findings, it should be easy to find all the grocery products that one needs online, it should be convenient for them to have access to online services for the purchase of grocery products, and finally, home delivery services should be affordable (Alaimo *et al*., 2022; Loketkrawee & Bhatiasevi, 2018; Rogus *et al*., 2020; Wang *et al*., 2023).

Furthermore, Attitude and Subjective Norms have direct influences on online grocery shopping Intention. Among these relationships, Subjective Norms indicates more importance than Attitude in determining behavioral intention to online grocery shopping. Similar to other empirical studies (Shih-Chih & Shing-Han, 2010), our research has shown that Subjective Norms have a positive and direct impact on Intention, compared to Attitude. In other words, the effect of "word-of-mouth" or "social pressure" have significant influence on shaping behavioral intention, which is in accordance with findings in a social media consumer related study (Elghannam & Mesias, 2019).

6. Conclusion

Online grocery shopping has grown in the past decade as a way to satisfy increased demand for convenience (Rogus *et al*., 2020) and the level of satisfaction (Alaimo *et al*., 2022) among consumers. Despite its fast growth rate in the past few years, primarily possibly due to Covid-19 pandemic related movement restrictions, online grocery shopping has ever gained little attention so far. It is notable that in Cyprus, the previous study on this topic occurred in 2006. This paper fills this research gap and opens a discussion about consumer behavior and online grocery shopping strategies. As reported earlier by Giannakopoulou *et al*. (2022), age and family income are key drivers of online grocery shopping. Decreasing age and increasing monthly family income were associated with an increased likelihood of online grocery shopping. Furthermore, our analysis enriches the literature on the consumers' profile characteristics of the adoption of online grocery shopping, especially in the Euro-Mediterranean region. This study provides new empirical evidence on the understanding of the determinants of online grocery shopping in Cyprus.

Big life changes (e.g., Covid-19 pandemic [Alaimo *et al*., 2022]) and transitions (e.g., caring for a sick relative or the birth of a child (Blitstein *et al*., 2020)), may trigger an initial adoption of a behavior (i.e., online grocery shopping), given the higher demand for social distancing, convenience and/or time-saving features. As noted by Goethals *et al*. (2012) and Frank and Peschel (2020), convenience and time saving aspects are also possible drivers for adopting online grocery shopping, especially for households with full-time employment and presence of young

children. Before the Covid-19 pandemic, only 7% of grocery retail sales worldwide involved e-commerce channels. During the peak of the pandemic (March-April 2020), the e-commerce share of grocery retail grew to 10%, while a recent study (Coppola, 2021), showed that the current share of global e-commerce penetration for online grocery shopping was valued at 9%. As suggested by Tyrväinen and Karjaluoto (2022), the managerial implications from such findings (e.g. perceived usefulness and attitude) can help retailers in improving their online grocery shopping services.

Another contribution of this research was the investigation of consumers' attitude towards online grocery shopping, the factors that enable online grocery shopping, behavioral intention and related social norms. As noted by Irianto (2015), attitude refers to "the extent to which an individual has good or bad evaluation or assessment on the concerned behavior". Results from Hansen (2008), showed that personal value affect consumers' attitude towards online grocery shopping. Our research has shown that Perceived Behavioral Control carries more weight than Intention in explaining online grocery shopping behavior, and that two antecedent constructs (Attitude and Subjective Norms) in TPB have direct influences on Intention for online grocery shopping.

One of the limitations of this study, as noted by several researchers, is that consumers' self-reported purchase intentions do not perfectly predict their future purchase behavior and the gap between intentions and behavior is not negligible (Sheeran, 2002). As suggested by Ajzen (1991), to improve behavioral predictions one should measure intentions after changes have occurred (e.g., follow-up research in the post-Covid era). This is a subject of a future research.

References

Adamides G., Marianthi G., Savvides S., 2006. *Traditional vs online attitudes towards grocery shopping in Cyprus*. Computers in Agriculture and Natural Resources, 23-25 July 2006, Orlando, Florida,

Ajzen I., 1991. The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2): 179-211.

Ajzen I., 2019. *Constructing a Theory of Planned Behavior questionnaire*. Retrieved June 11, 2024 from https://people.umass.edu/aizen/pdf/tpb.measurement.pdf.

Ajzen I., 2020. The theory of planned behavior: Frequently asked questions. *Human behavior and emerging technologies*, 2(4): 314-324.

Alaimo L.S., Fiore M., Galati A., 2022. Measuring consumers' level of satisfaction for online food shopping during COVID-19 in Italy using POSETs. *Socio-Economic Planning Sciences*, 82: 101064.

Anderson J.C., Gerbing D.W., 1988. Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin*, 103(3): 411.

Ben Hassen T., El Bilali H., Allahyari M.S., Charbel L., 2022. Food shopping, preparation and consumption practices in times of COVID-19: case of Lebanon. *Journal of Agribusiness in Developing and Emerging Economies*, 12(2): 281-303.

Blitstein J.L., Frentz F., Jilcott Pitts S.B., 2020. A mixed-method examination of reported benefits of online grocery shopping in the United States and Germany: Is health a factor? *Journal of Food Products Marketing*, 26(3): 212-224.

Blomqvist A., Nyman L., Lennartsson F., 2015. *Consumer attitudes towards online grocery shopping: A research conducted on Swedish consumers*. Bachelor's Thesis within Business Administration, Jönköping University.

Chandon P., Morwitz V.G., Reinartz W.J., 2005. Do intentions really predict behavior? Self-generated validity effects in survey research. *Journal of marketing*, 69(2): 1-14.

Coppola D., 2021. *Global development of e-commerce shares of grocery stores before and after the coronavirus (COVID-19) pandemic*. MasterCard data & services. Retrieved August 1, 2022 from https://www.statista.com/statistics/1229979/grocery-e-commerce-shares-development-during-pandemic/.

Dannenberg P., Fuchs M., Riedler T., Wiedemann C., 2020. Digital transition by COVID-19 pandemic? The German food online retail. *Tijdschrift voor economische en sociale geografie*, 111(3): 543-560.

Dominici A., Boncinelli F., Gerini F., Marone E., 2021. Determinants of online food purchasing: The impact of socio-demographic and situational factors. *Journal of Retailing and Consumer Services*, 60: 102473.

Driediger F., Bhatiasevi V., 2019. Online grocery shopping in Thailand: Consumer acceptance and usage behavior. *Journal of Retailing and Consumer Services*, 48: 224-237.

Elghannam A., Mesias F., 2019. Short food supply chains from a social media marketing perspective: A consumer-oriented study in Spain. *New Medit*, 18(1): 79-90.

Essakkat K., Mattas K., Unay-Gailhard I., Baourakis G., 2021. Youth's potential of adopting the Mediterranean diet lifestyle in response to climate change: Empirical study in Crete, Greece. *New Medit*, 20(5): 85-95.

Etumnu C.E., Foster K.A., Widmar N.O., Lusk J.L., Ortega D.L., 2019. *Drivers of online grocery shopping*. Agricultural and Applied Economics Association 2019 Annual Meeting, July 21-23, Atlanta, Georgia.

Eurostat, 2021 (June). *Population structure and ageing*. Retrieved 18 February, 2022 from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_structure_and_ageing.

Field A., Miles J., 2009. *Discovering statistics using SAS:(and sex and drugs and rock'n'roll)*. University of Sussex.

Fishbein M., Ajzen I., 2011. *Predicting and changing behavior: The reasoned action approach*. London: Psychology press.

Frank D.-A., Peschel A.O., 2020. Sweetening the Deal: The Ingredients that Drive Consumer Adoption of Online Grocery Shopping. *Journal of Food Products Marketing*, 26(8): 535-544. https://doi.org/10.1080/10454446.2020.1829523.

Giannakopoulou M., Adamides G., Stylianou A., 2022 (September 22-25). *Consumer Adoption of Online Grocery Shopping: Findings from a case study in Cyprus*. Athens: HAICTA.

Goethals F., Leclercq-Vandelannoitte A., Tütüncü Y., 2012. French consumers' perceptions of the unattended delivery model for e-grocery retailing. *Journal of Retailing and Consumer Services*, 19(1): 133-139.

Hansen T., 2008. Consumer values, the theory of planned behaviour and online grocery shopping. *International Journal of Consumer Studies*, 32(2): 128-137.

Irianto H., 2015. Consumers' attitude and intention towards organic food purchase: An extension of theory of planned behavior in gender perspective. *International journal of management, economics and social sciences*, 4(1): 17-31.

Jensen K.L., Yenerall J., Chen X., Yu T.E., 2021. US Consumers' Online Shopping Behaviors and Intentions During and After the COVID-19 Pandemic. *Journal of Agricultural and Applied Economics*, 53(3): 416-434. https://doi.org/10.1017/aae.2021.15.

Jribi S., Ismail H.B., Doggui D., Debbabi H., 2020. COVID-19 virus outbreak lockdown: What impacts on household food wastage? *Environment, Development and Sustainability*, 22(5): 3939-3955.

Kang H., Ahn J.-W., 2021. Model setting and interpretation of results in research using structural equation modeling: A checklist with guiding questions for reporting. *Asian Nursing Research*, 15(3): 157-162.

Lai H.-M., Chen C.-P., 2011. Factors influencing secondary school teachers' adoption of teaching blogs. *Computers & Education*, 56(4): 948-960.

Loketkrawee P., Bhatiasevi V., 2018. Elucidating the behavior of consumers toward online grocery shopping: The role of shopping orientation. *Journal of Internet Commerce*, 17(4): 418-445.

Maltese I., Le Pira M., Marcucci E., Gatta V., Evangelinos C., 2021. Grocery or @grocery: A stated preference investigation in Rome and Milan. *Research in Transportation Economics*, 87: 101096.

Minawi M., Kandil N., Qahoush N., 2020. *5 ways COVID-19 has impacted MENA's food habits*. Retrieved 6 October, 2023, from https://www.ipsos.com/sites/default/files/ct/news/documents/2020-06/5_ways_covid-19_impacted_menas_food_habits_-_ipsos_mena_0.pdf.

Mirkarimi K., Mansourian M., Kabir M.J., Ozouni-Davaji R.B., Eri M., Hosseini S.G., Qorbani M., Safari O., Rastgari Mehr B., Noroozi M., 2016. Fast food consumption behaviors in high-school students based on the Theory of Planned Behavior (TPB). *International journal of pediatrics*, 4(7): 2131-2142.

Nielsen, 2017. *What's in-store for online grocery shopping*. The Nielsen Company. Retrieved February 10, 2022 from https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/nielsen-global-connected-commerce-report-january-2017.pdf.

Nunnally J.C., 1967. *Psychometric theory*. New York: McGraw-Hill.

Prapavessis H., Gaston A., DeJesus S., 2015. The theory of planned behavior as a model for understanding sedentary behavior. *Psychology of Sport and Exercise*, 19: 23-32.

Research and Markets, 2022. *Food & Grocery Retail Market Size, Share & Trends Analysis Report by Product (Food Cupboard, Beverages), by Distribution Channel (Supermarkets & Hypermarkets, Online), by Region (APAC, Europe), and Segment Forecasts, 2022-2030*. Retrieved October 6 from https://www.researchandmarkets.com/reports/4613444/food-and-grocery-retail-market-size-share-and.

Rogus S., Guthrie J.F., Niculescu M., Mancino L., 2020. Online grocery shopping knowledge, attitudes, and behaviors among SNAP participants. *Journal of nutrition education and behavior*, 52(5): 539-545.

Rosseel Y., 2012. Iavaan: An R package for structural equation modeling. *Journal of statistical software*, 48: 1-36.

Sentosa I., Mat N.K.N., 2012. Examining a theory of planned behavior (TPB) and technology acceptance model (TAM) in internetpurchasing using structural equation modeling. *Researchers World*, 3(2 Part 2): 62.

Sheeran P., 2002. Intention—behavior relations: a conceptual and empirical review. *European review of social psychology*, 12(1): 1-36.

Sheppard B.H., Hartwick J., Warshaw P.R., 1988. The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of consumer research*, 15(3): 325-343.

Shewale R., 2023. *Internet user statistics in 2023 — (Global Demographics)*. Retrieved October 6 from https://www.demandsage.com/internet-user-statistics/.

Shih-Chih C., Shing-Han L., 2010. Consumer adoption of e-service: Integrating technology readiness with the theory of planned behavior. *African Journal of Business Management*, 4(16): 3556-3563.

Strydom W.F., 2018. Applying the theory of planned behavior to recycling behavior in South Africa. *Recycling*, 3(3): 43.

Symeou P., 2017. T*he Behavioral Model of the Cypriot Consumer of Organic Food: The Foundation of a National Strategic Plan for Promoting the Consumption of Organic Food*, Cyprus University of Technology.

Topa G., Moriano J.A., 2010. Theory of planned behavior and smoking: Meta-analysis and SEM model. *Substance abuse and rehabilitation*, 1: 23-33.

Tyrväinen O., Karjaluoto H., 2022. Online grocery shopping before and during the COVID-19 pandemic: A meta-analytical review. *Telematics and Informatics*, 71: 101839.

Tzeng S.-Y., Lin K.-Y., Lee C.-Y., 2022. Predicting college students' adoption of technology for self-directed learning: a model based on the theory of planned behavior with self-evaluation as an intermediate variable. *Frontiers in Psychology*, 13: 865803.

Van Droogenbroeck E., Van Hove L., 2017. Adoption of online grocery shopping: personal or household characteristics? *Journal of Internet Commerce*, 16(3): 255-286.

Wang K., Gao Y., Liu Y., Habib K.N., 2023. Exploring the choice between in-store versus online grocery shopping through an application of Semi-Compensatory Independent Availability Logit (SCIAL) model with latent variables. *Journal of Retailing and Consumer Services*, 71: 103191.

Wu J., Song S., 2021. Older adults' online shopping continuance intentions: Applying the technology acceptance model and the theory of planned behavior. *International Journal of Human–Computer Interaction*, 37(10): 938-948.

Xia Y., Yang Y., 2019. RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior research methods*, 51(1): 409-428.

The impacts of climate change on agriculture sector - Case study: Mediterranean countries

Yasmine Abd El-Nasser*, Ehab Abdelaziz Ibrahim*

DOI: 10.30682/nm2404c
JEL codes: C32, O13, Q01, Q54

Abstract

Climate change is a complex phenomenon, and the agricultural sector is the most vulnerable to its effects since it affects negatively the food security dimensions, which are directly associated with agriculture. Climate change influences food production and inflation; the impacts have been examined in twelve Mediterranean countries. The paper was equipped with simple and advanced statistics and econometric instruments like descriptive statistics, unit root tests, ARDL (2002-2022), and forecasting models. The results show that climate change has a negative impact on the agriculture sector for many Mediterranean countries in both the current situation and future scenarios. Both temperature and CO2 emissions have a significant negative effect on food production and inflation in the selected countries. Moreover, combining climate change with a rapid increase in population growth rate will put more pressure on the agriculture sector, resulting in an increase in the percentage of POU among the people of those countries. The findings confirm the interlinkage between the sustainable development goals in SDG 3 and SDG 13.

Keywords*: Climate change, Agricultural sector, Food security, ARDL model, Sustainable development goals*

1. Introduction

Climate change is one of the main global challenges, and all countries are affected by it in different degrees, levels, and negative impacts. The United Nations Framework Convention on Climate Change (UNFCCC, 1992) defines climate change as "a change of climate which is attributed directly or indirectly to human activity that alters the composition of the global atmosphere and which is in addition to natural climate variability observed over comparable time periods". Climate change increases risks to food security, particularly in coastal countries like Spain, France, Monaco, Italy, Slovenia, Croatia, Bosnia and Herzegovina, Montenegro, Albania, Greece, Turkey, Syria, Lebanon, Palestine, and Egypt (IPCC, 2012). The World Food Summit (1996) defines food security as "a situation when all people, at all times, have physical and economic access to sufficient, safe, and nutritious food that meets their dietary needs and food preferences for an active and healthy life". A new study and international reports showed that there is an increase in the expansion of food insecurity levels among coastal countries, leading to more conflicts, food prices, and catastrophic levels of hunger in those countries. The Agricul-

* Cairo University, Faculty of Agriculture, Giza, Cairo, Egypt.
Corresponding author: yasmine.abd.elnasser@agr.cu.edu.eg

ture Organization of the United Nations (1996) defines food insecurity as "when people do not have adequate physical and economic access to sufficient, safe, and nutritious foods that meet their dietary needs and preferences for an active and healthy life". According to the Global Food Security Index (GFSI) in 2022, Spain ranks 20 out of 113 countries with an overall food security score of 75.7, France ranks 4 with an overall score of 80.2, Italy ranks 27 with an overall score of 74, Greece ranks 31 with an overall score of 72.2, Turkey ranks 49 with an overall score of 65.3, Syria ranks 113 with an overall score of 36.3, and Egypt ranks 77 out of 113 countries with an overall food security score of 56. In most coastal countries, the agricultural sector is the primary sector that plays a vital role in their economies. For example, the agricultural sector in Egypt represents about 11% of the country's GDP, employs more than 50% of the labor force, and provides about 50% of the food of the country (The World Bank, 2023). So, a global challenge like climate change can affect the agricultural sector, leading to negative impacts on food security and causing national insecurity and political instability all over the country and/or region. Climate change has an effect on the four food security dimensions, which are food availability, food access, food utilization, and food stability. According to the IPCC in 2012, there are five key risks caused by climate change that have a direct negative impact on food security: (1) breakdown of food systems; (2) loss of marine and coastal ecosystems; (3) loss of rural livelihoods and income; (4) loss of terrestrial and inland water resources; and (5) negative effects on trade flows, food markets, and price stability. The main problem of this paper is that climate change has a negative effect on the agriculture sector, particularly in the Mediterranean countries whose economies are based on agriculture. Climate change has effects on temperature (rising it up), leading to a decrease in crop yields. In addition, it affects annual precipitation (decreasing), leading to a reduction in agricultural production. Moreover, climate change has an effect not only on the quantity of food production but also on the food quality standards, leading to the spread of diseases due to a decrease in air, water, and food quality, especially among children under the age of five. Therefore, climate change has a negative effect on all food security dimensions based on the strong direct relationship between agriculture and food security. According to the Global Hunger Index (GHI) in 2023, the global hunger index scores for coastal countries are as follows: Albania with a 6.1 score, Bosnia and Herzegovina with a 5 score, Croatia with a 5 score, Egypt with a 12.8 score, Libya with a 16.1 score, Morocco with a 9 score, Syria with a 26.1 score, Turkey with a 5 score, and Tunisia with a 5.9 score. So, this paper serves a main objective, which is to investigate the negative impacts of climate change on the agriculture sector of Mediterranean countries such as Albania, Algeria, Morocco, France, Croatia, Egypt, Syria, Tunisia, Turkey, Greece, Spain, and Italy. This paper serves a main objective, which is to investigate the impacts of climate change on the agriculture sector, with a focus on Mediterranean countries like Spain, France, Monaco, Italy, Slovenia, Croatia, Bosnia and Herzegovina, Montenegro, Albania, Greece, Turkey, Syria, Lebanon, and Egypt. In order to achieve the paper's main objective, this paper depends on using an auto-regressive distributed lag model (ARDL) during the period (2002-2022) to examine the current impacts of climate change on the agriculture sector for the selected Mediterranean countries and then estimate the potential impacts of climate change on the agriculture sector for those countries in the future. The data in this paper relies on different secondary sources like the Climate Change Knowledge Portal (CCKP), the Food and Agriculture Organization of the United Nations (FAO), and the World Bank (WB). This paper will contribute to the existing literature on the impacts of climate change by conducting a comprehensive study on those negative impacts on the agriculture sector in the selected Mediterranean countries, since less attention has been given to studying the actual and predicted impacts of climate change in the Mediterranean countries. So, the paper is organized as follows: Section (1) includes a literature review for previous studies that investigate the impacts of climate change on the agriculture sector, particularly in the Mediterranean countries; Section (2)

includes data sources, variables, and descriptive statistics beside the theoretical framework of the ARDL model; Section (3) illustrates the current and potential negative impacts of climate change on the agriculture sector in the selected Mediterranean countries; and Section (4) includes the expected recommendations and conclusions based on the results shown and collected by the ARDL model.

2. Literature review

This paper will follow two main studies (Yassin, 2016) and (Hashem, 2020), but the contribution in this paper will be on a larger scale. The study by Yasin (2016) focuses on the impacts of climate change on food security in Egypt by concentrating on their agricultural sector and the food production system. This study relies on interviews with many experts in the agricultural field from different organizational levels (governmental, local, and international). The study highlights the challenges and obstacles that the agricultural and food production systems faced because of climate change. Additionally, this study also investigates different management and adaptation policy options to reduce the negative impacts of climate change on the agricultural sector in Egypt. Moreover, the study finds that the current policies are not efficient or effective due to political, social, and organizational obstacles. So, the study recommends that establishing and enhancing a clear, comprehensive legal policy toward lowering the potential impacts of climate change will lead to a higher level of cooperation among different stakeholders. The other study that was conducted by Hashem (2020) focused on the negative impacts of climate change on agricultural, food, and environmental systems. The study uses the ARDL model to examine the expected impacts of climate change on the food security dimension. The results showed that there is a negative and significant impact of temperature on food production, while there is a positive and significant impact of temperature on food access. So, the study recommends that promoting a comprehensive adaptation system will reduce the impacts of climate change on the agricultural sector.

The impacts of climate change on the agriculture sector have been pointed out in many previous studies. Climate change has a negative effect on food availability. The U.S. Department of Agriculture (USDA, 2015) defines food availability as "the availability of sufficient food quantities in a stable way. It includes production, processing, packing, transport, storage of food, and all necessary systems included in doing those procedures". Climate change has a direct impact on food production since agricultural production depends on climate conditions; therefore, the agriculture sector is vulnerable to the risks of climate change effects (Greg, 2011). There is another indirect impact of climate change on agricultural production by effecting income distribution and growth, which therefore affects the demand for products from the agriculture sector (Schmidhuber and Tubiello, 2007). Based on several studies and reports, there is a potential disturbance in crop production because of climate change in the near future (2030-2050). Climate change can increase the frequency of heat stress waves, and more warm nights can lead to a reduction in rice production, especially in the Asia region, by 10% for every 1^{o}C above critical temperature (> 24^{o}C). It is expected to reduce rice production in north-eastern Thailand by 17.8% (without an adaptation strategy). Increasing the global temperature in the next few years will also affect the production of wheat, like in India, where the potential reduction will reach 50% by 2050 and in South East Australia by 24% in 2050 without an adaptation strategy (Nelson, 2009). Moreover, changes in rainfall systems caused by climate change can affect crop yields directly and indirectly, particularly in developing countries, since the rainfall provides the crops with their water and CO2 fertilization requirements. A study by Ray (2009) investigates the relationship between climate change effects on rainfall and crop yields for the main crops in the world. The results show that climate change can affect rainfall systems, leading to lower crop yields for most selected crops, like -13.4% oil palm production, -0.3% rice production, and -0.9% wheat production.

The effects of climate change on food access have been analyzed in several studies. The

World Food Programme (WFP, 2009) defines food access as "a household's ability to acquire an adequate amount of food regularly through a combination of purchases, barter, borrowings, food assistance, or gifts". Climate change has a negative effect on the food access dimension because a decrease in agricultural production (a decrease in food supply) will lead to a rise in food prices. Nelson (2014). Moreover, studies such as Nelson (2009) and Jobbins & Henley (2015) point to the negative effects of climate change on rural incomes that influence the agriculture sector, leading to a lowering purchasing power (decrease in incomes), which affects negatively economic access, and rural households will be the most vulnerable since their primary income comes from agriculture. Additionally, climate change can cause imbalances between demand and supply sides through increasing prices of food, which influence the demand side by decreasing incomes and the supply side by decreasing productivity. This Nelson's study estimates that prices will rise by 97% for maize, 67% for rice, and 56% for wheat by 2050 because of climate change effects. A study by FAO (2018) indicates the negative impacts of climate change on food access and that food prices will rise by an average of 15% by 2050. There is also a study by Hertel and Rosch (2010) that assesses the negative impact of climate change on staple prices, and the results showed that the global price of many staples will increase from 10 to 60% by 2030 and levels of poverty will rise from 20 to 50% in different areas of Asia and Africa. A study by Ivanic and Martin (2008) investigated the relationship between increased food prices and poverty in nine low-income countries, and the results showed that raising food prices led to an increase in poverty levels in these low-income countries. This study estimates that more than 100 million people (out of 2.4 billion people in low-income countries) who live in those low-income countries will suffer from poverty based on national rates of poverty (US $1 per day).

The negative impacts of climate change on food utilization have been discussed in many previous studies. The World Food Programme (WFP, 2009) defines food utilization as "safe and nutritious food that meets their dietary needs". Based on this definition, it is not enough that food be available and accessible for people; it also must be a safe and nutritious diet. A study conducted by Nelson (2009) showed that climate change can increase child malnutrition and lower calorie consumption. This study estimates that a decrease in calorie availability will lead to an increase in child malnutrition by 20% by 2050. A study by Jobbins and Henley (2015) investigates the negative impacts on food utilization dimensions caused by climate change, like spreading diseases (due to rising temperatures) and poor quality of water (due to droughts and floods). This study assessed the number of children malnourished under the age of five in MENA countries, and the results showed that about two million children under the age of five will suffer from malnutrition by 2050 because of climate change effects. Another study by Porter (2014) and Tirado (2010) analyzes the potential impacts of climate change on the food utilization dimension and how that influences the production and consumption of some foods, the nutritional quality of foods, and the spreading of diseases and illnesses like diarrhea. Moreover, a study by Simon J. Lioyed (2011) discussed the relationship between stunting and a lack of food because of climate change. This study showed that climate change will increase moderate stunting from 1% to 29% by 2050. In addition, climate change will also lead to a rise in severe stunting rates, which will reach 23% in central Sub-Saharan Africa (SSA) and 62% in South Asia by 2050 relative to the future scenario without a climate change adaptation strategy.

Stability is the fourth dimension of food security, and according to the World Food Summit, "to be food secure, a population, household, or individual must have access to adequate food at all times". Based on this definition, food stability points out the security of food, as all people must have access to sufficient food at all times, and there is no risk of sudden food supply shocks such as climatic crises that can cause seasonal food insecurity. Moreover, the agriculture sector has sensitivity and vulnerability to external shocks like economic crises, climatic crises, increases in food prices, and emergencies like spreading diseases, droughts, floods, and pests (FAO, 2006). So, food stability means a permanent situation

"sustainability" that does not occur in a moment, a day, or a season. According to FAO (2013), food stability indicators include political stability and absence of violence or terrorism, domestic food price volatility, per capita food production variability, per capita food supply variability, cereal import dependency ratio, percentage of arable land equipped for irrigation, and the value of food imports over total merchandise exports. Climate changes can affect the food stability dimension, which influences the stability of the other food security dimensions (stability of food availability, access, and utilization). Climate change can lead to food instability by decreasing global production (food supply shocks) and raising the real prices of food (food inaccessibility). A study by Nathan (2016) indicates that several crops have fluctuations in yields because of climate change effects on rainfall, temperature degrees, and the frequency of droughts and floods. Therefore, the World Food Programme (WFP, 2009) defines food insecurity as "a situation where some people do not have access to sufficient quantities of safe and nutritious food and hence do not consume the food that they need to grow normally and conduct an active and healthy life". Food insecurity can occur due to lack of food (food unavailability), lack of resources (food inaccessibility), improper use (no food utilization), and changes in time and/or dimension (food instability).

3. Data and methods

In this part of the paper, the author will describe the data characteristics, sources, selected variables, and selected countries that have been taken into consideration to obtain an appropriate result from the ARDL model, which tackles the similarity and differences for specific characteristics between the Mediterranean countries to measure the food security performances or scores for each MENA country in this study based on the data availability and limitations.

3.1. *Data sources*

The data was collected from different secondary sources and based on international organizations that specialize in collecting facts and databases about the global situation of climate change and/or food security, like the Food and Agriculture Organization of the United Nations (FAO), the World Bank (WB), and the Climate Change Knowledge Portal (CCKP). The data sources also include trusted websites such as the Population Pyramid, Country Economics, and Trading Economics databases. The main limitation of this paper is that only secondary data has been used. Conducting a survey across the selected countries could have strengthened the paper, but it was avoided due to the high financial cost required to gather accurate primary data.

3.2. *Data variables*

Climate change and food security variables that related to agriculture sector and used in this paper are as follows: (1) *Mean temperature*: "The average temperature of the air as indicated by a properly exposed thermometer during a given time period, usually a day, a month, or a year"; (2) *Per capita CO_2 emissions*: "represent the emissions of an average person in a country or region - they are calculated as the total emissions divided by population"; (3) *Food production index*: "Food production index covers food crops that are considered edible and that contain nutrients. Coffee and tea are excluded because, although edible, they have no nutritive value"; (4) *Consumer price index*: "The CPI is a measure of the change in the purchasing power of a country's currency and the price level of a basket of goods and services"; (5) *Prevalence of undernourishments*: "POU expresses the probability that a randomly selected individual from the population consumes a number of calories that is insufficient to cover her/his energy requirement for an active and healthy life. The indicator is computed by comparing a probability distribution of habitual daily dietary energy consumption with a threshold level called the minimum dietary energy Requirement. Both are based on the notion of an average individual in the reference population"; and (6) *Cereal import dependency ratio*: "The cereals import dependency ratio tells how much of the available domestic food supply of cereals has been imported and how much comes from the country's own production".

3.3. *ARDL model specifications*

In this section, the author will review the methodology used for this paper, which studies the impacts of climate change (as an independent variable) on selected indicators of the agricultural sector (as a dependent variable) through the use of time series analysis methods. So, the paper will rely on using the time series method through the co-integration method, as it must first ensure the stability of the time series, which is done by testing the presence of the unit root in the time series by the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests. The process of unit root testing is to perform the following regression:

$$\Delta y_t = \phi_1 y_{y-1} + \alpha + \beta t + \mu_t \quad (1)$$

On the time series data for which the unit root is to be tested, the test can be performed with the presence of the trend or not, and the extended formula is the most commonly used in performing the unit root test, and it is in the form shown below:

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \alpha_i \sum_{i=1}^{m} \Delta y_{t-1} + \varepsilon_t \quad (2)$$

But it should be noted here that the PP test offers another way to test the static of the time series and is considered better than the ADF test because the latter is based on the assumption that the time series is generated by the autoregressive process, while the PP test is based on a more general assumption, the time series is generated by the autoregressive integrated moving average process, whose methodology is attributed to Box-Jenkins.

After ensuring that the time series is stable, a co-integration test must be conducted. If there is co-integration between two or more variables, that means they participate in the same direction, i.e., if they have a long-term equilibrium relationship. Regression between them, and the most famous of these methods are the Engel-Granger method and the Johnson technique. The first method is often used in simple models that contain only two variables and have a large number of observations, as the method assumes that if the time series are integrated in the first order as individuals, CI~(1) (stable after the first difference) and regression can be performed in simple linear form according to the following equation:

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad (3)$$

As for the second method, it provides another method for testing co-integration. The idea of the Johansson test is to know the number of cointegration vectors between the variables based on the Max-Eigen statistic and the Trace statistic, and the general form of the Johansen technique is as follows:

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{p} In(1 - \hat{\lambda} t) \quad (4)$$

$$\lambda_{max}(r, r+1) = -T \, In \, (1 - \hat{\lambda}_{r+1}) \quad (5)$$

Where (r) represents the number of integration vectors under the null hypothesis, the null hypothesis is tested, which states that the number of cointegration vectors is at least equal to (r) vector, and this hypothesis is rejected in return for accepting the hypothesis that states that the number of integration vectors is more than (r) vector If the calculated value of the statistic for the maximum value (λmax) and the value of the effect statistic (λtrace) are greater than the critical values at an assumed significance level.

4. Results and discussion

The agricultural sector is one of the most vital sectors in most Mediterranean countries, and it is vulnerable to climate change. So, in this section, the paper will estimate the potential impact of climate change on agriculture sector of selected Mediterranean countries. In order to estimate those potential impacts, the author will adopt the ARDL model approach to investigate the negative impacts of climate change on the agricultural sector of those countries that affect food security dimensions (food availability, access, utilization, and stability). Therefore, the paper will focus on certain indicators related to climate change, agriculture, and food security, like mean temperature, per capita carbon dioxide emission, production index, consumer price index, POU, and cereal import dependency ratio.

4.1. *Descriptive statistics*

Descriptive statistics were conducted to describe or summarize the characteristics of the data for all variables mentioned earlier in Table 1. It includes three main categories of measures: (1) *Measures of central tendency*: summarize the center of the data set, like mean, median, and etc.; (2) *Measures of frequency distribution*: the occurrence of data within the data set as count; and (3) *Measures of variability*: summarize the dispersion of the data set with terms like standard deviation, min., max., kurtosis, and skewness.

Table 1 - Descriptive statistics of the variables for each selected country.

Countries	*Variables*	*Mean*	*Kurtosis*	*Skewness.*	*Mn.*	*Mix.*	*Count*
Albania	Mean Temperature	12	-0.6	-0.5	11	13	21
	Per capita CO_2	1	0.3	0.1	1	2	21
	Production index	85	-1.5	-0.2	61	106	21
	Consumer Price Index	4	2.6	1.1	2	8	21
	Undernourishment	6	-0.9	0.8	4.1	8.9	21
	Cereal import	43	-1.5	0.6	37.3	51.7	21
Algeria	Mean Temperature	24	-2	-0.5	23.0	24.0	21
	Per capita CO_2	3	-1	-0.7	2.5	3.9	21
	Production index	85	-1	-0.4	55.9	103.8	21
	Consumer Price Index	5	1	1.0	1.0	12.2	21
	Undernourishment	4	-1	0.4	2.0	8.1	21
	Cereal import	73	0	0.5	68.0	79.7	21
Morocco	Mean Temperature	18.2	0.0	0.2	17	19	21
	Per capita CO_2	1.5	-1.4	0.5	1.3	1.8	21
	Production index	93.6	-0.9	-0.4	69	111	21
	Consumer Price Index	2.4	2.8	1.7	0.60	7.3	21
	Undernourishment	5.0	-1.5	-0.2	3.6	6.3	21
	Cereal import	48.0	-0.1	0.7	36.4	65.7	21
France	Mean Temperature	11.4	0.2	0.3	10.0	13.0	21
	Per capita CO_2	5.7	-1.7	0.2	4.6	6.7	21
	Production index	101	0.3	0.5	92.2	112.0	21
	Consumer Price Index	2.7	1.6	0.8	1.0	5.5	21
	Undernourishment	3.0	1	-1	3.0	3.0	21
	Cereal import	-89.9	-0.3	-0.8	-123	-69.5	21
Croatia	Mean Temperature	11.4	0.2	0.3	10.0	13.0	21
	Per capita CO_2	1.0	-1.1	2.0	0.1	9.3	21
	Production index	-0.7	3.2	1.4	-0.3	2.7	21
	Consumer Price Index	10.0	1113.0	4.2	62.2	1.0	21
	Undernourishment	12.0	1113.0	4.4	90.3	2.6	21
	Cereal import	1.0	-1.2	2.0	0.1	9.3	21
Egypt	Mean Temperature	23.0	18.0	2.1	100.7	12.0	21
	Per capita CO_2	0.0	-0.8	-0.1	0.0	1.4	21
	Production index	22.0	18.0	1.8	91.3	0.9	21
	Consumer Price Index	24.0	18.0	2.4	111.0	38.9	21
	Undernourishment	10.0	-2.9	-1.8	-0.8	3.4	21
	Cereal import	0.01	-0.5	-0.1	0.0	1.4	21

Countries	*Variables*	*Mean*	*Kurtosis*	*Skewness.*	*Mn.*	*Mix.*	*Count*
Syria	Mean Temperature	18.7	0.3	-0.3	0.0	18	21
	Per capita CO_2	2.3	0.4	-0.9	-0.4	1.2	21
	Production index	118.3	357	-1.0	0.4	88.4	21
	Consumer Price Index	28.2	1406	2.2	1.7	0.4	21
	Undernourishment	11.4	61.6	-0.4	1.0	4.0	21
	Cereal import	34.7	24.3	0.3	1.0	28.8	21
Tunisia	Mean Temperature	20.4	-2.1	0.3	20.0	21.0	21
	Per capita CO_2	2.4	-0.4	0.8	2.1	2.9	21
	Production index	93.7	1.4	-0.4	67.4	116.7	21
	Consumer Price Index	4.8	-0.9	0.0	2.2	7.3	21
	Undernourishment	3.5	-1.5	0.3	2.6	4.5	21
	Cereal import	63.7	-1.2	-0.3	54.5	71.5	21
Greece	Mean Temperature	14.3	652.0	7.5	101.5	2.3	21
	Per capita CO_2	1.0	0.2	0.0	0.5	0.8	21
	Production index	14.0	652.0	5.7	93.6	0.5	21
	Consumer Price Index	15.0	652.0	9.2	109.3	5.4	21
	Undernourishment	14.3	652.0	7.5	101.5	2.3	21
	Cereal import	-1.1	-2.3	-1.8	0.2	-0.8	21
Spain	Mean Temperature	14.0	636.0	6.6	101.5	2.7	21
	Per capita CO_2	4.6	-1.8	0.0	0.2	0.8	21
	Production index	14.0	636.0	5.2	85.4	1.0	21
	Consumer Price Index	15.0	-0.38	1	6	8.1	21
	Undernourishment	21.0	-2.7	-1.8	-0.4	-0.5	21
	Cereal import	4.6	-0.1	0.0	0.2	0.8	21
Italy	Mean Temperature	13.0	832.0	7.2	108	2.1	21
	Per capita CO_2	4.6	-1.6	-0.4	0.3	1.1	21
	Production index	13.0	832	5.7	95	0.6	21
	Consumer Price Index	14.0	832	8.4	124	5.4	21
	Undernourishment	21.0	-2.3	-1.8	-1.1	1.3	21
	Cereal import	4.6	-0.9	-0.4	0.3	1.1	21

Source: Author.

4.2. *Unit root test*

In order to capture the order of integration for the selected variables in this paper, we have chosen to do the unit root test by using both the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests. The results of the unit root test for variable stationarity are shown in Table 2 at both level (I_0) and level (I_1). The results showed that almost all variables are stationary for all selected countries and also for both tests at level (I_1), which means variables are integrated in order one, known as "first differences", except the Consumer Price Index variable is stationary at level (I_0), which means variables are integrated in order zero, known as "no differences".

4.3. *ARDL model estimations*

To estimate the impacts of climate change on food security in Mediterranean countries, we use the ARDL model Autoregressive distributed lag). The author assesses the impacts of climate change (temperature and CO_2 emission per capita) variables on food availability, which was measured by the food production index, and food access, which was measured by food inflation. The results showed that there is a cur-

Table 2 - Unit root results: ADF and PP tests.

Countries	*Variables*	*ADF Level (I_0)*	*ADF Level (I_1)*	*PP Level (I_0)*	*PP Level (I_1)*
Albania	Mean Temperature	-1.396 (0.561)	-5.990* (0.001)	-1.231 (0.639)	-6.665* (0.001)
	Per capita CO_2	-1.995 (0285)	-4.858* (0.001)	-1.996 (0.287)	-5.082 (0.007)
	Production index	-1.282 (0.616)	-3.292* (0.030)	-1.228 (0.641)	-3.408* (0.023)
	Consumer Price Index	-3.439* (0.023)	-	-4.146* (0.004)	-
	Undernourishment	-1.528 (0.518)	-2.447* (0.014)	-1.121 (0.685)	-3.382 (0.015)
	Cereal import	-1.905 (0.323)	-3.201* (0.021)	-1.006 (0.730)	-4.153* (0.027)
Algeria	Mean Temperature	-1.242 (0.634)	-4.358* (0.003)	-1.236 (0.637)	-5.354* (0.004)
	Per capita CO_2	-1.579 (0.473)	-5.302* (0.001)	-1.584 (0.480)	-5.210* (0.001)
	Production index	-1.867 (0.339)	-3.356* (0.026)	-1.861 (0.342)	-3.371* (0.025)
	Consumer Price Index	-3.937* (0.007)	-	-3.894* (0.008)	-
	Undernourishment	-1.418 (0.552)	-5.022* (0.008)	-1.639 (0.445)	-5.071* (0.008)
	Cereal import	-3.452* (0.023)	-	-3.292* (0.029)	-
Morocco	Mean Temperature	-4.282* (0.005)	-	-4.091* (0.006)	-
	Per capita CO_2	-0.411 (0.883)	- 2.383* (0.086)	0.708 (0.989)	-4.493* (0.003)
	Production index	-2.411 (0.151)	-6.234* (0.001)	-2.318 (0.176)	-10.12* (0.000)
	Consumer Price Index	-4.172* (0.004)	-	-4.24* (0.005)	-
	Undernourishment	-1.808 (0.364)	-3.444* (0.01)	-1.295 (0.611)	-3.831* (0.015)
	Cereal import	-0.864 (0.992)	-7.162* (0.000)	-2.069 (0.257)	-7.263* (0.000)
France	Mean Temperature	-0.601 (0.849)	-5.297* (0.005)	-0.121 (0.932)	-5.320* (0.004)
	Per capita CO_2	-0.545 (0.862)	-5.871* (0.001)	-0.413 (0.889)	-5.875* (0.001)
	Production index	-2.435 (0.145)	-8.438* (0.000)	-2.436 (0.146)	-14.64* (0.000)
	Consumer Price Index	-6.217* (0.001)	-	-5.477* (0.003)	-
	Undernourishment	-0.541 (0.862)	-9.746* (0.000)	-1.411 (0.556)	-10.63* (0.000)
	Cereal import	1.213 (0.996)	-7.341* (0.000)	-2.775 (0.079)	-7.988* (0.000)
Croatia	Mean Temperature	-0.930 (0.753)	-6.393 (0.001)	-1.421 (0.601)	- 5.026* (0.008)
	Per capita CO_2	-2.723 (0.086)	-4.319* (0.003)	-2.731 (0.087)	-4.907* (0.001)
	Production index	-2.296 (0.182)	-6.707* (0.000)	-2.297 (0.183)	-7.505* (0.000)
	Consumer Price Index	-3.961* (0.007)	-	-3.967* (0.008)	-
	Undernourishment	-2.001 (0.281)	-4.457* (0.002)	-2.595 (0.190)	-7.228* (0.000)
	Cereal import	-1.488 (0.998)	-3.463* (0.021)	-1.719 (0.999)	3.468* (0.022)
Egypt	Mean Temperature	-7.348* (0.000)	-	-7.348* (0.001)	-
	Per capita CO_2	-0.995 (0.734)	-3.013* (0.047)	-1.059 (0.710)	-3.017* (0.048)
	Production index	-1.048 (0.714)	-4.451* (0.002)	-1.048 (0.714)	-4.473* (0.003)
	Consumer Price Index	-3.080* (0.044)	-	-3.080* (0.044)	-
	Undernourishment	-1.396 (0.563)	-3.767* (0.02)	-1.440 (0.542)	-3.767* (0.02)
	Cereal import	-0.235 (0.918)	-3.031* (0.044)	-0.493 (0.873)	-3.832* (0.001)
Syria	Mean Temperature	-2.396 (0.155)	-6.509* (0.000)	-2.371 (0.161)	-7.975* (0.000)
	Per capita CO_2	-0.253 (0.969)	-4.268* (0.004)	-0.087 (0.956)	-4.348* (0.003)
	Production index	-1.243 (0.633)	-5.287* (0.001)	-1.050 (0.714)	-6.304* (0.001)
	Consumer Price Index	-2.359 (0.164)	-5.281* (0.001)	-2.294 (0.182)	-7.059* (0.000)
	Undernourishment	-2.845 (0.999)	-3.831* (0.014)	-2.201 (0.998)	-3.832* (0.015)
	Cereal import	-2.560 (0.117)	-4.529* (0.002)	-2.475 (0.135)	-5.496* (0.003)
Tunisia	Mean Temperature	-0.809 (0.794)	-4.358* (0.003)	-0.809 (0.795)	-4.358* (0.003)
	Per capita CO_2	-0.663 (0.987)	-4.929* (0.001)	-1.837 (0.994)	-4.938* (0.001)
	Production index	-3.597* (0.015)	-	-6.848 (0.000)	-
	Consumer Price Index	-2.757 (0.082)	-4.932* (0.001)	-2.757 (0.082)	-7.923* (0.000)
	Undernourishment	-1.212 (0.647)	-5.308* (0.001)	-0.937 (0.754)	-5.753* (0.000)
	Cereal import	-2.435 (0.145)	-4.748* (0.002)	-2.465 (0.138)	-4.495* (0.003)

Countries	*Variables*	*ADF Level (I_0)*	*ADF Level (I_1)*	*PP Level (I_0)*	*PP Level (I_1)*
Greece	Mean Temperature	-2.396 (0.155)	-6.508* (0.000)	-2.371 (0.161)	-7.975* (0.000)
	Per capita CO_2	-0.519 (0.982)	-2.637* (0.06)	-1.418 (0.998)	-4.988* (0.000)
	Production index	-0.496 (0.873)	-5.531* (0.003)	-0.179 (0.923)	-5.775* (0.002)
	Consumer Price Index	-3.001* (0.044)	-	-3.413* (0.024)	-
	Undernourishment	-4.929* (0.009)	-	-4.913* (0.009)	-
	Cereal import	-3.345* (0.027)	-	-1.652 (0.486)	-3.398* (0.155)
Spain	Mean Temperature	-0.262 (0.913)	-9.746* (0.000)	-1.397 (0.563)	-10.63* (0.000)
	Per capita CO_2	-0.239 (0.918)	-3.265* (0.031)	-0.371 (0.896)	-3.19 4* (0.036)
	Production index	-2.501 (0.130)	-3.959* (0.039)	-2.466 (0.137)	-9.446* (0.000)
	Consumer Price Index	-2.754 (0.083)	-6.032* (0.001)	-2.754 (0.083)	-7.915* (0.000)
	Undernourishment	-6.062* (0.001)	-	-5.892* (0.001)	-
	Cereal import	-2.216 (0.206)	-5.043* (0.001)	-2.201 (0.211)	-5.035* (0.001)
Italy	Mean Temperature	-0.644 (0.839)	-4.358* (0.003)	-0.644 (0.839)	-4.358* (0.003)
	Per capita CO_2	-0.113 (0.958)	-3.295* (0.029)	-0.005 (0.947)	-3.244* (0.033)
	Production index	-1.595 (0.466)	-5.470* (0.003)	-1.463 (0.531)	-9.170* (0.000)
	Consumer Price Index	-2.779 (0.07)	-5.001* (0.001)	-2.641 (0.101)	-8.094* (0.000)
	Undernourishment	-3.919* (0.007)	-	-3.919* (0.007)	-
	Cereal import	-3.296 (0.999)	-3.8723* (0.006)	-0.691 (0.988)	-3.995* (0.007)

() = the values between parentheses represent the value of prob. (*) at significant level 5%. Source: Author.

Table 3 - ARDL models estimation.

Country	*Variables*	*Food Production Index*		*Food Inflation Rate*	
		B	*Prob.*	*B*	*Prob.*
Albania	Mean Temp.	4.789	0.041	-0.354	0.398
	Per capita CO_2	7.409	0.285	-1.934	0.232
Algeria	Mean Temp.	18.478	0.004	-1.387	0.797
	Per capita CO_2	-9.482	0.284	1.777	0.789
Morocco	Mean Temp.	-5.997	0.031	0.378	0.639
	Per capita CO_2	3.421	0.938	-8.188	0.043
France	Mean Temp.	-3.479	0.037	-0.751	0.041
	Per capita CO_2	0.043	0.993	-1.131	0.045
Croatia	Mean Temp.	0.587	0.950	-1.991	0.043
	Per capita CO_2	2.924	0.851	6.945	0.022
Egypt	Mean Temp.	5.241	0.081	-4.333	0.551
	Per capita CO_2	10.394	0.039	21.641	0.023
Syria	Mean Temp.	4.956	0.028	4.089	0.842
	Per capita CO_2	-3.995	0.741	-46.098	0.014
Tunisia	Mean Temp.	7.277	0.029	-2.214	0.041
	Per capita CO_2	41.56	0.044	4.488	0.011
Turkey	Mean Temp.	0.346	0.843	-1.736	0.574
	Per capita CO_2	8.964	0.042	4.771	0.008
Greece	Mean Temp.	0.346	0.843	-1.736	0.574
	Per capita CO_2	8.964	0.042	4.571	0.041
Spain	Mean Temp.	-0.749	0.928	1.464	0.364
	Per capita CO_2	-5.583	0.012	2.331	0.038
Italy	Mean Temp.	1.331	0.754	-0.928	0.373
	Per capita CO_2	-1.938	0.739	0.841	0.035

Source: Author.

rent negative impact of temperature and/or CO_2 emissions per capita on food availability and access dimensions in which mean temperature as an independent variable shows a significant effect on food production index as it is the dependent variable where is the value of probability is less than 0.05 in most selected Mediterranean counties while CO_2 emissions per capita as an independent variable show a significant effect on both food production index and food inflation where is the value of probability is less than 0.05 in most selected Mediterranean counties due to CO_2 emissions is the main contribution on greenhouse gases effects and driven-force to climate change phenomena (Table 3).

Based on results obtained from Tables 2 and 3, the main finding of this paper is that climate change has a negative impact on both the food security dimensions and the agriculture sector for many Mediterranean countries, and therefore, our findings show that they are in synergy with what has been published on the topic in the last decade, like Ehab (2023), Antonelli (2022), Del Pozo (2019), Sameh (2015), and Jeder (2011).

4.4. *Potential impact of climate change by 2070*

The potential impacts of climate change on the agriculture sector of selected Mediterranean countries were estimated and measured by temperature since it has effects on both food availability and access dimensions by 2070. Food production will be decreasing from 11% to 37%, while food prices will be increasing from 10% to 24%. We expect a population growth of 57% in 2070 compared to 2022 and a rise in temperature from 1.0 to 2.5 Celsius, and POU will increase by 1% every 10 years. All those impacts will take place without an adaptive strategy and only during the next 45 years. So, there is a negative relationship between population growth and food availability, while there is a positive relationship between population growth and food access. Moreover, there is a negative relationship between temperature and food availability, while there is a positive relationship between temperature and food access. Therefore, climate change has a negative impact on the agriculture sector of selected countries in terms of both food availability and access.

5. Recommendations and conclusions

Climate change is a complex phenomenon, and the agricultural sector is the most vulnerable to its effects since climate change affects negatively the food security dimensions, which are directly associated with agriculture. The main climate change variables (temperature and CO2 emissions) influencing food production index and food inflation rate in Mediterranean countries have been examined in this paper for the following twelve countries: Algeria, Morocco, France, Croatia, Egypt, Syria, Tunisia, Turkey, Greece, Spain, and Italy. Equipped with simple and advanced statistics and econometric instruments like descriptive statistics, unit root tests, ARDL, and forecasting models, time series data was used on the actual period from 2002 to 2022, while it was used on the forecast period from 2023 to 2070. The results show that climate

Table 4 - Forecasting the impact of climate change on agriculture sector of selected Mediterranean countries.

Year	** Population of Selected countries (No.)*	*Yearly Change on population (%)*	*Temperature Scenario (°C)*	*Impact on Food Production (%)*	*Impact on Food Prices (%)*	*Prevalence of Undernourishment (%)*
2030	555,130,470	1.24	1.0	-9%	7%	5%
2040	621,360,990	1.19	1.2	-11%	10%	6%
2050	687,591,510	1.13	1.5	-17%	13%	7%
2060	774,448,568	1.08	1.7	-24%	17%	8%
2070	861,305,626	1.02	2.0	-37%	24%	9%

(*) Selected Mediterranean countries: Albania, Algeria, Morocco, France, Croatia, Egypt, Syria, Tunisia, Turkey, Greece, and Spain. Source: Author, results produced by EViews (version 20) software.

change has a negative impact on the agriculture sector for many Mediterranean countries in both the current situation and future scenarios. Both temperature and CO_2 emissions have a significant effect on food production and have a negative relationship with it, as well as a significant effect on food inflation and a positive relationship with it in those countries. Moreover, combining climate change with a rapid increase in population growth rate will put more pressure on the agriculture sector, which will accelerate the decreasing rate of food production and increase the rate of food inflation, causing an increase in POU among the people of those countries. The findings confirm the interlinkage between sustainable development goal number three (ending hunger and ensuring food security) and sustainable development goal number thirteen (climate action), in which climate indicators affect food availability by decreasing agriculture productivity, leading to a decrease in food supply, resulting in negative effects on the dimension of food access that are due to a shortage in food supply facing an increase in food demand.

The governments, policymakers, and stakeholders in those Mediterranean countries must pay more attention to the negative impact of climate change on their agriculture sector before it is too late, at which point the costs (economically, socially, and environmentally) will be destructive and irreversible. Therefore, we are recommending the following: (1) The Mediterranean governments must invest more in rural development to increase the food availability dimension and create more economic opportunities to increase the food access dimension; (2) there is a necessity to design and apply new national adaptation strategy targeting the reduction of climate change impacts on the agriculture sector and this adaptation strategy to be based on different climate change pathways and according to each Mediterranean country scenario; (3) establishing and activating a different comprehensive legal framework in those countries aims to increase cooperation between all stakeholders, including both governmental and non-governmental organizations, to raise awareness of the climate change effects and find effective solutions to reduce the expected climate change impacts on the agriculture sector; (4) The Mediterranean countries should invest more in new adaptation options such as clean technologies (environment-friendly) like renewable energy (RE) technologies and other agriculture adaptation technologies (i.e., green ones), with a focus also on training stakeholders on using them to reduce GHG emissions. (5) Reactivating and revitalizing the role of agricultural research institutes in Mediterranean countries, particularly the low-income ones, to improve soil fertility and crop traits, manage water and soil resources, and provide crop protection and insurance to farmers against climate change effects; and (6) Mediterranean countries must establish strong and high levels of cooperation with each other to devise creative adaptation strategies and options against climate change risks, especially in the agriculture field, to reduce potential negative impacts.

References

Antonelli M., Basile L., Gagliardi F., Isernia P., 2022. The future of the Mediterranean agri-food systems: Trends and perspectives from a Delphi survey. *Land Use Policy*, 120, 106263. DOI:10.1016/j.landusepol.2022.106263.

Del Pozo A., Brunel-Saldias N., Engler A., Ortega-Farias S., Acevedo-Opazo C., Lobos G., Jara-Rojas R., Molina-Montenegro M., 2019. Climate Change Impacts and Adaptation Strategies of Agriculture in Mediterranean-Climate Regions (MCRs). *Sustainability*, 11, 2769. https://doi.org/10.3390/su11102769.

Edame G.E., 2011. Climate change, food security and Agricultural productivity in Africa: Issues and policy directions. *International Journal of Humanities and social science*, 1(21). https://www.researchgate.net/publication/256401074_Climate_Change_Food_Security_and_Agricultural_Productivity_in_Africa_Issues_and_Policy_Directions.

Engle R.F., Granger C.W., 1987. Co-integration and error correction: representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*: 251-276.

FAO, 1996. *World food summit: Rome Declaration on world food security and world food summit plan of Action*. Rome. http://www.fao.org/3/w3613e/w3613e00.htm.

FAO, 2006. *Food and Agriculture Organization: food security, policy brief*, June 2006, Issue 2. https://

www.fao.org/fileadmin/templates/faoitaly/documents/pdf/pdf_Food_Security_Cocept_Note.pdf.

FAO, 2013. *New approaches to the measurement of food security*. https://www.fao.org/fileadmin/templates/ess/documents/afcas23/Presentations/AFCAS_9d_New_approaches_to_the_measurement_of_food_security.pdf.

FAO, 2018. *How climate change may affect global food demand and supply in the long term*, Rome: FAO, Global perspective studies team. https://www.fao.org/climatechange/42904-061f6861c5c2331aa1eafe445e822d498.pdf.

GFSI (Global Food Security Index), 2022. *Countries – overall food security score*. https://impact.economist.com/sustainability/project/food-security-index/.

GHI (Global Hunger Index), 2023. *Scores of coastal countries by 2023*. https://www.globalhungerindex.org/ranking.html.

Hashem E.A., 2020. The Impacts of Climate Change on Food Security - Case Study: Egypt. *Journal of Economics and Business*, 3(2): 868-884. https://doi.org/10.31014/aior.1992.03.02.244.

Hertel T.W., Rosch S.D., 2010. Climate Change, Agriculture, and Poverty. Applied Economic Perspectives and Policy, 32: 355-385. https://doi.org/10.1093/aepp/ppq016.

Ibrahim E.A., 2023. The Impact of Climate Change on Food Security Dimensions in Egypt by 2070. *New Medit*, 22(2). https://Doi.Org/10.30682/Nm2302i.

IPCC, 2012. Glossary of terms. In: Field C.B., Barros V., Stocker T.F., Qin D., Dokken D.J., Ebi K.L., Mastrandrea M.D., Mach K.J., Plattner G.-K., Allen S.K., Tignor M., Midgley P.M. (eds.), *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge-New York: Cambridge University Press, pp. 555-564.

Ivanic M., Martin W., 2008. Implications of Higher Global Food Prices for Poverty in Low-Income Countries. *Agricultural Economics*, 39: 405-416.x. https://www.researchgate.net/publication/227811263 Implications_of_Higher_Global_Food_Prices_for_Poverty_in_Low-Income_Countries.

Jeder H., Abdelhamid A., Salah A., 2021. Smallholder Farmers' Perceptions and Adaptation Strategies to Mitigate the Effect of Climate Change in the Oases of South-Eastern Tunisia. *Mew Medit*, 20(5): 3-16. https://doi.org/10.30682/nm2105a.

Jobbins G., Henley G., 2015. *Food in an uncertain future: the impacts of climate change on food security and nutrition in the Middle East and North Africa*. Cairo-London: World Food Programme-Overseas Development Institute. https://www.researchgate.net/publication/284727004_Food_in_an_uncertain_future_the_impacts_of_climate_change_on_food_security_and_nutrition_in_the_Middle_East_and_North_Africa.

Lloyd S.J., Kovats R.S., Chalabi Z., 2011. Climate change, crop yields, and undernutrition: development of a model to quantify the impact of climate scenarios on child undernutrition. *Environ Health Perspect*, 119(12): 1817-1823. DOI: 10.1289/ehp.1003311. Epub 2011 Aug 15. Erratum in: *Environ Health Perspect*, 123(5): A118. PMID: 21844000; PMCID: PMC3261974.

Loks N.A., Umar A.M., Mamzing D., Akila L.K., Nyazi Majak C., 2015. Potential Impacts of Climate Change on Food Security. *Journal of Agricultural Science and Technology*, 5: 799-810. https://pdfs.semanticscholar.org/6463/5adfcc4f13a6fad071bcdc2de5f010c33ce1.pdf.

Nelson G.C., 2009. *Climate change: Impact on Agriculture and costs of Adaptation*, IFPRI, International food policy Research Institute. https://www.ifpri.org/publication/climate-change-impact-agriculture-and-costs-adaptation.

Nelson G.C., Valin H., Sands R.D., Havlik P., Ahammad H., Deryng D., Elliott J., Fujimori S., Hasegawa T., Heyhoe E., Kyle P., Von Lampe M., Lotze-Campen H., d'Croz D.M., van Meijl H., van der Mensbrugghe D., Müller C., Popp A., Robertson R., Robinson S., Schmid E., Schmitz C., Tabeau A., Willenbockel D., 2014. Climate change effects on agriculture: economic responses to biophysical shocks. *PNAS*, 111(9): 3274-3279.

Perron P., Ng S., 1996. Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *The Review of Economic Studies*, 63(3): 435-463.

Phillips P.C., Ouliaris S., 1990. Asymptotic properties of residual based tests for cointegration. *Econometrica: Journal of the Econometric Society*: 165-193.

Porter J.R., Xie L., Challinor A.J., Cochrane K., Howden S.M., Iqbal M.M., Lobell D.B., Travasso M.I., 2014. Food security and food production systems. In: Field C.B., Barros V.R., Dokken D.J., Mach K.J., Mastrandrea M.D., 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.), *Climate change 2014: impacts, adaptation, and vulnerability*. *Part A: global and sectoral aspects*, Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge-New York: Cambridge University Press, pp. 485-533.

Ray D., Gerber J., MacDonald G. *et al*., 2015. Climate variation explains a third of global crop yield variability. *Nat Commun*, 6: 5989. https://doi.org/10.1038/ncomms6989.

Schmidhuber J., Tubiello F.N., 2007. Global food security under climate change. *PNAS*, 104(50): 19703-19708.

Tirado M.C., Clarke R., Jaykus L.A., McQuatters-Gallop A., Frank J.M., 2010. Climate change and food safety: a review. *Food Research International*, 43(7): 1745-1765. https://www.sciencedirect.com/science/article/abs/pii/S0963996910002231.

UNFCCC, 1992. *United Nations Framework Convention on Climate Change*. Rio de Janeiro, Brazil and New York, United States, 9 May 1992.

USDA, 2015. *Climate change, Global food security and the U.S food system*, Dec. 2015. https://www.usda.gov/oce/energy-and-environment/food-security

WB (World Bank), 2023. *World Development Indicators: Egypt database*. https://databank.worldbank.org/source/world-development-indicators.

WFP, 2009. "Hunger and Markets", World Hunger Series. Rome: WFP, London: Earthscan. https://www.fao.org/fileadmin/templates/ERP/uni/F4D.pdf.

Yassin L., 2016. *Climate change and food security in Egypt*, Master's Thesis, the American University in Cairo, AUC Knowledge Fountain. https://fount.aucegypt.edu/etds/315.

On farm non-agricultural activities: Recent evolution and dynamics in Portugal

António Xavier*, Maria do Socorro Rosário**,
Maria Leonor Silva Carvalho***,
Maria de Belém Costa Freitas****

DOI: 10.30682/nm2404d
JEL codes: Q15, Q24

Abstract

An increasing share of farmer's revenues derives from on-farm non-agricultural activities (OFNAA), which constitute a complement to the farmer's income and can function as a factor for the development of farms, enhancing the endogenous resources of the territories and contributing to the multifunctionality of rural areas. Therefore, it is relevant to understand the importance of these non-agricultural activities in the territory, their diversification, spatial trends at local level and the relation with farm's orientation. This paper intends to analyse the OFNAA, using as object of study the Portuguese municipalities. To analyse the diversification of the OFNAA, a diversification index based on entropy is proposed. The relationships between OFNAA diversification and the farms' technical-economic orientation (TEO) are also analysed using correlation matrixes, while the spatial patterns are studied, using the global Moran I and local Moran-LISA. The results provide important insights of the OFNAA dynamics and diversification. Therefore, this study provides an important tool for policy management and implementation.

Keywords*: On-Farm Non-Agricultural Activities (OFNAA), Diversification Index (DI), Normalised entropy, Global Moran I, Local Moran's I - LISA.*

1. Introduction

Agriculture is typically an activity subject to risk and uncertainty. According to Bairwa *et al*. (2013) "Risk can be defined as imperfect knowledge where the probabilities of the possible outcomes are known, and uncertainty exists when these probabilities are not known". Farmers carry out their activity in a highly uncertain environment. Moschini and Hennessy (2001) state that this uncertainty can have several sources, such as production uncertainty; technological uncertainty, and political uncertainty. In the Mediterranean area, Céu and Gaspar (2024) studied the financial distress of farms cultivating vines and olives and highlighted that the effects of climate change and free trade, imply interventions to enhance the economical context of rural areas and reduce risk. To control risk and reduce uncertainty regarding

* CEFAGE-UE (Center for Advanced Studies in Management and Economics), Universidade de Évora, Portugal.
** Direção de Serviços de Estatística, GPP (Gabinete de Planeamento e Políticas), Lisbon, Portugal.
*** Independent researcher.
**** University of Algarve, Faro, Portugal; MED - Mediterranean Institute for Agriculture, Environment and Development; CHANGE - Global Change and Sustainability Institute.
Corresponding author: mbfreitas@ualg.pt

different factors, farmers can implement different strategies (Boncinelli *et al*., 2017), one of which involves the development of other non-agricultural activities, such as tourism, food production, or renewable energies, among others.

Several investigations carried out at the international level indicate that an increasing part of farm income comes from on-farm non-agricultural activities (OFNAA) (Boncinelli *et al*., 2017), with many empirical studies supporting the contribution of diversification to increase household income and reduce uncertainty (Reardon *et al*., 1992; Mishra and Sandretto, 2001; de Janvry *et al*., 2005; Owusu *et al*., 2011). Besides stimulating family income, OFNAA can provide a substantial contribution to rural development, namely by reducing the level of poverty, especially in developing countries (Lanjouw and Lanjouw, 2001; Lanjouw and Shariff, 2004) and are gaining importance in rural economies. According to Boncinelli *et al.* (2017), based in Eurostat's Farm Structure Survey (2008), 12% of European farms had OFNAA. For the USA, Vogel (2012), cited by Boncinelli *et al.* (2017), provides an estimation that 40% of the total agricultural production originates from OFNAA.

In the last 40 years, the European Union has promoted agricultural diversification through direct support programs at both national and supranational levels (Chaplin *et al*., 2004). On the other hand, as Elisiário (2018) points out, there is a growing demand for "rural" and society expects diverse functions from the rural world, having sparked interest in studying the means of achieving progress in rural areas and how this impacts the farmer's activity. In Portugal, OFNAAs constitute a complement to the farmer's income and can contribute also to the development of agricultural holdings, through the enhancement of traditional activities, using existing endogenous resources, both on the farm itself and in the surrounding territories. The sustainable development of Portuguese agriculture must be related to an equitable development of profitable non-agricultural activities, which can support the development of agricultural holdings and the rural communities where they are located. Moreover, there are also new dynamics on farms, through new activities arising from market demand and new food diets, which give rise to new needs to be met, particularly in the energy and tourism areas, where the demand for clean energy and amenities make the study of OFNAA interesting.

The collection of OFNAA data has been carried out since 1999 in agricultural census, whose definition says that they are on-farm profitable activities, which use the farm's resources, and which can be carried out by family labour and/or hired labour. Studying it over time shows trends that can later be considered by agricultural policy to develop additional income for farmers, creating conditions for improving rural development and the well-being of agricultural entrepreneurs.

However, few studies analyse these activities in Portugal. Elisiário (2018) stands out in a study at regional level, from 1999 to 2016, relating the OFNAA trends to various factors related to "geographical distribution and the physical, economic and labour dimensions" of agricultural holdings. To date, there appears to be no research that analyses the degree of diversification of different Portuguese municipalities in non-agricultural activities and that relates these activities to the dominant farms' technical economic orientation (TEO). However, the Agricultural Census of 1999, 2009, and more recently of 2019, provide information regarding OFNAA with a spatial disaggregation at the municipality and parish level.

In this context, it was considered pertinent to study the dynamics of OFNAA's evolution in small territorial units, namely municipalities, identifying the differences in terms of diversification. A more diversified OFNAA will imply more resilience.

To measure diversification, entropy provides an interesting tool. The entropy concept, from physics, was introduced by Shannon (1948) in information theory and used by Jaynes (1957) in statistics. Several authors have used the normalised entropy measure (Golan, 1994 cited by Golan *et al*., 1996; Xavier *et al*., 2018), to measure uncertainty. The relation with the TEO, may be analysed with statistical methods such as correlations and the differences in diversification among them may be analysed employing a One-Way ANOVA. Finally, spatial statistics such as Global Moran I (Moran, 1948) and Local Indicator of Spatial Association (LISA) (local Moran's I) (Anselin, 1995) can identify spatial autocorrelation.

This article intends to analyse the OFNAA, disaggregated at the level of municipality, having as object of study the territory of mainland Portugal. More specifically, the objectives are the following: analyse the OFNAA and their dominant categories; create a novel index to study OFNAA diversification; analyse the relation among OFNAA diversification and TEO, and finally to analyse spatial trends, namely the existence of autocorrelation.

In addition to the introduction, the article is organized into 5 more sections, literature review, methodology, empirical implementation, results and discussion, and final conclusions.

2. Literature review

A farm strategy of diversification into non-farm activities can be a form of self-insurance, helping farmers to increase and stabilize their incomes. (Alasia *et al*., 2009; Seng, 2015). In Portugal, only a few recent studies address the issue of OFNAAs. Elisiário (2018) study stands out studying the evolution of OFNAAs by agrarian region from the year 1999 to 2016. The author states that the trends that led to the intensification and specialisation of agricultural production are felt in OFNAAs. In Portugal, diversification based on non-farm activities has declined over the past two decades, coinciding with a trend toward farm specialization. Specialized farms typically focus on one or a few activities, while non-specialized or mixed farms not only diversify their agricultural production but also appear more likely to engage in a wider range of OFNAAs.

The theoretical basis of the farm diversification process has its origins in the farm household model (Mishra *et al*., 2001; Boncinelli *et al*., 2017). Grilli *et al*. (2024) review the determinants of agricultural diversification, and highlight that the number of factors that may influence diversification is large and many models don't integrate all the important factors, possibly due to lack of some important data. Nevertheless, the literature has emphasized the structural factors that shape farm diversification, such as the characteristics of farms, farmers, and households (Meraner *et al*., 2015). McNamara and Weiss (2005) found that young persons will diversify their activities since they have less aversion to risk and old farmers will tend to diversify due to the reduced workload. Lipshits and Barel-Shaked (2021) analyse the importance of policy reforms in the farmers' diversification decision concluding that "younger, educated, and wealthier farmers who are more peripheral, prone towards diversifying income".

McNamara and Weiss (2005) highlighted that on-farm diversification and off-farm labour allocation are related to farm characteristics and that farm size is one important factor to consider with larger farms tending to be more diversified. Chmieliński *et al*. (2023) on the other hand studied the diversification strategies of family farms in Poland. The authors concluded that small farms will tend to diversify, while large farms won't.

Other studies highlight the importance of geographical location.

Meraner *et al*. (2015) analyse the determinants of farm diversification in the Netherlands. The authors found that specialization can impact diversification decisions, since mixed farming systems tend to be diversified. On the other hand, farm size will contribute to diversification, but the authors also highlight that geographical conditions and farm location (namely socio-demographic, economic, and geographical factors) are related to farm diversification.

Boncinelli *et al*. (2017) studied the determinants of farm diversification, using as case study Tuscany, a region in the centre of Italy, concluding that diversification is important in reducing risk and that there is a tendency for farms situated in marginal regions to be more dependent on agriculture as their main source of income.

Pfeifer *et al*. (2009) analyse the importance of location, for agricultural diversification in the Gelderse Vallei area, Netherlands, concluding that landscape conditions may influence farm's diversification.

Zasada (2011) highlights that farmers that develop their activity in peri-urban areas have increased the importance of different activities to answer to the necessities of the urban society. Zasada and Piorr (2015) studied the local conditions for development policy in Brandenburg, Germany. Žibert *et al*. (2021) examined the key

factors and causal relationships involved in the diversification of non-agricultural activities in Slovenia. Their findings suggest that the development of farm tourism activities is influenced by the level of tourism in a particular local. These studies showed that market access affects diversification, with farmers closer to touristic or urban areas easily accessing the market and tending to diversify to satisfy touristic and urban needs (Zasada, 2011; Zasada and Piorr, 2015; Žibert *et al*., 2021).

On the other hand, Bartolini *et al*. (2014) highlights that farmers located in more distant and remote areas will tend to diversify, while being close to "urban areas reduces the probability of diversifying".

Numerous studies calculate technical efficiency and productivity for farms separately according to their specialisation (e.g. Ameur *et al*., 2024; Sintori, 2023; Kashiwagi and Kamiyama, 2023). The Commission Regulation (EC) No. 1242/2008, of 8 December, introduced a new methodology for classifying agricultural holdings in terms of Economic Dimension (DE) and Technical-Economic Orientation (TEO), revoking the Commission Decision 85/377/EEC. The DE of a farm is now based on the Standard Output of its activities (and not, as before, in the Standard Gross Margin), and the TEO is calculated considering the relative contribution of each activity standard output to its total standard output (GPP, 2011).

There is evidence that agricultural productivity also plays a role in OFNAAS availability. Takeshima *et al*. (2018) studied a sample of farms in Nigeria, concluding that an higher agricultural productivity led to the investment of capital and labour in non-agricultural activities, namely linked to agri-environmental concerns. Additionally, higher agricultural productivity can increase the returns on capital and labour invested in non-agricultural activities, potentially boosting the contribution of these activities to economic growth. Seng (2015) studied the consequences of non-agricultural activities on farm household food consumption, concluding that by developing non-agricultural activities farmers had gains in food consumption thus perceiving the benefits of these activities. Concerning farm households, in the rural Himalayas, Scharf and Rahut (2014) investigated the distributional and welfare impacts of engagement in nonfarm work, concluding that by engaging in nonfarm activities, rural farm households make positive gains in per capita food consumption, thus confirming the hypothesis that engagement in nonfarm activities exerts positive effects on household food consumption. Tesfaye and Nayak (2022) analysed the impact of non-agricultural activities on food security of family farms in Ethiopia. The authors concluded that the significant determinants of non-farm income-generating activities are the age of the household head, family size, landholdings, access to extension services, total household income, and membership in agricultural organizations. Khan *et al*. (2024) examines the adoption of renewable energy as a supplemental income source for Pakistani farmers, specifically focusing on solar energy production. The findings indicate a positive correlation between solar energy generation and increased farmer income.

In Europe, several studies were carried out. Chaplin *et al*. (2004) analysed the diversification of family and corporate farms in Central Europe (Czech Republic, Hungary and Poland) and concluded that the diversification is relatively limited and that enterprise diversification by farmers is unlikely to generate a sufficient number of new jobs to address the prevailing high rural unemployment rate. Trnková (2021) studied the socioeconomic relevance of diversification to farm non-agricultural activities in 2018 for 135 European regions in 28 EU countries. The authors concluded that these activities are more important in Central and Northern Europe than in South and South-Eastern Europe. Salvioni *et al*. (2021) discuss the influence of on-farm non-agricultural activities diversification on the financial performance of family farms. The authors use as case study Italy, concluding that diversification strategies have a positive impact on the financial performance of family farms. Tafidou *et al*. (2023) try to determine if diversification of farms by integrating tourism infrastructure has positive effects on farms' income. The authors use a bootstrap regression analysis, implementing the approach in a case study in Greece, and concluding

that tourism increases farms' performance. Rosa and Francescone (2023) analyse the situation of multifunctional farming activities in Italy, which concluded had benefits for farm income. Ohorodnyk and Finger (2024) analyse the agri-tourism in Ukraine, which can play an important role in farm income diversification, and has the potential to generate positive effects in Ukraine not only on the agricultural sector but also on the country's sustainable development. Grillini *et al*. (2023) analyse the impact of agri-turism on traditional agricultural activities. The authors used as a case study the Tyrol-South Tyrol-Trentino Euroregion. The authors conclude that it contributes positively to economic sustainability outcomes, such as increased income, while adversely affecting production (in quantity and value). Agir *et al*. (2023) studied the Turkish farmers' perspectives on the opportunities and challenges of Agrivoltaics, concluding that it constitutes an income opportunity.

Finally, other authors tried to help farmers implementing these activities. Baghernejad *et al*. (2024) use a combined SWOT-AHP-TOWS model to help farmers choosing strategies regarding non-agricultural activities in the Guilan province, Iran.

One important aspect related to diversification of farms' activities, is the diversification of agricultural practices. Zabala *et al*. (2023) carry out a study in which they review crop diversification practices in Europe, concluding that it presents an added value to farms "representing an adaptive management strategy for ecological transition, without compromising economic sustainability". Alcon *et al*. (2024) analyse the costs and benefits of diversified farming systems in Europe and highlight that in the long-term crop diversification has benefits.

These results suggest that the relation among diversification of OFNAAS and farms' technical economic orientation (TEO) is interesting and needed. Particularly in Portugal, and after an analysis by experts, it would be relevant to study the relation among diversification of OFNAAS and the following technical economical orientations: Farms specialized in vegetable productions (FEVP); Farms specialized in livestock breeding (FSLV) and Mixed orientation farms (MOF).

For building the nonfarm activities, a diversification index will be interesting, as it allows easy comparisons among territorial units. Regarding composite indexes, several references were considered in the theoretical background, such as OECD (2008) and Bathei and Štreimikienė (2023). OECD (2008) present a handbook that details methods for creating composite indicators, which include additive aggregation methods, geometric aggregation, and non-compensatory multi-criteria approach (MCA). Bathei and Štreimikienė (2023) provide a careful review of agricultural sustainability indicators identifying 101 indicators from previous studies. Within the scope of entropy that provided a solution to the problem, Golan *et al*. (1996) proposed a generalised maximum entropy methodology for data estimation, including error terms; Xavier *et al*. (2018) used this methodology in goal programming to analyse the sustainability of agricultural farms. Chen and Zhang (2023) use an entropy indicator to study the level of green agricultural development in Mianyang, China.

Portugal is very diverse in what concerns the TEO (Figure 1). This is linked with both the biophysical characteristics of the territory and the land ownership (property dimension and socioeconomic characteristics of the owner).

The Portuguese Agricultural sector and TEO are also related to agricultural policies, namely Common Agricultural Policy (CAP). For a summary of the Common Agricultural Policy History, see for instance European Council of the European Union (2024) or Giuliani and Baron (2023).

Several important marks of this policy are presented as follows. The first one, was the 1992 reform (REG. CEE 1765/92 and 1766/92) which intended to deal with several problems such as agricultural surplus production. This reform favoured the reduction of support to prices and the granting of direct subsidies to farmers to compensate for the drop in prices, tentatively aligned with world prices. These had new obligations relating to environmental protection, as well as incentives to improve the quality of the food produced. The set aside was introduced, which implicated that farmers with an area equivalent to a production of 92 tonnes of cereals per year would be forced to reduce 15% of their arable land in order to benefit from CAP support.

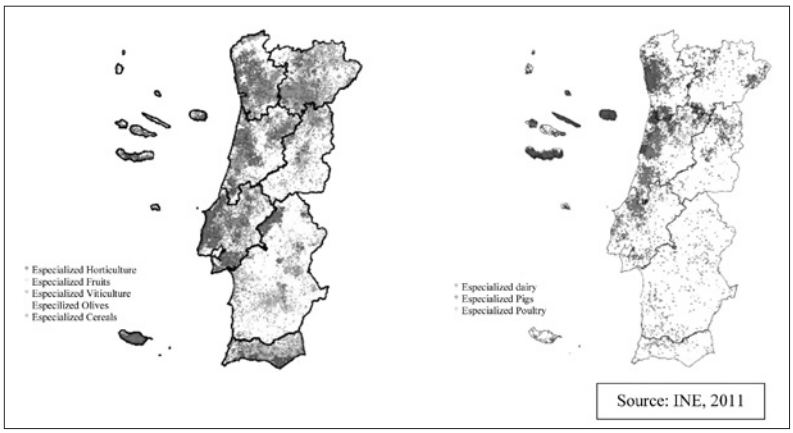


Figure 1 - Technical Economic Orientation of Portuguese farms.

The reform of 1999, considered future expansion of the European Union, was within the scope of "Agenda 2000". This reform had several objectives, among others: improving the competitiveness of community agriculture; defining a rural development policy (2nd pillar of CAP); contributing to animal welfare and plant health; guaranteeing an equitable standard of living for the agricultural population and valuing the multifunctional role of agriculture.

The CAP reform of 2003 was another major step that involved replacing support linked to income with a single farm payment (SFP), disconnected from production and conditioned by standards of food safety, animal and plant health, and animal welfare. Modulation allowed the reduction of direct payments to larger farms as a way of financing rural development.

In 2013, the reform attempted to respond to new challenges to ensure food supply and respond to challenges such as animal welfare, sustainable use of resources, and food security. This reform included the greening of payments, limitations to the support for big farms, and provided additional support for smaller ones (European Council of the European Union, 2024).

Finally, in 2021, the reform "aims to introduce a new strategic approach, giving member states the autonomy to put together strategic plans based on their needs and in line with EU-wide goals" (European Council of the European Union, 2024).

Bearing this in mind, another important issue is the spatial statistics and the relations among the territorial units. The existing of clustering of the specialisation or diversification of OFNAAS implies several different territorial strategies and can be related with the technical orientation of Portuguese farms as well as other biophysical and socioeconomic characteristics. This may be a clustering of high or low values, and its statistical significance has to be tested. For the analysis of spatial autocorrelation, a widely used measure is the global Moran index (I) (Moran, 1948), which evaluates the relationship of spatial interdependence between all polygons in the study area and expresses it through a single value for the study area (Moran, 1950, cited by O'Sullivan and Unwin, 2010; Luzardo *et al*., 2017; Alidadi *et al*., 2023). The global Moran's I analyse spatial autocorrelation for an entire area providing a single value and the local indicators of spatial association (LISA) measure spatial autocorrelation at each location (Anselin, 1995; Fu *et al*., 2014). Several studies have implemented this methodology (Global and Local Moran), from which we present the following examples: Luzardo *et al*. (2017) carried out an analysis of geospatial data associated with area features, in which the variable chosen for the study was the Municipal Human Development Index (HDI-M); Davarpanah *et al*. (2017) presented a study within the scope of geology; Almeida *et al*. (2009) analysed the dengue epidemic concerning the socioeconomic context according to geographic areas; Alidadi *et al*. (2023) studied the spatial distribution of COVID-19 cases in Tokyo; Tang and Werner (2023) analysed the global mining activity.

3. Methodological approach

The methodological approach used in this work follows different steps, as can be seen in Figure 2. The first step concerns the bibliographic review and analysis of available information. The agricultural census of 1999, 2009 and 2019 are the primary source of information. In the second step, the situation in Portugal and Mainland Portugal regarding OFNAAs and their trends throughout the period under study was analysed. The diversification index (DI) was calculated by municipality, in Mainland Portugal, for the years 1999, 2009 and 2019, as well as interannual dynamics considering this index. In a third step the spatial autocorrelation of the index was evaluated using the Global Moran index and the Local Moran I-LISA. A statistical analysis was carried out in which a relationship was established between the farms' orientation and the diversification index and a one-way ANOVA was used to verify if the differences among the different types of farms' orientation are significant.

3.1. *The Diversification Index*

The maximum entropy principle, formulated by Jaynes in 1957, provides a systematic approach to inferring probability distributions from partial knowledge (Golan *et al*., 1996). To measure the discrepancy between two probability distributions, Good (1963) proposed the concept of minimum cross-entropy. Golan *et al*. (1996) developed the Generalised Cross-Entropy (GCE) and the Generalized Maximum Entropy (GME), which considers the unknown distribution and the measurement of errors.

In agricultural censuses there is information, at municipality level, on the number of farms for each OFNAA, however, each farm can have more than one OFNAA. Therefore, the sum of the number of farms in all OFNAA categories will not coincide with the total number of farms. To overcome this problem, the concept of entropy was used, which evaluates the shape of the distribution independently of its numerical quantities. The concept of entropy has been used to construct sustainability indices by several authors, such as Chen and Zhang (2023). While its roots lie in physics, the entropy concept was adapted to the field of information science by Shannon (1948), who used an axiomatic method to define a single function that measures the uncertainty of a set of events.

Thus, the entropy of the probability distribu-

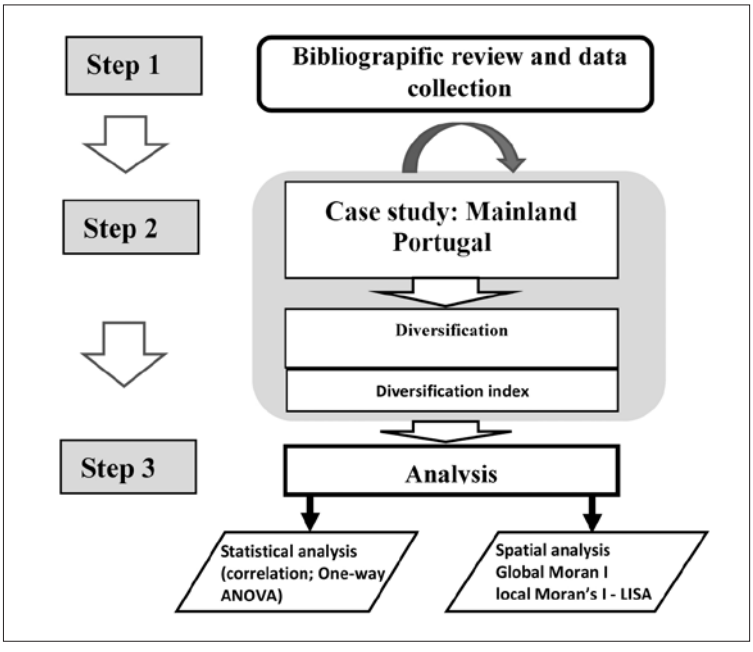


Figure 2 - Methodological approach.

tion, p = (p1, p2, …, pk), reaches a maximum when the probabilities are uniform. To measure the information content of a system and the contribution of each data to reduce uncertainty, several authors have used the normalized entropy measure (Golan, 1994 cited by Golan *et al*., 1996; Xavier *et al*., 2018). In this study, this concept is used to create the proposed diversification index (DI). Following the concept of normalized entropy, the weight of each category in relation to their total sum (and not the actual number of farms) is considered. Therefore, the DI calculation is as follows:

$$DI_i = \frac{-\sum_{i=1}^{I} p_{ik} ln(p_{ik})}{\ln(K)} \quad (1)$$

where DI_i is the diversification index of each unit i considered, p_{ik} corresponds to the weight of each OFNAA category, k, in the total K, in territorial unit i. A DI equal to 1 will imply a uniform distribution of OFNAAs in territorial unit i, that is, maximum diversification. A value of 0 will imply the concentration of OFNAA in a single category, that is, a strong specialization of the territorial unit.

3.2. *Statistical analysis*

Regarding the relationship between DI_i and TEO, correlation matrices were constructed and p values were calculated (Marôco, 2014; Field, 2024). A correlation matrix allows analysing individually the bivariate relations among several variables in the dataset. In this case, the dataset will contain, for all the municipalities, the diversification index and the percentage of each TEO. The results of the matrix contain a correlation coefficient, which in this case is the Pearson correlation coefficient (see Field, 2024 for more information). The correlation does not imply causation, so it is only intended to assess the existing relationship.

For assessing the statistical differences of the diversification among the types of TEO considered, a One-Way ANOVA (Analysis of Variance) was implemented (Marôco, 2014; Field, 2024). A One-Way ANOVA allows comparing the means of multiple groups (categorical variables) and defines if their difference is statistically meaningful. In this case, the mean difference of the *DI* for the groups of municipalities associated with a TEO. The first step of this analysis allows identifying if the mean of two groups is different. To identify the groups, post-hoc tests have to be implemented.

For statistical spatial analysis of the results the Global Moran I and the Local Moran I – Lisa were considered.

Autocorrelation relates to the level of similarity between the values of a variable with spatial references (Davarpanah *et al*., 2017), being based on Tobler's first law of geography (Tobler, 1979): "the pairs of adjacent values or spatial characteristics nearby are probably more similar than the values of more distant territorial units". Autocorrelation may be measured with the global Moran I index (Moran, 1948), which analyses the relationship of spatial interdependence between all polygons in the study area and expresses it through a single value for the entire area (Moran, 1950, cited by O'Sullivan and Unwin, 2010; Luzardo *et al*., 2017; Alidadi *et al*., 2023) and allows identifying the existence of spatial clusters (Davarpanah *et al*., 2018). It tests if the linked areas show more similarity than expected in a random pattern. It ranges from -1 to +1, being positive for direct correlation, and negative when inverse (Almeida *et al*., 2009).

The global Moran's I is then calculated using the following equation (Davarpanah *et al*., 2018; Alidadi *et al*., 2023):

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{(\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}) \sum_{i=1}^{n} (x_i - \bar{x})^2} \quad (2)$$

where n is the total number of municipalities, x_i and x_j are the values of the municipalities i and j, $\bar{x}$ is the average of x_i across the study area, and w_{ij} is the spatial weights matrix that represents proximity (distance) or contiguity relations between a municipality i and its surrounding municipalities j (Davarpanah *et al*., 2018).

After calculating the Moran index, it is important to test its statistical validity using "inferential statistics" to assess if the values represent a statistically significant spatial autocorrelation and are not the result of chance. The null hypothesis (H0) to be tested is "there is no spatial autocorrelation", and relates to a complete spatial

randomness (CSR) of the spatial variable values in geographic space (Davarpanah *et al*., 2018). To accept or reject the null hypothesis (H0), the z score provides a measure for the standard deviation and the p-value indicates the probability of the spatial pattern being created by a random process (the null hypothesis) (Mitchell, 2005; Fox *et al*., 2012; Davarpanah *et al*., 2018).

The local indicator of spatial association (LISA) – local Moran I – measures the level of spatial autocorrelation at each municipality (Anselin, 1995; Feng *et al*., 2014) and (Levine, 2004; Fu *et al*., 2014; Alidadi *et al*., 2023) and can be expressed as:

$$I_i = \frac{z_i - \bar{z}}{\sigma^2} \sum_{j=1, j \neq i}^{n} \left[W_{ij}(z_i - \bar{z}) \right] \qquad (3)$$

where $\bar{z}$ is the mean value of z with the municipality number of n; z_i is the value of the variable at municipality i; zj is the value at other municipalities; σ^2 is the variance of z; and W_{ij} is a distance weighting between z_i and z_j. The weight W_{ij} can also be determined using a distance band. It should be noted that the results of the index are affected by the definition of weight, data transformation, and extreme values (Fu *et al*., 2014).

The local version of the statistics aims to show where spatial patterns are located within the area under study. Thus, spatial units are defined in one of the following four classes (Tartaruga, 2020): high-high: positive autocorrelation, values of the unit variable and neighbours, on average, are high (the set of contiguous areas of this type is called hot spot); low-low: positive autocorrelation, values of the unit variable and neighbours, on average, are low (the set of these areas is known as cold spots); high-low: negative autocorrelation, unit variable value is high, however, those of its neighbours are, on average, low; low-high: negative autocorrelation, unit variable value is low, however, those of its neighbours are, on average, high.

4. Empirical implementation

The study area is Mainland Portugal. The study at the national level characterises the sector, at the level of municipalities in Mainland Portugal, analysing the importance of farms with OFNAAs and their diversification, through ID. The autonomous regions of the Azores and Madeira, although not considered in the spatial analysis, are considered when framing the data.

Figure 3 shows the territory of Mainland Portugal and its municipalities. Mainland Portugal is currently made up of 278 municipalities and has a territory of 88,889 km^2.

The approach was applied to data referring to the last 3 agricultural census: 1999, 2009 and 2019. There are some methodological differences related to changes between 1999 and 2009, through Regulation (EC) No. 1166/ 2008 of the European Parliament and the Council of 19 November 2008, regarding surveys on the structure of agricultural holdings and Agricultural Census. Thus, in 1999, the non-agricultural profitable activity item "Forest production" appears with a

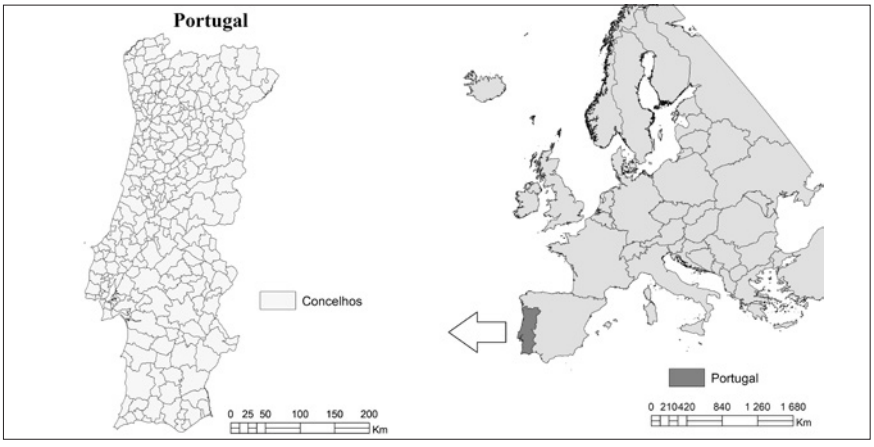


Figure 3 - Geographical location of Portugal Mainland.

value of 0 in all units. There was still no evidence of the notoriety of this variable as its collection was only considered important later, because of afforestation of agricultural land policy. In this situation, the OFNAAs present this year were not collected. Also, the same farm may have more than one OFNAA. Therefore, the sum of partial data will not coincide with the total data regarding the number of farms with OFNAAs.

When calculating the DI, the following OFNAAs were considered, as set out in the Agricultural Census: Rural tourism and directly related activities, Crafts and processing of non-food agricultural products, Processing of agricultural food products, Forestry production, Provision of services, Wood processing, Aquaculture, Production of renewable energy, Other profitable activities.

The proposed diversification index was applied to all municipalities in mainland Portugal. The IDs were classified by classes: 0-0.25; 0.25-0.50; 0.50-0.70; 0.70-0.80 and >0.80, with different amplitudes, resulting from the analysis of histograms of data distribution, to better capture spatial diversity.

To evaluate the relationship between the proportion of farms with OFNAAs and ID with TEO, the following TEO were considered: Farms specialised in vegetable productions (FEPV), Farms specialised in livestock breeding (FSLV) and Mixed orientation farms (MOF). These are general divisions that reflect the main orientations of the territory.

The kurtosis and asymmetry of the data were analysed. It was found that the data on the proportion of farms with OFNAAs had values that showed the existence of non-normal distribution in all years considered.

One requirement of parametric tests is normality. In cases where the distribution is non-normal, data may be transformed, to ensure that the distribution is as close as possible to normality (see Field, 2024 for details).

Therefore, we decided to transform the data for all variables using a base 10 logarithm. Feng *et al*. (2014) state that some of the problems and advantages of using this approach include correcting asymmetric data to achieve normality and such transformation should be used with caution. In cases where the distribution contained 0 values, a parameter common to all municipalities in the variable was placed, taking into account the guidelines of Feng *et al*. (2014). Feng *et al*. (2014) also stated that, as the logarithmic transformation can only be used for positive results, the common practice of adding a small positive constant to all observations before applying this transformation can have a relevant effect on the statistical significance of the hypothesis test.

When applying the global Moran index (I) and the Local Moran I-LISA, an Edges and Corner (Queen Contiguity) spatial continuity matrix was considered and the cells were normalised using the ARCGIS 10.4 software.

Finally, regarding technical implementation, Excel and SPSS were used to perform statistical calculations. The cartography construction was carried out using ARCGIS 10.4.

5. Results and discussion

The results are presented in accordance with the methodological approach previously analysed.

In Table 1, it can be seen that, in Portugal and on the Mainland, the total number of agricultural holdings and the number of holdings with OFNAAs decreased between 1999 and 2019. The proportion of holdings with OFNAAs also shows a decrease in its weight from 1999 to 2009, maintaining its relevance from 2009 to 2019. The generalized decrease in the number of agricultural farms was accompanied by increases in the farms' average area, due to the existence of more competitive and more specialised agriculture.

Table 2 presents the relevance of each of the OFNAA categories in Portugal and on the mainland. It appears that in the two territorial units considered, in 1999, "processing of agricultural food products" dominated, followed by "provision of services" and "other profitable activities". In 2009, the most important OFNAAs concerned "forestry production", followed by "service provision activities", "other profitable activities" and "processing of agricultural food products". Finally, in 2019, OFNAAs relating to "forest production" and "service provision" continued to dominate, with OFNAA, "Rural tourism and directly

Table 1 - Evolution of the OFNAAs in Portugal and Mainland Portugal.

Indicator	*1999*		*2009*		*2019*	
	Portugal	*Mainland*	*Portugal*	*Mainland*	*Portugal*	*Mainland*
Total number of farms (nº)	415969	382163	305266	278114	290229	266039
Number of farms with OFNAAs (nº)	33885	32721	15284	15045	14739	14463
Farms with OFNAAs (%)	8.0	9.0	5.0	5.0	5.0	5.0

Source: INE, National Agricultural Census 1999, 2009 e 2019.

Table 2 - Dynamics of the OFNAAs by category, Portugal and Mainland Portugal.

Activities	*1999*		*2009*		*2019*		*Dynamic % (1999-2009*		*Dynamic % (2009-2019*	
	Portugal	*Mainland*	*Portugal*	*Mainland*	*Portugal*	*Mainland*	*Portugal*	*Mainland*	*Portugal*	*Mainland*
Total	33885	32721	15284	15045	14739	14463	-54.9	-54.0	-3.6	-3.9
Rural tourism and directly related activities	444	418	606	573	1406	1320	36.5	37.1	132.0	130.4
Crafts and processing of non-food agricultural products	369	299	78	71	49	42	-78.9	-7.3	-37.2	-40.8
Processing of agricultural food products	29992	29009	1148	1114	1231	1172	-96.2	-96.2	7.2	5,2
Forestry production	0	0	10842	10836	9816	9809			-9.5	-9.5
Provision of services	2185	2109	1740	1616	1682	1604	-20.4	-23.4	-3.3	-0.7
Wood processing	684	674	118	111	227	216	-82.7	-83.5	92.4	94.6
Aquaculture	32	32	16	16	8	8	-50.0	-50.0	-50.0	-50.0
Production of renewable energy	24	24	101	99	485	482	320.8	312.5	380.2	386.9
Other profitable activities	923	907	1305	1270	1264	1211	41.4	40.0	-3.1	-4.6

Source: INE, National Agricultural Census 1999, 2009 e 2019.

related activities", gaining a prominent position and becoming the third most important.

Regarding the dynamics between 1999, 2009 and 2019 (Table 2), it is worth highlighting the strong growth of "Production of renewable energy", "Rural tourism and directly related activities" and "Wood Processing". It should also be noted that "Processing of agricultural food products" in the last decade showed a slight positive trend. The remainder revealed decreases, in absolute numbers. The high increases seen show the support of both the population in general and the response of farmers in the provision of services, particularly cleaner energy and, on the other hand, the new vision and adherence to the scenic and bucolic environment of the rural landscape. The fluctuations in wood processing are probably related to new environmental and energy concerns and the adaptation of the market to offer suitable products. Regarding

the processing of food products, production has been adjusted to new hygiene, health and environmental rules in the manufacture of food products, which were implemented and closely monitored with the closure of many production units.

The spatial dynamics, in each year, of the proportion of farms with OFNAAs per municipality are presented, for 1999, 2009 and 2019. Figure 4 shows the spatial distribution of the proportion of farms with OFNAAs in the municipalities of the Mainland.

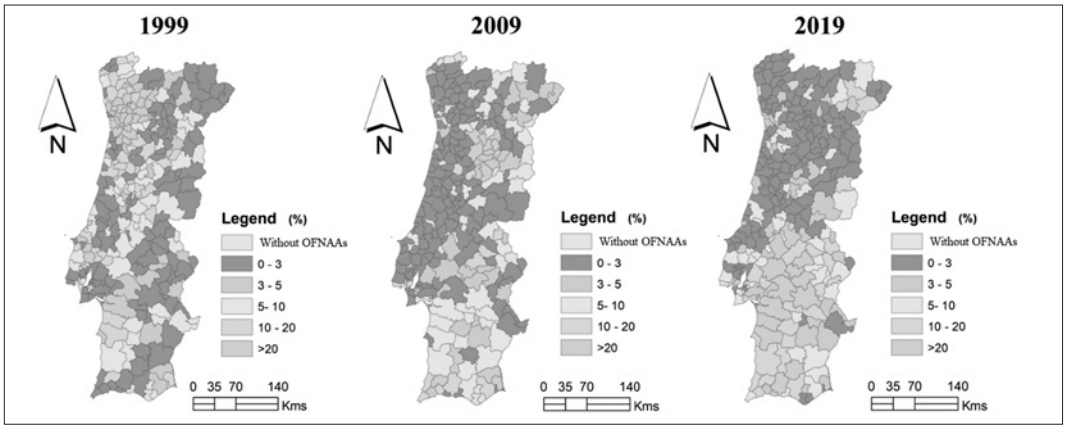
In 1999, the highest values were found in the northwest of Portugal in an area predominantly of smallholdings, where there appears to be a spatial clustering of data, and in the municipalities of the central littoral, where the size of the ownership is also very low (territories with high population density). The lowest values (below 3%) are recorded in large areas throughout the interior. To the south of the Tagus River, this class is predominant in the inland and almost throughout the south littoral. These are areas of low population density and tend to have a lower proportion of farms with OFNAAs. In the area of Lisbon and the Setúbal Peninsula, the reduced number of OFNAAs' proportions is linked to the competition with the industry in that territory.

In 2009, it was observed that the lowest proportions of farms with OFNAAs were found north of the Tagus, in the centre, and north of Portugal. It is also in this area that most municipalities without OFNAAs are located. The majority of municipalities with a proportion of 20% or more of farms with OFNAAs are also located in the previously mentioned area, corresponding to the international Douro territorial zone and the Serra da Estrela region. To the south of the Tagus, values greater than 3% are recorded, which may indicate changes in some policy measures such as decoupling.

Finally, in 2019, it was observed that the highest values for the proportion of farms with OFNAAs were recorded in the south of the Tagus (Alentejo and Algarve), as opposed to the centre and north where the lowest values were recorded. It is also concluded that there was a reduction in the number of farms without OFNAAs. The distribution of the results indicates the possibility of their spatial autocorrelation, and this distribution may be related to the edaphoclimatic and social conditions of the territory, with the OTE of farms and the adoption of policy measures.

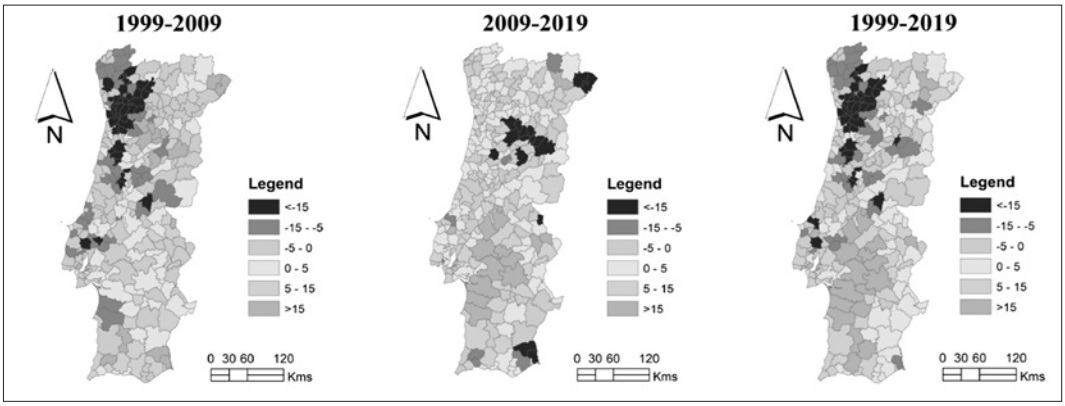
The temporal variation in the proportions of farm values with OFNAAs is shown in Figure 5 for the periods: 1999-2009; 2009-2019 and 1999-2019. Between 1999 and 2009, there was a clear decrease in the proportion of farms with OFNAAs in the municipalities of northwest Portugal, where in 1999, the highest values in the aforementioned series were recorded. It should also be noted that there is a group of nearby municipalities in the south of the Mainland where the proportion of farms with these activities grown. In the period 2009-2019, it appears that

Figure 4 - Spatial distribution of the proportion of farms with OFNAAs.



Source: INE, National Agricultural Census 1999, 2009 e 2019.

Figure 5 - Spatial distribution of the temporal variation of the proportion of farms with OFNAAs.



Source: INE, National Agricultural Census 1999, 2009 e 2019.

positive dynamics tend to be registered in the majority of Portuguese territory, with emphasis on the south of the Tagus River. It is also observed that municipalities in the central inland, which had prominent positive dynamics in the previous period, showed a strongly negative trend, namely in Serra da Estrela and adjacent municipalities, as well as in northeast. Finally, regarding the global evolution from 1999-2019, it is observed that there is a negative evolution, especially in the northwest of Portugal, the north and central coast, while in the south of the Tagus there are positive dynamics, as well as in the centre and interior north.

The diversification index (DI) for the Mainland and Portugal is presented in Figure 6. The analysis shows that the structure in 1999 was highly specialized, and in 2009 and 2019, there was an increase in the diversification of OFNAAs on Portuguese farms. This strong specialization may be related to Common Agricultural Policy (CAP) measures that maintained aid linked to production. In the middle of the decade there were profound changes in the CAP with the single payment regime completely decoupled, leaving the farmers more freedom to choose the possibility of freeing up land and labour for other uses.

Figure 7 presents the spatial dynamics of the proposed activity diversification index.

In 1999, the analysis of the DI shows that the northern municipalities had a specialized structure (vineyards and wine areas and intensive dairy farming). It can also be seen that the areas south of the Tagus River present higher values, which indicate greater diversification of OFNAAs. There is a tendency for several municipalities in the north to have greater specialization in OFNAAs, in contrast to several municipalities in the south of the Tejo River, with the southern regions concentrating diversity (0.25-0.50).

In 2009, several spatial patterns changed. Although the municipalities with the greatest diversification of OFNAAs on farms are located in the south of Portugal, it is observed that the municipalities in the north of Portugal have increased their degree of diversification. The lowest values of the diversification index (DI) are mainly located in the centre.

Figure 6 - Temporal evolution of the diversification index (DI) for Portugal and the Mainland.

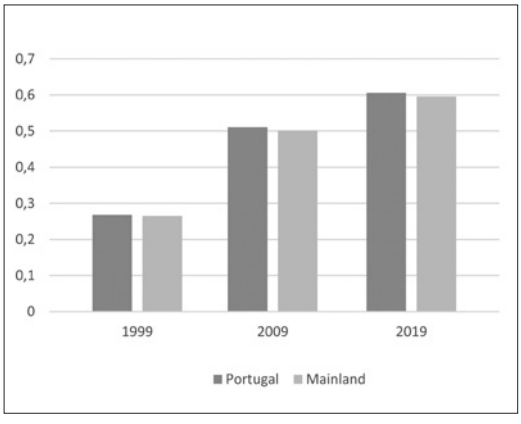
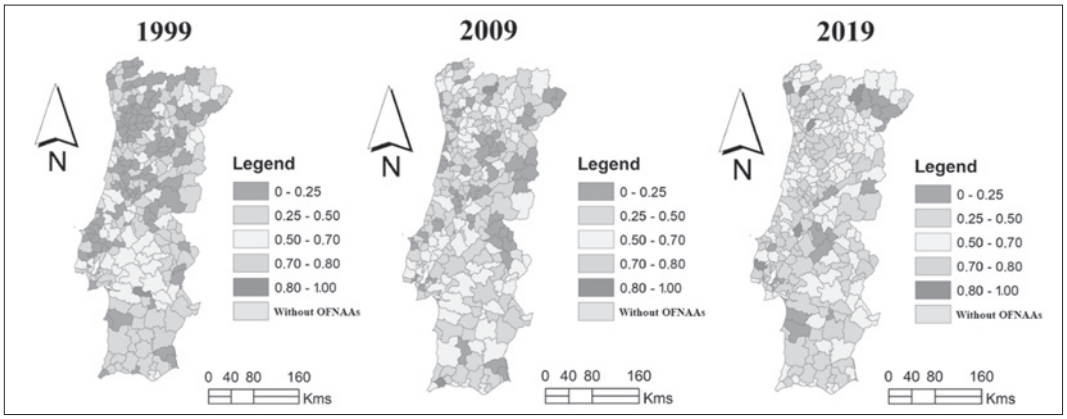


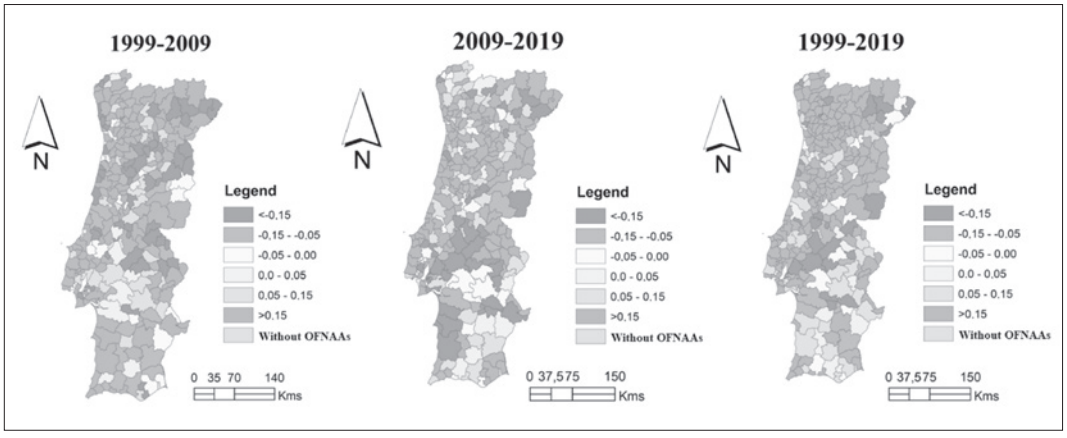
Figure 7 - Spatial distribution of the diversification index (DI).



In 2019, there was an increase in DI in the centre and north of Portugal, with many municipalities with a more diversified structure now located in these areas, that is, with 0.7 or more ID. The south of the Tagus River no longer has a prominent position in ID values and now have a structure that tends to show specialisation in some OFNAAs. Compared to the proportion of farms with OFNAAs, it can be observed that in some more diversified areas they end up being those in which the weight of farms with profitable non-agricultural activities is smaller. This is visible, for example, in 2019, where it was found that the municipalities with the highest proportions of OFNAAs are in the southwest littoral, but that they tend to have lower diversification rates.

Figure 8 shows the dynamics between the 3 time periods under analysis: 1999-2009 and 2009-2019 and 1999-2019 (global dynamics). When analysing these situations, it can be seen that, in the period 1999-2009, positive dynamics were observed throughout Mainland Portugal, with emphasis on the regions south of the Tagus River and northern Portugal (especially the northwest). The municipalities with negative dynamics are mainly located in the northern and central interior, also on the central coast. In the period 2009-2019, positive dynamics were noted, which tended to be located in the centre and north as well as in the interior of some southern regions. Regarding global dynamics 1999-2019, it can be concluded that positive dynamics were recorded throughout the country.

Figure 8 - Spatial distribution of the temporal variation of the diversification index (DI).



5.1. *Statistical analysis of the results*

Before proceeding with data analysis, a study was carried out on the skewness and kurtosis of the diversification index (DI).

Skewness assesses the extent to which a distribution of a variable is symmetrical. A skewness value between −1 and +1 is considered excellent, for assessing normality but a value between −2 and +2 is considered acceptable (Hair *et al*., 2022). Kurtosis analyses the flatness of the distribution: a positive value indicates a distribution more peaked than normal, while a negative kurtosis indicates a flatter shape. If the kurtosis is between -2 and +2 the values are accepted as indicators of normality. When both skewness and kurtosis are close to zero the distribution is considered as normal (Hair *et al*., 2022).

Kline (1998), cited by Marôco (2014), states that parametric test models are robust to asymmetry values below 3 and absolute kurtosis values of 8-10. Thus, according to Marôco (2014), the t test is considered to be robust to violations of normality assumptions when the asymmetry and kurtosis values are not very high.

These measures were calculated for the various logarithmized indicators (using the base 10 logarithm mentioned above) in the years 1999, 2009 and 2019. The results indicate the normality of most variables, with values of kurtosis of asymmetry between -2 and 2, many of which are between -1- and 1. Mixed orientation farms (MOF) present kurtosis values higher than the limit of 2 in all years and asymmetry values higher than the limit of 2 in 1999 and in 2009. Farms oriented to vegetable production (FEPV), also present a kurtosis slightly higher than 2 in 1999. Regarding the normality testing, the Kolmogorov-Smirnov, cannot lead us to accept the null hypothesis of normality. However, and considering Marôco (2014) and Kline (1998) (cited by Marôco (2014)), the application of parametric tests is possible in most of the situations presented.

Thus, correlation matrices were constructed, and the respective p-value calculated, for the 278 municipalities in Mainland Portugal for 1999, 2009, 2019.

The 1999 correlation matrix (Table 3) reveals that between the mentioned TEO and the DI, the correlation coefficients are weak. There is a positive correlation between DI of 0.095 with FEPV and 0.072 with FSLV and a slightly negative correlation with EXM of -0.216 ($p<0.01$). EXM and FSLV have a positive correlation of 0.693 ($p<0.01$).

Table 3 - Correlation matrix for the year 1999.

	FEPV	*FSLV*	*EXM*	*DI*
FEPV	1			
FSLV	0.034	1		
EXM	0.144*	0.693**	1	
DI	0.095	0.072	-0.216**	1

**Correlation is significant at the 0.05 level (two-tailed). **Correlation is significant at the 0.01 level (two-tailed).*

Table 4 - Correlation matrix for the year 2009.

	FEPV	*FSLV*	*EXM*	*DI*
FEPV	1			
FSLV	-0.609**	1		
EXM	-0.721**	0.385**	1	
DI	-0.068	0.085	0.039	1

**Correlation is significant at the 0.05 level (two-tailed). **Correlation is significant at the 0.01 level (two-tailed).*

Table 5 - Correlation matrix for the year 2019.

	FEPV	*FSLV*	*EXM*	*DI*
FEPV	1			
FSLV	-0.670**	1		
EXM	-0.742**	0.493**	1	
DI	0.100	0.036	-0.122*	1

**Correlation is significant at the 0.05 level (two-tailed). **Correlation is significant at the 0.01 level (two-tailed).*

In 2009 (Table 4), it can be concluded that the relationships between the diversification index and TEO are quite weak. There are also notable negative correlations between FEPV and FSLV (-0.609 and $p<0.01$) and EXM (-0.721 and $p<0.01$). Between EXM and FSLV there continues to be a positive correlation (0.385 and $p<0.01$).

In 2019 (Table 5), there was again a weak correlation between TEO and DI, as well as some tendency towards a correlation among TEO.

There is an inverse correlation of -0,122 (p=0.05) between EXM and DI. In TEO, a strong inverse correlation (p<0.01) between FEPV and FSLV and EXM and a positive correlation between EXM and FSLV(p<0.01) can also be observed.

To assess the differences in the diversification index among the several farms' orientations, a One-Way ANOVA was conducted on the years of 1999, 2009 and 2019. The one-way ANOVA compares the means of the diversification index (DI) between the groups (in this case the farms' orientation, which is a categorical variable) allowing to identify if they are statistically significantly different.

Firstly, the data was tested regarding normality and homoscedasticity, which are requirements for this test. For normality, the Kolmogorov-Smirnov was used. Results show that for the 1999 data, we must reject the null hypothesis of normality for all farms' orientation. The skewness and kurtosis show some violation of the values mentioned before (Hair *et al*., 2022), however as mentioned by Marôco (2014), the parametric tests are robust to higher limits of asymmetry and kurtosis. Regarding 2009, for the Kolmogorov-Smirnov test, we accept the null hypothesis of normality for FSLV and MOF and reject it for FEPV. Nevertheless, in FEPV the kurtosis and skewness values are acceptable and are always below 2. Regarding 2019, the Kolmogorov-Smirnov test showed that only for FEPV we must reject the null hypothesis of normality, nevertheless, the skewness and kurtosis are below 2.

Table 6 - Levene statistics of variance homogeneity.

Year	*Levene statistic*	*df1*	*df2*	*Sig.*
1999	0.784	2	264	0.458
2009	1.298	2	275	0.275
2019	2.075	2	270	0.128

The Levene test for the homogeneity of samples is presented next (Table 6). We accept the null hypothesis of Homoscedasticity for all years of the series.

After verifying the ANOVA requirements, the first step is to test the effects of the several technical-economical farms' orientation in the diversification index (DI), therefore we're able to test if the mean of the diversification index is different in at least two groups or not. The results of the one-way ANOVA for 1999 (Table 7) and 2009 (Table 8) show that there is no effect of the several technical-economical farms' orientation in the diversification index (DI). For nominal (categorical) variables, the ANOVA analysis showed that there were no statistically significant relationships between those and the dependent variable and the mean differences of the DI among technical-economical orientations are not significant.

For the year of 2019 (Table 9), the effect of the several technical-economical orientations in the diversification indexes is significant (p=0.01), being the differences between the means statistically significant. These results don't allow identifying the groups in which there are the differ-

Table 7 - One-way ANOVA results for 1999.

DI99	*Sum of squares*	*df*	*Mean square*	*Z*	*Sig.*
Between groups	6.46	2	3.23	2.815	0.062
Within groups	302.941	264	1.148		
Total	309.401	266			

Table 8 - One-way ANOVA results for 2009.

DI09	*Sum of squares*	*df*	*Mean square*	*Z*	*Sig.*
Between groups	0.011	2	0.005	1.2	0.303
Within groups	1.212	275	0.004		
Total	1.223	277			

Table 9 - One-way ANOVA results for 2019.

DI19	*Sum of squares*	*df*	*Mean square*	*Z*	*Sig.*
Between groups	0.032	2	0.016	4.672	0.010
Within groups	0.927	270	0.003		
Total	0.960	272			

Table 10 - One-Way ANOVA post-hoc test results.

Dependent variable			*Average difference (I-J)**	*Std. Error*	*Sig.*	*Confidence interval 95%*	
						Lower limit	*Upper limit*
DMS	1	2	-0.0134	0.0148	0.364	-0.0425	0.0156
		3	0.0029*	0.0103	0.005	0.0087	0.0492
	2	1	0.0134	0.0148	0.364	-0.0156	0.0425
		3	0.0424*	0.0171	0.014	0.0087	0.0760
	3	1	-0.0289*	0.0103	0.005	-0.0492	-0.0087
		2	-0.0424*	0.0171	0.014	-0.0760	-0.0087
Bonferroni	1	2	-0.0134	0.0148	1.000	-0.0490	0.0221
		3	0.0289*	0.0103	0.016	0.0041	0.0538
	2	1	0.3421	0.0148	1.000	-0.0221	0.0490
		3	0.0424*	0.0171	0.042	0.0012	0.0836
	3	1	-0.0289*	0.0103	0.016	-0.0538	-0.0041
		2	-0.0424*	0.0171	0.042	-0.0836	-0.0012

** The mean difference is statistically significant at level 0.05. 1-FEPV; 2-FSLV; 3-MOF.*

ences. So, post-hoc tests have to be carried out. Given the different existing tests, two of them were selected: the DMS and the Bonferroni.

The DMS and the Bonferroni tests were used to identify the differences among groups for the year 2019 (Table 10). The DMS shows that the difference among FEPV and MOF is statistically significant (p=0.05) and the difference between FSLV and MOF (p=0.014) are statistically significant. The Bonferroni shows that the difference among FEPV and MOF (p=0.016) and with FSLV and MOF (p=0.042) are statistically significant.

5.1.1. *Spatial statistics*

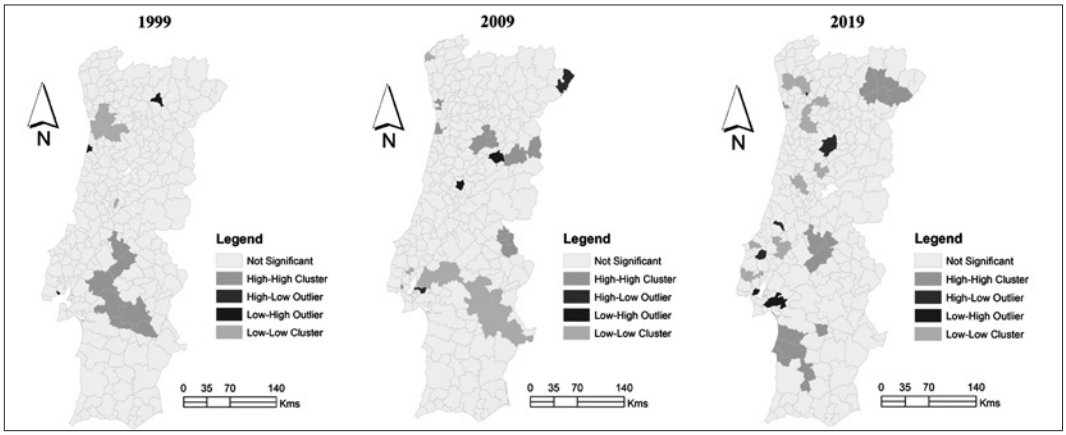
To evaluate spatial autocorrelation, the global Moran index was calculated, together with a significance test. Spatial autocorrelation analysis was applied to the diversification index (DI) in each of the years mentioned. The global Moran Index (I) for the diversification index of OFNAAs is found in Table 11. As mentioned, the

Table 11 - Global Moran I for the DI in 1999, 2009 e 2019.

	1999	*2009*	*2019*
Moran Index	0.19789	0.19996	0.15480
Expected Index	-0.00357	-0.00357	-0.00357
Variance	0.00143	0.00143	0.00143
z-score	5.32962	5.38537	4.19422
p-value	0.00000	0.00000	0.00003

global Moran Index I test whether connected areas have greater similarity regarding the indicator studied than expected in a random pattern. The degree of existing autocorrelation can be quantified, ranging from -1 to +1, being positive for direct correlation, negative when inverse. When carrying out the test on Moran's I index, positive values were obtained for the indicator. The Z-score values that determine p-value allow rejecting the null hypothesis at 1% of absence of

Figure 9 - Local Moran I results for the DI.



spatial autocorrelation (i.e., in which this clustering pattern would result from chance). It is therefore important to carry out a more careful spatial analysis looking for spatial factors that can explain these situations of spatial clustering.

Figure 9 presents the Local Moran results. In 1999 a high-high spatial cluster is observed in the central southern area, while a low-low cluster is present at the northwest area of Mainland Portugal. In 2009 we observe a low-low cluster in the central southern area, while a high-high cluster is present at centre of Portugal. Finally, in 2019, a cluster is present at the southern littoral and in the north of Portugal.

5.2. *Discussion*

Non-agricultural profitable activities are a source of additional income for the farmer and their analysis constitutes a relevant aspect when thinking about policies for rural areas and their multifunctionality.

The first aspect to discuss will be the relationship between the proportion of farms with OFNAAs and the diversification index (DI). Regarding the relationship between the diversification index (DI) and the TEO, the research shows that the relationship does not seem to be significant with the general groups of TEO. However, it is important to analyse this relationship within the subgroups of each of the TEO, which was not done. On the other hand, it should be noted that this relationship was limited since these TEO values refer to the total number of farms existing in each municipality and farms with OFNAAs often represent a small percentage of existing farms.

Another aspect to discuss will be the analysis of the identified spatial dynamics conditioning factors. It has been proven that there is a spatial autocorrelation between the results of the proportion of farms with OFNAAs and the DI, however, nothing has been advanced as to why these patterns occur in each year and what factors lead to common geographic groupings, in which there are geographic relationships guided by a clear positive autocorrelation of the results. This may be related to natural conditions or even socioeconomic factors and requires careful spatial analysis, never forgetting Tobler's law "All things are related to everything else, but close things are more related than distant things". A brief discussion is provided on the explaining factors for the diversification index (see Figures 7 and 9).

In 1999, there is a lower DI in northwest Portugal, as well a statistically significant spatial cluster of Low-Low values, which is explained by the fact that there is a specialized farm structure on vineyards and quality wine production, namely of Protected Designation of Origin (PDO) but also in Dairy farming for milk production. In south Portugal, in the northern part of Alentejo Region, there is a High-High Cluster with good values of the Diversification Index (DI). This is related to the fact that, in this year, farmers continued their land uses and activities and there wasn't a special-

ization of the agricultural activity. Farmers were oriented mostly to cereals, tomatoes and industrial crops, which are seasonal and had more free time to carry out other activities in their farms. There were still coupled subsidies, which had a role on farmers options.

In 2009 there are changes linked to decoupling of payments, introduced some years before the single farm payment (SFP). The farmer is free to choose the production and activities that are more favourable. In the South, a low-low cluster was identified in part of the Alentejo, which is mostly related to intensive irrigated crops, in the Alqueva Dam area, that imply specialization and are time intensive. In the north of Portugal there is specialization, but there are improvements in diversification. The liberalization of the Dairy milk production quotas also contributed to the diversification of OFNAA in this area. These and other conditions, together with the biophysical and natural context, created conditions to develop activities such as tourism.

In 2019, the diversification tendency continues in most of the Portuguese territory which is explained by decoupling and the agricultural policy context.

There is also a conceptual aspect that must be underlined when analysing these dynamics, as some statistical concepts changed between 1999 and 2009 and mark some differences in the matrices that were used to calculate the diversification index, and therefore, in its dynamics. These conceptual changes must be taken into account at the level of general dynamics' analysis, but also at the level of spatial patterns that were established in 1999, as opposed to other years.

6. Concluding remarks

This study presented an analysis of non-agricultural activities in mainland Portugal and made it possible to present a preliminary analysis of the situation of these activities at municipality level, including data from the 2019 agricultural census, which until now has not been analysed in previous studies. The structure of activities and the trend of farms at national level were studied. A recent growth trend was noted in the following OFNAAs: Production of renewable energy, Rural tourism and directly related activities and Wood Production. To analyse the diversification of activities, a diversification index was created based on the concept of entropy. This showed that Portugal registered an increasing diversification of OFNAAs on agricultural holdings, which is an important aspect to consider when we think about the multifunctionality of Portuguese agriculture.

A territorial analysis was carried out in the municipalities of Portugal, studying the dynamics of the proportion of farms with OFNAAs and their diversification over time. Regarding the proportion of OFNAAs in total farms, it appears that positive dynamics tend to be registered in the majority of Portuguese territory in the most recent period of 2009-2019, with emphasis on the south of the Tagus River. It is also observed that municipalities in the inland centre of Portugal that had prominent positive dynamics in the previous period showed a strongly negative trend. Regarding the diversification index, it is concluded that between 2009-2019 positive dynamics tend to be located in the centre and north.

An attempt was made to establish a relationship between DI and the proportion of farms with OFNAAs with TEO, using correlation matrices. Clustering and spatial grouping were also analysed using the global Moran I index, for the DI and proportion of farms with OFNAAs, where the existence of spatial clustering was concluded, and a hypothesis test was developed that proved that such positive spatial autocorrelation does not result from chance.

Therefore, based on the results obtained in this study, the following lines of investigation are now established:

- The first line of investigation will be the study of these activities including detailed data that allows their characterisation. The present data is aggregated, meaning it is not possible to have detailed information to characterise farms in terms of their TEO, main uses and crops, agricultural population and other relevant indicators.
- The second line of investigation concerns the analysis of spatial patterns and conditioning factors. The global Moran index indicates the existence of spatial autocorrelation. Such groupings of data dictate often common behaviours that must be better understood.

- A third line of research concerns the creation of a typology of farms taking into account their orientation towards the different OFNAAS, as well as analysing the relationship between the OFNAAs existing on the farms.

Funding

This research was funded by National Funds through Fundação para a Ciência e Tecnologia under the project grants number UIDB/05183/2020 and UID/04007/2020.

References

Agir S., Derin-Gure P., Senturk B., 2023. Farmers' perspectives on challenges and opportunities of agrivoltaics in Turkiye: An institutional perspective. *Renewable Energy*, 212: 35-49.

Alasia A., Weersink A., Bollman R., Cranfield J., 2009. Off-farm Labour Decision of Canadian Farm Operators: Urbanization Effects and Rural Labour Market Linkages. *Journal of Rural Studies*, 25(1): 12-24.

Alcon F., Albaladejo-García J.A., Martínez-García V., Rossi E.S., Blasi E., Lehtonen H., Martínez-Paz J.M., Zabala J.A., 2024. Cost benefit analysis of diversified farming systems across Europe: Incorporating non-market benefits of ecosystem services. *Science of the Total Environment*, 912: 169272.

Alidadi M., Sharifi A., Murakami D., 2023. Tokyo's COVID-19: An urban perspective on factors influencing infection rates in a global city. *Sustainable cities and society*, 97: 104743.

Almeida A.S.D., Medronho R.D.A., Valencia L.I.O., 2009. Análise espacial da dengue e o contexto socioeconômico no município do Rio de Janeiro, RJ. *Revista de Saúde Pública*, 43: 666-673.

Ameur F., Oulmane A., Boudedja K., Bouzid A., Benmehaia M.A., 2024. Assessing technical efficiency and its determinants for dairy cattle farms in northern Algeria: The two-step DEA-Tobit approach. *New Medit*, 23(1): 129-142.

Anselin L., 1995. Local indicators of spatial association - LISA. *Geographic Analysis*, 27(2): 93-115.

Baghernejad J., Sabouri M.S., Shokati Amghani M., Norozi A., 2023. Developing strategies for stabilizing the livelihood of smallholder farmers through non-farm activities: the application of the SWOT-AHP-TOWS analysis. *Frontiers in Sustainable Food Systems*, 7: 1199368.

Bairwa S.L., Kushwaha1 S., Bairwa S., 2013. Agricultural Education, Research and Extension in India. Chittupure, Varanasi: Poddar Publication, 221005. Ed. Goyal *et al*. https://www.researchgate.net/publication/303224394_MANAGING_RISK_AND_UNCERTAINTY_IN_AGRICULTURE_-_A_REVIEW [accessed: Jan 02, 2024].

Bartolini F., Andreoli M., Brunori G., 2014. Explaining determinants of the on-farm diversification: empirical evidence from Tuscany region. *Bio-based and Applied Economics*, 3: 477-493.

Bathaei A., Štreimikienė D., 2023. A systematic review of agricultural sustainability indicators. *Agriculture*, 13(2): 241.

Boncinelli F., Bartolini F., Casini L., Brunori G., 2017. On farm non-agricultural activities: Geographical determinants of diversification and intensification strategy. *Letters in Spatial and Resource Sciences*, 10: 17-29.

Céu M.S., Gaspar R.M., 2024. Financial distress in European vineyards and olive groves (No. 2023/0266). New Medit, 1(23): 31-53.

Chaplin H., Davidova S., Gorton M., 2004. Agricultural adjustment and the diversification of farm households and corporate farms in Central Europe. *Journal of rural studies*, 20(1): 61-77.

Chen C., Zhang H., 2023. Evaluation of green development level of Mianyang agriculture, based on the entropy weight method. *Sustainability*, 15(9): 7589.

Chmieliński P., Pawłowska A., Bocian M., 2023. On-farm or off-farm? Diversification processes in the livelihood strategies of farming families in Poland. *Social Sciences & Humanities Open*, 8(1): 100575.

Davarpanah A., Babaie H.A., Dai D., 2018. Spatial autocorrelation of neogene-quaternary lava along the Snake River plain, Idaho, USA. *Earth Science Informatics*, 11: 59-75.

de Janvry A., Sadoulet E., Zhu N., 2005. *The Role of Non- Farm Incomes in Reducing Rural Poverty and Inequality in China*. UC Berkeley Department of Agricultural and Resource Economics, UCB. CUDARE Working Paper No. 1001. http://escholarship.org/uc/item/7ts2z766.

Elisiário R.M., 2018. *Multifuncionalidade das explorações agrícolas no desenvolvimento rural*, Dissertação para obtenção do Grau de Mestre em Engenharia Agronómica (Sem Área de Especialização), Instituto Superior de Agronomia, Universidade de Lisboa.

European Council of the European Union, 2024. *Timeline - History of the CAP*. Accessed online 30 August 2024: https://www.consilium.europa.eu/en/policies/cap-introduction/timeline-history-of-cap/.

Feng C., Wang H., Lu N., Chen T., He H., Lu Y.,

Tu X.M., 2014. Log-transformation and its implications for data analysis. *Shanghai Arch Psychiatry*, 26(2): 105-109. DOI: 10.3969/j.issn.1002-0829.2014.02.009.

Field A., 2024. *Discovering statistics using IBM SPSS statistics*. Thousand Oaks, CA, USA: Sage.

Fox E., Balram S., Dragicevic S., Roberts A., 2012. Spatial analysis of high resolution aerial photographs to analyze the spread of Mountain pine beetle infestations. *Journal of Sustainable Development*, 5(9): 106.

Fu W.J., Jiang P.K., Zhao K.L., 2014. Using Moran's I and GIS to study the spatial pattern of forest litter carbon density in a subtropical region of southeastern China. *Biogeosciences*, 11: 2401-2409.

Gabinete de Planeamento e Políticas (GPP), 2011. *Evolução da tipologia comunitária das explorações agrícolas*, Gabinete de Planeamento e Políticas, Lisbon.

Giuliani A., Baron H., 2023. The CAP (Common Agricultural Policy): A Short History of Crises and Major Transformations of European Agriculture. *Forum for Social Economics*, 1-27. DOI: 10.1080/07360932.2023.2259618.

Golan A., Judge G., Miller D., 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York, USA: John Wiley & Sons.

Good I., 1963. Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables. *The Annals of Mathematical Statistics*, 34(3): 911-934.

Grilli G., Pagliacci F., Gatto P., 2024. Determinants of agricultural diversification: What really matters? A review. *Journal of Rural Studies*, 110: 103365.

Grillini G., Sacchi G., Streifeneder T., Fischer C., 2023. Differences in sustainability outcomes between agritourism and non-agritourism farms based on robust empirical evidence from the Tyrol/Trentino mountain region. *Journal of Rural Studies*, 104: 103152.

Hair J.F., Hult G.T.M., Ringle C.M., Sarstedt M., 2022. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) (3 ed.). Thousand Oaks, CA, USA: Sage.

INE - Instituto Nacional de Estatística, 2011. *Recenseamento geral da agricultura de 2009*, Instituto Nacional de Estatística, Lisboa, Portugal.

INE - Instituto Nacional de Estatística, 2021. *Recenseamento Agrícola. Análise dos principais resultados: 2019*. Lisboa, ISBN 978-989-25-0562-6.

Jaynes E.T., 1957. Information theory and statistical methods. *I. Phys. Rev*., 106: 620-630.

Lanjouw J.O., Lanjouw P., 2001. The Rural Non-Farm Sector: Issues and Evidence from Developing Countries. *Agricultural Economics*, 26(2001): 1-23.

Kashiwagi K., Kamiyama H., 2023. Effect of adoption of organic farming on technical efficiency of olive-growing farms: empirical evidence from West Bank of Palestine. *Agricultural and Food Economics*, 11(1): 26.

Khan N., Elhindi K.M., Kassem H.S., Kazim R., Zhang S., 2024. Unveiling the nexus between solar energy adoption and crop farmer income: evidence from Pakistan. *Frontiers in Sustainable Food Systems*, 8: 1364040.

Lanjouw P., Shariff A., 2004. Rural Non-Farm Employment in India: Access, Income and Poverty Impact. *Economic and Political Weekly*, 39(40): 4429-4446.

Levine N., 2004. *CrimeStat III: a spatial statistics program for the analysis of crime incident locations*. Ned Levine & Associates, Houston, TX., and the National Institute of Justice, Washington, DC.

Lipshits R., Barel-Shaked S., 2021. Policy reforms in agriculture and farmer's income diversification decision: The case of eggs farms. *New Medit*, 20(2): 65-78.

Luzardo A.J.R., Castañeda Filho R.M., Rubim I.B., 2017. Análise espacial exploratória com o emprego do Índice de Moran. *GEOgraphia*, 19(40): 161-179.

Marôco J., 2014. *Análise Estatística com o SPSS statistics*. Lisboa, Report Number.

McNamara K.T., Weiss C., 2005. Farm household income and on-and off-farm diversification. *Journal of Agricultural and Applied Economics*, 37: 37-48.

Meraner M., Heijmanb W., Kuhlmanc T., Finger R., 2015. Determinants of farm diversification in the Netherlands. *Land Use Policy*, 42: 767-780.

Mishra A., Sandretto C., 2001. Stability of Farm Income and the Role of Nonfarm Income in U.S. Agriculture. *Review of Agricultural Economics*, 24(1): 208-221.

Mitchell A., 2005. *The ESRI guide to GIS analysis*, vol 2. Environmental Systems Research Institute (Redlands).

Moran P.A.P., 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society*, Series B, 10: 243-251.

Moschini G., Hennessy D.A., 2001. Uncertainty, risk aversion and risk management for agricultural producers. In: Gardner B., Rausser G. (eds.), *Handbook of Agricultural Economics*, vol. 1, pp. 88-153. Amsterdam: Elsevier.

OECD, Organization for Economic Co-operation and Development - JRC, Joint Research Centre, 2008.

Handbook on constructing composite indicators. Methodology and user guide. OECD, Paris.

Ohorodnyk V., Finger R., 2024. Envisioning the future of agri-tourism in Ukraine: from minor role to viable farm households and sustainable regional economies. *Journal of Rural Studies*, 108: 103283.

O'Sullivan D.; Unwin D.J., 2010. *Geographic information analysis*. New Jersey: John Wiley and Sons, 405 pp.

Owusu V., Awudu A., Seini A., 2011. Non-Farm Work and Food Security among Farm Households in Northern Ghana. *Food Policy*, 36(2): 108-118.

Pfeifer C., Jongeneel R.A., Sonneveld M.P., Stoorvogel J.J., 2009. Landscape properties as drivers for farm diversification: a Dutch case study. *Land Use Policy*, 26: 1106-1115.

Reardon T., Delgado C., Matlon P., 1992. Determinants and Effects of Income Diversification Amongst Farm Households in Burkina Faso. *Journal of Development Studies*, 28(2): 264-296.

Reg. (CE) n. 1166/2008. Available at: https://eur-lex.europa.eu/legal-content/PT/ALL/?uri=CELEX-:32008R1166.

Rosa M., Francescone M., 2023. The Evolution of Multifunctional Agriculture in Italy. *Sustainability*, 15: 11403. https://doi.org/ 10.3390/su151411403.

Salvioni C., Henke R., Vanni F., 2020. The impact of non-agricultural diversification on financial performance: Evidence from family farms in Italy. *Sustainability*, 12(2): 486.

Scharf M., Rahut D.B., 2014. Nonfarm Employment and Rural Welfare: Evidence from Himalaya. *American Journal of Agricultural Economics*, 96(4): 1183-1197.

Seng K., 2015. The Effects of nonfarm activities on farm households' food consumption in rural Cambodia. *Development Studies Research*, 2(1): 77-89, DOI: 10.1080/21665095.2015.1098554.

Shannon C., 1948. A mathematical theory of communication. *The Bell System Technical Journal*, 27: 379-423.

Sintori A., Konstantidelli V., Gouta P., Tzouramani I., 2023. Profitability, Productivity, and Technical Efficiency of Cretan Olive Groves across Alternative Ecological Farm Types. *Agriculture*, 13(12): 2194.

Tafidou A., Lialia E., Prentzas A., Kouriati A., Dimitriadou E., Moulogianni C., Bournaris T., 2023. Land Diversification and Its Contribution to Farms' Income. *Land*, 12(4): 911. DOI: https://doi.org/10.3390/land12040911.

Takeshima H., Amare M., Mavrotas G., 2018. *The Role of Agricultural Productivity in Non-farm Activities in Nigeria*, IFPRI Discussion Paper 01761.

Tang L., Werner T., 2023. Global mining footprint mapped from high-resolution satellite imagery. *Communications Earth & Environment*, 134. DOI: https://doi.org/10.1038/s43247-023-00805-6.

Tartaruga I.G.P., 2020. A economia metropolitana e os seus impactos regionais em Portugal: uma análise espacial exploratória de Lisboa, Porto e Coimbra. Cuyonomics. *Investigaciones en Economía Regional*, 4(5): 87-103.

Tesfaye T., Nayak D., 2022. Does participation in non-farm activities provide food security? Evidence from rural Ethiopia. *Cogent Social Sciences*, 8(1): 2108230. DOI: 10.1080/23311886.2022.2108230.

Tobler W., 1979. Cellular geography. In: Gale S., Olsson G. (eds.), *Philosophy in geography*. Dordrecht: Reidel, pp. 379-386.

Trnková G., 2021. *The Economic and Social Importance of Farm Diversification towards Nonagricultural Activities in EU*. HED Hradec Edonomic Days. DOI: 10.36689/uhk/hed/2021-01-081.

Vogel S., 2012. *Multi-enterprising Farm Household: the Importance of Their Alternative Business Ventures in the Rural Economy EIB-10*. US Department of Agriculture, Economic Research Services.

Xavier A.S., Freitas M.B.C., Fragoso R., Rosário M.S., 2018. A regional composite indicator for analysing agricultural sustainability in Portugal: a goal programming approach. *Ecological Indicators*, 89: 84-100. https://doi.org/10.1016/j.ecolind.2018.01.048.

Zabala J.A., Martínez-García V., Martínez-Paz J.M., López-Becerra E.I., Nasso M., Díaz-Pereira E., Sánchez-Navarro V., Álvaro-Fuentes J., González-Rosado M., Farina R., Di Bene C., Huerta E., Jurrius A., Frey-Treseler K., Lóczy D., Fosci L., Blasi E., Lehtonen H., Alcon F., 2023. Crop diversification practices in Europe: An economic cross-case study comparison. *Sustainability Science*, 18.6(2023): 2691-2706.

Zasada I., 2011. Multifunctional peri-urban agriculture – a review of societal demands and the provision of goods and services by farming. *Land Use Policy*, 28: 639-648.

Zasada I., Piorr A., 2015. The role of local framework conditions for the adoption of rural development policy: an example of diversification, tourism development and village renewal in Brandenburg. Germany. *Ecological Indicators*, 59: 82-93.

Žibert M., Prevolšek B., Pažek K., Rozman Č., Škraba A., 2021. Developing a diversification strategy of non-agricultural activities on farms using system dynamics modelling: a case study of Slovenia. *Kybernetes*, 51(13): 33-56.

Timely repayment of agricultural loans: Evidence from Serbian farmers

Dejan Mirč*, Marina Ignjatović**, Dragan Milić*,
Vladimir Crnojević***, Nedeljko Tica*

DOI: 10.30682/nm2404e
JEL codes: E5, Q10, Q14

Abstract

Our study aimed to determine factors influencing the timely loan repayment of smallholder farmers. We used data from 1735 liquidated loans, collecting a set of 36 feasible determinant variables. The study was two-folded. In the first step, with a 64% accuracy, a Logit model revealed 18 significant predictors of timely repayment. Previously credited clients, special agricultural accounts, average monthly inflow, loan amount, age when applying for a loan, clean credit history, and no credit in the past have a positive influence. In contrast, the number of transactions, profiling, owned farm area, past due records over five days, tax debt status, and livestock negatively influenced timely repayment. In the second step, we used machine learning algorithms to enhance model prediction performance. XGBoost model has envisioned timely repayment with 92% accuracy. As significant predictors, Shapley's additive explanations identified clean credit history, average monthly inflow, time of owning the account, age when applying for a loan, and horticulture. The study's findings provide insight into the critical factors in substantially achieving a high repayment rate on borrowed funds.

Keywords*: Agricultural loans, Timely repayment, Smallholder farmers, Banks, Machine learning algorithms*

1. Introduction

Agriculture is among the world's largest industries and is paramount for economic prosperity and development. It is described as a backbone and the primary sector of the industry for many countries in the world (Murungi *et al*., 2023; Pejak *et al*., 2022), as well as a critical link in the food supply chain (Đokić *et al*., 2022). Growth in this sector is two to four times more efficient in raising incomes among the poorest than other sectors (World Bank, 2023). In terms of economic growth, according to the World Bank report, agriculture accounts for 4% of global GDP; in some developing countries, it can even account for more than 25% of GDP (World Bank, 2023). Since the world population is projected to increase to nine billion people by the year 2050, specific estimates indicate that agricultural production will have to increase by at least 70% (Ljubičić *et al*., 2023; Maricic *et al*., 2016; McKenzie and Williams, 2015).

* University of Novi Sad, Faculty of Agriculture, Serbia.
** University of Belgrade, Faculty of Organizational Sciences, Serbia.
*** University of Novi Sad, Faculty of Technical Sciences, Serbia.
Corresponding author: marina.ignjatovic@fon.bg.ac.rs

Despite its importance, sustaining a business in this sector is difficult, especially for small local farmers. One of the main reasons is insufficient assets for funding a business, and among the various factors contributing to low profitability of agriculture, the foremost is lack of access to credit (Bharti, 2018). Farmers often require considerable loans to transform their agricultural business into a profit-oriented farm business (Rathore *et al*., 2017). Thus, there is undoubtedly a need for improving access to agricultural financing for producers (Kusek *et al*., 2017). Unfortunately, despite the high demand for agricultural loans, financial institutions tend to avoid them, particularly for small local farmers (Kong and Loubere, 2021).

Different authors have even determined that state support is needed to provide the agricultural industry with affordable long-term lending resources and introduce referential credit mechanisms for agricultural commodity producers (Poliatykina *et al*., 2022, p. 126). In Serbia, for example, there has been a considerable improvement in the quality and volume of institutional support for developing the small and medium enterprises (SME) sector in all activities. Thus, with appropriate agrarian policy instruments, the business performance of SMEs in agribusiness is being improved year by year (Jovanović and Zubović, 2019).

The other side represents banks' risk when approving agricultural loans. Banks often encounter the problem of loans that are not repaid on time or fully repaid at all. For financial institutions, the recovery of agricultural loans is often crucial because timely repayment ensures the recycling of funds and strong confidence among the parties (Rathore *et al*., 2017). Also, it helps farmers trust in their ability to develop their business. Various factors, including government policies, demographics, institutional, cultural, and environmental elements, directly and indirectly impact loan repayment (Sileshi *et al*., 2012). Swift repayment of credit is crucial for maintaining good creditworthiness. Consequently, the inability of borrowers to repay their loans is a critical concern for the long-term viability of credit institutions (Kassegn and Endris, 2022). Our research particularly strives to determine the factors influencing the timely repayment of agricultural loans.

Most banks have a problem calculating creditworthiness for clients who do not have financial reports, i.e. audit reports. For clients with balance sheets, it is relatively easy. Banks apply standardized procedures in line with Basel II adopted principle of risk management. Still, the main challenge comes when the customer does not have an official financial report – which is often the case with small farmers in many countries. In that case, some banks on the market usually do not provide loans for this client segment, or if they do, they develop some internal methodology and logic which must be in line with Basel II and approved by the national bank of the local county. This is precisely the situation we investigate in our paper: a large group of small producers, approximately 449,000 of them, without financial reports but with a real need for investments in their agricultural production. Government subsidies and guaranty schemes can cover one part of their needs, but producers must also use commercial bank loans.

In the Serbian market, it is not possible to insure the debt of small agricultural producers to the bank. Still, despite that fact, commercial banks have provided working capital and investment loans to producers very actively for the last 20 years. In general, the National Bank of Serbia supports the development of new credit products for agricultural producers because it is apparent that they are in continuous need of them.

Credit in Serbia policy reveals that the budget for subsidies in agriculture is growing every year. The government is providing cheap or free loans for investments and working capital (WC) and 30% to 50% cash back (grants) for defined types of investments in agriculture. The most subsidy types of investments are new tractors, machinery, irrigation systems, and hail nets. In 2024, the Ministry of Agriculture's most significant budget for cattle and milk production (around 70%) is proposed. In general, Agricultural unions - cooperatives also provide, in some cases, benefits for farmers by postponing payments up to 12 months for WC or smaller machinery. The Ministry of Agriculture and the European Union (EU) are the main supporters of direct finance to small farmers.

We performed a study in the Serbian region, in one of the central banks that provides agricultural loans to small farmers. Although agriculture plays a crucial role in Serbia, the sector faces challenges such as outdated equipment, fragmentation of holdings, lack of modern technologies, and problems in the distribution and marketing of products. Serbia is located on a total area of 8,840 thousand hectares. The area of agricultural land includes 5,734 thousand hectares, and around 4,867 thousand hectares of that area is arable land. There are over 450 thousand smallholder farmers in Serbia. When it comes to context analysis of agricultural needs in Serbia, according to data from the Credit Bureau (CB), only 10% of smallholder farmers use loans, and around 30% of legal entities in agro-businesses use loans. Generally, 80% of approved loans are used for WC, such as seeds, fertilizers, chemistry, etc., and just 20% for investment purposes. Around 30% of small farmers postpone paying (up to 6 months) provided by the distributor of WC. In the coming years, the use of loans will increase because distributors will slowly postpone paying services. Serbia's challenge is finding a balance that increases loan accessibility without compromising responsible lending practices or violating regulatory requirements. Striking this balance can help foster financial inclusion while mitigating the associated risks. One of the pre-requirements of increasing access to loans is the profound insight into their timely repayment.

This study aims to determine the significant drivers of timely repayment of agricultural bank loans. The results could help distinguish between lenders who would be acceptable and those who would possibly have problems with the repayment of loans (Hardy and Weed, 1980). To create the appropriate model, we derived an extensive series of latent input variables, which are specific and unique to the agricultural sector. We imported the latent inputs into the logit model to extract the most crucial input variables, predict timely repayment, and interpret the results justly. We later developed machine learning algorithms to enhance the performance of the model prediction.

The next chapter offers a literature review on agricultural loans and timely repayment. The following chapters briefly describe the methodology, the study results, and the discussion. Finally, the conclusions of the study are given.

2. Literature Review

Advancements in agriculture significantly contribute to the availability, accessibility, and stability of food resources. The Economist Intelligence Unit has developed The Global Food Security Index (GFSI), measuring the level to which countries provide safe food to their citizens (Izraelov and Silber, 2019; Maricic *et al*., 2016). Chavas *et al*. (2022) have analyzed the yield risk and its implications for the economics of food security. They have investigated agricultural diversification throughout different regions in Italy and found a close relationship between agricultural development and reductions in food insecurity (Chavas *et al*., 2022). According to the authors, agricultural technology has been a key driver in reducing food insecurity through increasing food production and reducing risk exposure in agriculture.

One of the most significant problems of agricultural development is the lack of access to finance (Endris and Kassegn, 2023; Mirč *et al*., 2023; Ozalp, 2019), which is majorly contributing to low profitability in agriculture (Bharti, 2018). Many of the bank's rules and regulations disadvantage farmers' ability to access finance because of a lack of collateral, regulatory matters, lending criteria, and the short timeframe for land leases, thus predominantly affecting production (Amadhila and Ikhide, 2016). As Huang and Wang (2014) show, agriculture is much underinvested. In developing countries, domestic and foreign aid has not increased appropriately to maintain sustainable agriculture (Huang and Wang, 2014). Moreover, transitioning to sustainable agricultural systems is imperative to meet the global Sustainable Development Goals (SDGs). However, achieving more sustainable agricultural production systems will require significant additional capital (Havemann *et al*., 2020; Xia *et al*., 2022).

Agricultural loans affect the value of agricultural production (Chandio *et al*., 2018; Kadanali and Kaya, 2020). Besides savings and insurance prod-

ucts, credits largely influence the capacity of small local farmers and are crucial for inclusive finance and sustainable agricultural production (Peprah *et al*., 2021). In developing countries, smallholder farmers face two significant barriers to agricultural investment: weather-related risk and credit. In that sense, rural areas suffer from several financial market imperfections that hinder credit market access and agricultural investment, particularly among smallholders (Mishra *et al*., 2021).

Agricultural loans have been studied worldwide. Kaya and Kadanalı (2022) found indications that the effect of agricultural loans in Turkey has increased over the past time, and the deposit banks have a high impact on agricultural production. Moreover, the ratio of agricultural credits to total loans in Turkey has increased from 3% to 6% between 1999 and 2014, and during this process, new actors, such as private banks, agricultural lenders, agricultural product marketing firms, etc., have entered the agricultural credit market (Kusek *et al*., 2017). In China, the government has taken measures to increase access to agricultural loans to improve farmers' social welfare (Feder *et al*., 1989; Gong and Elahi, 2022). For example, rural land management mortgage loans can enable farmers to gain more credit funds, which is conducive to agricultural development and revitalization (Zheng and Zhang, 2021). China has successfully achieved its intended policy goal of boosting the agriculture sector (Lin and He, 2020). Also, through fiscal incentives, financial institutions are encouraged to increase agricultural loan offers, leading to a significant decline in urban-rural income inequality, particularly in underdeveloped areas of China (Tang and Sun, 2022).

On the other hand, credit availability is much lower in less developed countries. Agricultural credit is a significant factor in the Indian agricultural sector (Behera and Behera, 2022). Still, credit utilization for productive purposes is limited due to the frequent abuse for nonagricultural purposes (Rathore *et al*., 2017). In Nigeria, the agricultural sector is characterized by low productivity due to the lack of modern technologies, and poor access to credit is seen as a critical barrier to their adoption (Balana and Oyeyemi, 2022). One of the main challenges the agricultural banks in Iran encounter is the high probability of repayment failure (Pishbahar *et al*., 2015). In Ghana, the mistrust of financial institutions in small local farmers is seen through activities such as requesting enormous collateral, guarantors, high savings, high interest rates for agriculture loans, delinquency, and bureaucratic processes in accessing loans (Teye and Quarshie, 2022).

Some authors propose introducing insured loan products to increase credit market access through an increase in the supply of credit (Mishra *et al*., 2021). Teye and Quarshie (2022) suggested that enabling policy environment and frameworks with supportive rural infrastructure, such as warehouse receipt systems, can make major increases in farmers' access to loans for investment in modern technologies, which can further increase agricultural productivity, essential to address issues of food uncertainty and rural poverty in Ghana. Some authors propose blended finance as a novel alternative to financing the agricultural sector (Dey and Mishra, 2022).

In the past, for example, private banks typically had little interest in agricultural finance. Thus, states became lenders of last resort for local farmers, setting up agricultural finance institutions (Martin and Clapp, 2015). The state hence played a crucial role in providing agricultural credit and capital to farmers, and, importantly, the state protected banks and finance by preventing foreclosures and other losses on loans (Martin and Clapp, 2015; Onyiriuba *et al*., 2020). In Azerbaijan, for example, agriculture financing has positive potential since the government provides financial support through investments and loans that may positively affect farmers' financial sustainability and competitiveness (Humbatova and Hajiyev, 2021). However, given that government subsidies and banking loans do not always provide sufficient funds, according to some authors, new sources of financing are in need, and Mirovic and Bolesnikov are pointing out the possibilities of applying asset securitization in financing agriculture (Mirović and Bolesnikov, 2013).

Agricultural credit is characterized not only by the approval of the loan but also by its timely repayment. The effective performance of financial institutions can only be judged when the farm-

er-borrowers repay their loans and when they fall due to the farm credit agency (Rathore *et al*., 2017).

Timely loan repayment is an essential issue in all sectors. Increasing the number of payment loan models is vital for financial institutions to attract more clients (Eroglu and Ozturk, 2016). To ensure approved loan repayment, they adapt innovations and new technologies to payment intermediation (Miglionico, 2023). It was common for bank clients to desire not to make payments in some periods because of income uncertainty over time. Formato (1992) and Moon (1994) used this to study the models of arbitrary skips.

Jet, when it comes to agricultural loans, the determinants of timely repayment are pretty specific. Olagunju has, for example, found that a ration rate, among others, was an important factor in ascertaining the rate of repayment at different significant levels for crop farmers, leading to a conclusion that credit rationing did have a positive influence on agricultural credit repayment (Olagunju *et al*., 2023). When observing the size of local farmers, findings show that, compared to medium farmers, marginal and small farmers diverted a portion of the loans. The extent of loan repayment by medium farmers was higher than that of small and marginal farmers (Ray and Das, 2023). A study by Amedi, Dumayiri, and Mohammed (2019) showed that the factors that significantly influence loan repayment are sex, household size, group size output value, and loan disbursement timeliness. Moreover, over-indebtedness was found to be higher among the more experienced farmers and the farmers having more percentage of cash crop and with the increase of overdue amount and credit demand per acre, but lower with the rise of per acre cost for production (Das and Sharma, 2023).

Pishbahar and his coworkers found that extra activities besides farming, the extension of the loan repayment period, and a large volume of received loans were the causes that had significant negative impacts on loan repayment. On the other hand, causes like high interest rates of loans, existing collaterals or different types of guarantors, services received from the banks, and long-term maturity periods significantly boost the probability of timely loan repayment (Pishbahar *et al*., 2015). Other methods, such as crop insurance, protect lenders by increasing the likelihood of loan repayment when revenue declines (Ifft *et al*., 2023). Agricultural insurance can positively affect agricultural loans by reducing the risk for lenders, thereby encouraging more favourable loan terms and increasing credit availability to farmers. However, some authors have found that crop subsidies negatively impact farmers' insurance policies and premiums, with public aid disincentivizing agricultural insurance, leading to their low penetration, which is the case in Italy (Miglietta *et al*., 2020).

According to the Common Agricultural Policy (CAP), the EU promotes food security and sustainable farming by providing income support to farmers. CAP Strategic Plans support the resilience of the agricultural sector by supporting viable farm income through direct payments to active farmers (European Commission, 2024). CAP financial instruments have significant potential to contribute to the "Farm to Fork" and "Biodiversity strategies", but the financial needs of agriculture and the agri-food sector remain high. In Poland, for example, the permanent domination of the subsidies under CAP and its first pillar (market management and income support) radically narrows the space for using financial instruments (Kulawik *et al*., 2018). To receive total payments under the first pillar, farmers must comply with cross-compliance, which covers statutory management requirements and standards for good agricultural and environmental conditions and the requirements of 'greening' (Heyl *et al*., 2021). Kulawik *et al*. (2018) consider that financial instruments under the CAP are adjusted mainly to achieve allocation and stabilisation objectives and that only larger farms may be interested in them. Staniszewski and Borychowski (2020) have also found that the impact of subsidies on efficiency depends on the size of farms and that the significant, stimulating effect of subsidies was identified only in the group of the largest farms. Thus, even in the EU, it is more difficult for smallholder farmers to manage. Serbia has over 85% of small-sized agricultural companies, while only 1.5% are large farmers.

As Western Balkan (WB) countries are in the process of integration into the EU, agricultural

efficiency and the actuality of the problems of the agricultural sector are paramount (Đokić *et al*., 2022). Still, as the same authors show, there is a significant difference in technical efficiency between WB and the EU, which is significantly lower in WB. This might be an alarm for policymakers in the WB, in the sense that agricultural policy measures should encourage more intensive agricultural production, which could create a better foundation for agricultural growth (Matkovski *et al*., 2022).

Agriculture is essential in Serbia's economy, with a high share of the Gross Domestic Product (GDP), which, according to the World Bank, was 6.29% in 2021 (Radulović *et al*., 2023). In a study on micro-sized, small-sized, and medium-sized agricultural companies in Serbia, results indicate that the micro-sized agricultural companies were at greater risk of bankruptcy. In contrast, small- and medium-sized companies were more stable (Milić *et al*., 2022). Regarding the reliability of reported earnings of agricultural companies in Serbia, profitability and leverage significantly influence the scope and direction of earnings management. Income-decreasing behaviour is observed in more profitable firms, while income-increasing behaviour is observed in highly leveraged and more profitable firms (Milić *et al*., 2018). Further, Tekic *et al*. (2021) studied the model of agricultural loan approval in Serbia. They have found, from several models, that consistent factors influencing loan approval were profitability and solvency of the smallholders (Tekić *et al*., 2021).

As mentioned before, the agricultural land area in Serbia is more than 5.7 million hectares, and more than 4.8 million hectares of that area is arable land. The most developed branches of agriculture are animal husbandry (43%) and arable farming (42%), followed by fruit growing and viticulture (12%), while other crops are represented by 3%. Family farming and private ownership (smallholders) dominate, while the average size of a commercial farm (enterprise) is 500 to 700 hectares. Family farming is fragmented and has a pronounced natural consumption and a significantly lower degree of commercialization than European farms. Approximately 600 thousand tractors, 38 thousand harvesters, and more than 3 million attachment machines cultivate agricultural land (SORS, 2023a). Regarding livestock, according to data from 2018, there were a total of 424,155 dairy cows, 3,266,102 pigs, 1,799,814 sheep, and 218,397 goats (SORS, 2023b). In Serbia, the irrigation system is installed on 180 thousand hectares. Still, only 30-40 thousand hectares are irrigated, which, together with gardens and some newer systems, represents less than 1% of the arable land.

Table 1 - Number of loans of smallholder farmers in Serbia.

Date	*Number of loans*	*Number of clients*
December 31, 2022	70,545	45,522
November 30, 2023	68,348	43,747

Data source: CB (2023)

From more than 450 thousand registered agricultural holdings in Serbia, out of which 433,217 are active, barely 10% use loans. Table 1 shows the number of loans for smallholder farmers in Serbia.

As a candidate country for membership in the EU, Serbia is working to harmonize its agricultural policy with European standards to improve the competitiveness and sustainability of its agricultural sector. The agricultural market and its upstream food and beverage sector have always been considered the most valuable resource and have tremendous potential that can and should be used. Serbia's most significant traditional comparative advantage lies in favourable climate conditions and rich and fertile land.

3. Methodology

3.1. *Respondents*

Our research observed small local farmers in Serbia. Data were provided by the authors from the OTP bank, one of the largest banks in Serbia, with a highly developed agricultural department. Data contain 1,735 liquidated loans from Serbian smallholders. We observed the *timely repayment* of these loans, representing our primary dependent variable. Of 1,735 loans, 926 were timely repaid, while 809 were defaults. Liquidated loans are dated from 2018 to 2023.

3.2. *Variables and Models*

In this paper, we have modelled the timely repayment of agricultural loans, which represents the binary variable, as noted above. Initially, we extracted a vast set of 36 input variables to evaluate the output. Input variables were chosen for dual reasons. First, it was in accordance with the presented literature review. Second, it was derived from the discussion sessions with leading experts in the bank's agricultural sector.

Input variables include bank data on the client (previously credited client, the average monthly number of salary transactions, profiling, segment, years between the first product approval and a report), data on special account (SA) for agricultural purposes (period of owning SA, SA payment card, period of owning E-bank SA, SA average monthly inflow, SA average monthly number of transactions), client descriptives (age, mail address, employment data availability, occupation, employment status, education), loan data (loan type, loan amount in RSD, related parties), CB data (client's exposure on application date, clean credit history, CB report pulls in 30 days, past due records over five days, no credit in the past, tax debt status, number of credit products), farm data (farm existence in years, farm area, owned farm area, share of owned area, number of crops, number of farm members 18 to 72, farm holder age on application date, farming, horticulture, fruit growing, livestock, other). The input variables encompass a comprehensive range of data, including the clients' banking history, SA details, personal demographics, loan specifics, CB information, and detailed farm data. This diverse set of factors is crucial in assessing the risk and potential for timely loan repayment, as they provide a holistic view of the client's financial stability, creditworthiness, and the operational aspects of their agricultural activities.

We have used a two-fold approach to predict the timely repayment of agricultural loans. Firstly, we used the list of input variables as predictors in the Logistic regression model. The purpose was to obtain an explainable model that would be useful to stakeholders and policymakers. The selected variables that would show the highest importance for predicting timely repayment will be described in detail in the results section.

Secondly, looking up to some authors (Chen *et al*., 2021; Elnaggar *et al*., 2020), we have applied machine learning algorithms for prediction. The primary purpose was to downsize the list of our input variables. Based on the evidence on best prediction performance from the literature, the models we have focused on are XGBoost, Random Forests, and Support Vector Machines (Chen *et al*., 2021; Elnaggar *et al*., 2020). The results of our analyses are given in the following chapter.

4. Results and Discussion

In the first part of our research, we have created the Logit model for the timely repayment of agricultural loans. As described in the previous section, we have included 36 predictor variables in our model. The following formulas define the model:

$$P_i = \frac{e^{\beta_0 + \sum_{j=1}^{k} \beta_j X_{ji}}}{1 + e^{\beta_0 + \sum_{j=1}^{k} \beta_j X_{ji}}} \quad (1)$$

or

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \sum_{j=1}^{k} \beta_j X_{ji} + \varepsilon_i \quad (2)$$

where P_i presents the probability of the i-th client to repay the loan on time, i=1…n, L_i presents the Logit model, X_{ji} represents the j-th predictor variable for the i-th client, β_0 represents the constant and βj represents the coefficient for the j-th predictor variable.

We have performed a backwards Logit modelling to reduce the model to the significant input variables for the prediction. The initial model represents the model with all 36 predictors, while the final model includes 18 predictors. The results of the model are given in Table 2.

The Naglekerke R of the initial model is 0.164, and for the final model is 0.156, which is a bit lower than preferable. However, the omnibus test is statistically significant ($p<0.001$). At the same time, the Hosmer and Lemeshow Test is not statistically significant ($p=0.472$ for the initial model and $p=0.242$ for the final model), which shows that our model is not a bad fit. The overall accuracy of the final model is 64.4%.

The variables shown to be significant for predicting timely repayment in the final Logit model

Table 2 - Results of the Logit model.

	Initial model			*Final model*		
Variable	*B*	*Wald*	*Exp(B)*	*B*	*Wald*	*Exp(B)*
Constant	1.254	3.820	3.505	1.492	6.812*	4.446
Period of Owning SA (years)	0.003	0.067	1.003			
Previously Credited Client	0.259	2.510	1.296	0.290	4.874*	1.336
Payment Card for SA	0.180	1.787	1.197			
Owning E-Bank SA (years)	0.068	4.075*	1.070	0.073	4.969*	1.075
Average Monthly Inflow to SA (100 EUR)	0.014	5.911*	1.015	0.014	6.297*	1.015
Average Monthly Number of SA Transactions	-0.024	3.967*	0.976	-0.020	3.264	0.981
Profiling in Months	-0.455	5.963*	0.635	-0.470	6.507*	0.625
Age	-0.199	7.256*	0.820	-0.190	7.665**	0.827
Mail Availability	-0.285	4.355*	0.752	-0.265	3.985*	0.767
Employment Data Availability	-0.107	0.425	0.898			
Segment	0.042	0.130	1.043			
Occupation	0.174	2.651	1.190			
Related Parties	-0.163	1.015	0.850			
Loan Type	-0.300	4.042*	0.741	-0.293	4.005*	0.746
Loan Amount (1000 EUR)	0.028	3.725	1.028	0.024	3.126	1.025
Employment Status	-0.179	6.899*	0.836	-0.199	9.841**	0.819
Education	0.021	0.073	1.021			
Farm Existence in Years	-0.001	0.006	0.999			
Farm Area	0.031	0.033	1.032			
Owned Farm Area	-0.008	0.457	0.992	-0.014	2.901	0.986
Share of Owned Area	-0.024	0.015	0.976			
Number of Crops	-0.019	0.774	0.981			
Number of Farm Members 18 to 72	-0.015	0.075	0.985			
Farm Holder's Age When Applying	0.208	7.886**	1.232	0.199	8.377**	1.220
Clean Credit History	0.635	12.878***	1.887	0.622	12.583***	1.862
CB Report Pulls in the Last 30 Days	-0.132	0.155	0.877			
Past Due Records over 5 Days	-0.735	7.532**	0.479	-0.779	8.763**	0.459
No Credit in Past	0.277	4.185*	1.319	0.267	4.349*	1.306
Years Between First Product Approval and Report	5.455	5.960*	234.011	5.642	6.510*	282.145
Tax Debt Status (1000 EUR)	-0.045	4.541*	0.956	-0.047	5.454*	0.954
Number of Credit Products	-0.020	0.869	0.980			
Farming	-0.035	0.040	0.966			
Horticulture	-0.083	0.260	0.921			
Fruit Growing	-0.010	0.003	0.990			
Other	-0.078	0.220	0.925			
Livestock	-0.001	5.151	0.999	-0.002	6.443*	0.998
Goodness of fit measures						
Naglerke R	0.164			0.156		
Omnibus test	227.43***			215.55***		
Hosmer and Lemeshow Test	7.617			10.340		
		Predicted			*Predicted*	
Confusion matrix		Yes	No		Yes	No
	Yes	68.25%	31.75%	Yes	66.63%	33.37%
	No	39.06%	60.94%	No	38.07%	61.93%

**p<0.05; **p<0.01; ***p<0.001.*

include the set of 18 input variables, which will be further discussed. *Previously Credited Client* is a client who has already used the bank's services in terms of credit, and the bank has a lot of data about this client and his repayment history. When the client applies for a loan again, the bank can generate additional interest income and increase client-level profitability. If the previous loan was repaid per the agreed terms, it could increase the bank's faith in the client's ability to meet obligations, reducing the risk of defaulting in the repayment period. This information provides benefits for the bank in terms of reduction of marketing costs – it is more likely that a previously credited client will use the services of the same bank again, which can reduce the need for intensive marketing campaigns to attract new clients; better understanding of credit risk – the bank already has historical data on the client's behavior regarding loan repayment, which enables a better assessment of credit risk and making informed decisions regarding the approval of new loans; personalized offers – bank may use historical data on previous loans to provide customized offers to the client, including better interest rates or other benefits that reflect the client's financial profile and needs; maintaining existing relationships – a long-term relationship between the client and the bank can positively affect the overall loyalty of the client; more efficient approval processes - faster and more efficient for previously loaned clients, which can increase client satisfaction and reduce costs; expanding the portfolio – by re-crediting existing clients, the bank can gradually expand its loan portfolio and diversify risk. This variable's odds ratio (OR) is 1.336, meaning that if the client is previously credited, the higher the odds are that he will repay the loan in time.

Average Monthly Inflow to SA, measured in hundreds of EUR, represents the amount regularly paid into the client's SA every month. This information has several benefits for the bank: ensuring the stability of deposits – regular inflow of funds to the SA contributes to the stability of deposits, which is critical to maintaining bank liquidity and stability; potential for additional services – clients with a regular flow of money to their SA may be more inclined to use other products and services of the bank, such as savings, loans, investments or cards, which can provide additional profits to the bank; reducing the risk of overdraft – clients with stable incomes are less likely to face unauthorized overdrafts, which reduces the need for the bank to intervene and charge additional fees; better assessment of credit risk – clients with a stable income usually represent a lower risk for the bank compared to clients whose income varies or is irregular; opportunity for personalization of services; more efficient liquidity management. This variable's OR is 1.015, meaning that the higher the monthly inflow to SA, the higher the odds the client will repay the loan in time. Similarly, annual farm income was a significant factor for small-scale farmers (Isibor and Nkamigbo, 2019).

Average Monthly Number of SA Transactions refers to the total number of financial transactions that the client performs on his SA during the month. Some of the benefits that the bank can have from a high average monthly number of transactions on the SA include: income from fees – the bank may charge fees for certain types of transactions, increasing the bank's income; increased client interaction – more frequent interaction enables the bank to provide additional services, transaction notifications, personalized offers or advice; encouraging the use of digital channels – reducing operational costs compared to traditional methods, contributing to more efficient operations; better understanding of client needs – the bank can adjust its services, offers and marketing strategies; reduced risk of overdraft – clients who regularly transact on their SA often have better control over their finances, reducing the risk of unauthorized transactions or overdrafts; ability to offer personalized products; better risk management - helps the bank identify potential risks, such as suspicious transactions or fraud. However, as opposed to the expectations, this variable's OR is 0.981, meaning that the higher the monthly number of SA transactions, the lower the odds that the client will repay the loan in time.

Profiling in Months refers to the time that has passed since opening the SA. Benefits for the bank include long-term client loyalty, monitoring financial behaviour, increasing proactivity in providing services, target groups, personalizing services and communicating. However, as opposed to the expectations, this variable's OR is 0.625, meaning that the longer the profiling, the lower the odds that the client will repay the loan in time.

Age of the client at the time of applying the bank's product can significantly affect various aspects of the banking relationship: credit risk - younger clients often have longer service life and generate income for a longer time; types of products and services; long-term client relationship; marketing strategies - better targeting different age groups; risk insurance; customized offers. This variable's OR is 0.827, meaning that the older the client, the lower the odds that he will repay the loan in time. The results are in favour of younger clients. Some authors have also found that age negatively affects the repayment rate of smallholder farmers (Kassegn and Endris, 2022).

Mail Availability is important to the bank for personalized communication since they may use e-mail to inform clients about new services, updates, interest rates, policy changes, and other significant information. Clients can receive electronic reports instead of paper statements via e-mail, which is more environmentally friendly and convenient. The bank can send reminders about upcoming payment deadlines, which can help clients avoid delays and unwanted fees. Clients can receive e-mail notifications about changes in laws and regulations that affect banking services or their accounts. However, as opposed to expectations, this variable's OR is 0.767, meaning that if the mail is available, the odds that the client will repay the loan in time are lower.

Loan Type monitoring compares new and previous applications and possible predictions of new loans by maturity and amount. Clients may use three types of loans: working capital, overdraft, or investment. Working capital loans provide short-term funding for operational needs, overdrafts offer flexible borrowing up to a certain limit for immediate expenses, and investment loans are long-term financing for major purchases or projects. Clients benefit by receiving funds to start or expand business or maintain liquidity, while banks benefit by earning income from interest and fees. In addition, loans often have certain conditions that help the bank monitor and manage risks.

Loan Amount, measured in thousands of EUR, provides valuable information about the client. A bank can consider offering long-term loans if the client has demonstrated the ability to maintain financial discipline in the long term. This variable's OR is 1.025, meaning that the higher the loan amount, the higher the odds the client will repay the loan in time. The loan amount was also found significant by Isibor and Nkamigbo (2019).

Employment Status is read from the Excerpt from the register of agricultural holdings and includes (1) exclusively engaged in agriculture, (2) engaged in countryside tourism, (3) agriculture is the predominant activity (more than 50% of working time), (4) pensioner (in case of non-existence of formal status of pensioner, persons over 65 years of age), and (5) agriculture is an additional activity (less than 50% of working time). The data can help make the final loan decision correctly and with the rating. Clients who are exclusively engaged in agriculture, engaged in countryside tourism, or have agriculture as the predominant activity (more than 50% of working time) are considered more desirable when assessing the rating. On the other hand, pensioners or clients for whom agriculture is an additional activity have additional income other than agriculture. Their taxes are settled from other incomes (no tax debt), so they are more acceptable from the point of view of risk.

Owned Farm Area is also read from the Excerpt from the register of agricultural holdings. This variable's OR is 0.986, meaning that the larger the farm area, the lower the odds that the client will repay the loan in time. The results favour smaller farms, as opposed to some findings that land size positively influences the repayment rate of smallholder farmers (Kassegn and Endris, 2022).

Farm Holder's Age When Applying is a significant factor when assessing the creditworthiness and general risk for the bank. Younger smallholders without financial history present a challenge in evaluating creditworthiness. Older smallholders often have a more stable financial situation, long-term work experience, and more experience in business but may be closer to retirement. This variable's OR is 1.220, meaning that the older the client is when applying for a loan, the higher the odds he will repay the loan in time. As opposed to the total age, this result favours older clients at the time of application. This result is in accordance with the previous literature studies. We showed that it is more probable for older farmers at the time of application to repay the loan in time. Like-

ly, Das and Sharma found that the over-indebtedness is related to the farmers from a lower age group (Das and Sharma, 2023).

Clean Credit History is read from the CB report and is generally one of the parameters of the loan rating. The bank usually analyzes this information to make an informed decision and manage risks per policy and objectives. Suppose the client has regularly met his financial obligations in the past and has no late payments. In that case, this usually positively affects his credit rating and relaxes the terms of loan approval by the risk. A clean credit history without historical delay is essential when evaluating a client's credit rating. Positive impacts include: bank's trust – the bank usually evaluates repayment history to assess the level of risk associated with granting a loan; increase in credit rating – regular settlement of obligations usually results in an improvement in credit rating, which can lead to more favourable loan terms; more favourable loan terms; faster approval process; lower risk of non-payment – clients with a good credit score and a clean repayment history represent a lower risk for the bank in terms of potential unpaid debt; beneficial impact on the bank's portfolio. As expected, this variable's OR is 1.862, meaning that if the client has a clean credit history, the odds are higher that he will repay the loan in time.

Past Due Records over 5 Days are read from the CB report. If the smallholder has a recorded past delay in the CB for more than five days, this can significantly impact the bank's risk assessment and approval of the loan. It can serve as an indicator of financial problems and increased risk. These clients are categorically considered different from clients with a clean credit history. They are often asked for additional collateral or evidence to reduce the potential risk of a possible delay in repayment of the new loan. The bank can more carefully analyze their financial ability to meet obligations before approving a new loan. Interest rates could even be increased to compensate for increased approval risk and stricter lending conditions. As expected, this variable's OR is 0.459, meaning that if the client has past due records over five days, the lower the odds that he will repay the loan in time.

No Credit in Past is read from the CB report. The absence of earlier data on credit behaviour can affect the client's rating through a lack of data for analysis, increased uncertainty, or difficulty in determining creditworthiness. Possible impacts on ratings include a lower initial rating category, gradual rating increase, and additional requirements or collection of alternative evidence. However, this variable's OR is 1.306, meaning that if the client had no credit in the past, the odds that he will repay the loan in time are higher.

Years Between First Product Approval and Report is read from the system, which counts the years from the approval of the first placement until today. It provides a complete picture of the client: he used/did not use the bank's products, was late/regular in repayment, and sufficient information for the risk parameter in which this is contained. Potential benefits include: long-term loyalty – if a long period has passed from the approval of the first product to the date of the report, this may indicate a long-term relationship with the bank and client loyalty; insight into loan repayment history; product resale – a long period may indicate the need to re-engage the client and resell the product; personalized offers; more efficient loan approvals. This variable's OR is 282, meaning that the longer the time between the first product approval and report, the higher the odds that the client will repay the loan in time.

Tax Debt Status, measured in thousands of EUR, is read from a tax certificate the client must provide. The client is registered to pay taxes from agriculture if the basis of the work activity is only agriculture or rural tourism. Tax debt in some parts of Serbia is present; some farmers owe 10-20,000 EUR with interest. The state has not resolved this yet. High debts of 10-20,000 EUR with interest can significantly burden the financial situation of farmers. It impacts liquidity, creating a need for tax solutions and government actions. This variable's OR is 0.954, meaning that the higher the debt status, the lower the odds that the client will repay the loan in time.

Livestock is read from Excerpt from the register of agricultural holdings (cattle, sheep, pigs, etc.). Clients who have more livestock are better rated. Livestock includes dairy cows and bulls, sows and fattening pigs, sheep, goats, etc. Calculation of creditworthiness for a working

Table 3 - Comparison of machine learning algorithms performance.

Model	*F1*	*Accuracy*	*ROC AUC*	*Timely repayment*	*Confusion matrix*	
					Yes	*No*
XGBoost	92.78%	92.28%	92.23%	Yes	92.98%	7.02%
				No	8.53%	91.47%
Random forests	65.37%	61.14%	60.79%	Yes	68.6%	31.4%
				No	47.02%	52.98%

capital loan implies indebtedness concerning production costs. In the case of an investment loan, the instalment and annual repayment are calculated based on the net profit. However, this variable's OR is 0.998, meaning that the more livestock, the odds that the client will repay the loan in time are lower. This result is opposed to the findings of Kassegn and Endris (2022), who showed that livestock has a significant positive influence on the repayment rate of smallholder farmers.

The second part of our research includes applying and comparing specific machine learning algorithms to find the model that best predicts the data. As discussed before, our output was a binary variable, namely, a *timely repayment* that denotes whether the bank client - a smallholder – repaid his loan on time. Based on the previous literature, we have chosen three models that showed the best results: Random Forests, SVM, and XGBoost. The SVM model gave very low accuracy prediction results when modelling timely credit repayment; thus, we have focused on XGBoost and Random Forests. We have divided the dataset into train and test sets encompassing 70% and 30% of the dataset. We have compared the performance of these machine learning algorithms. The results are given in Table 3.

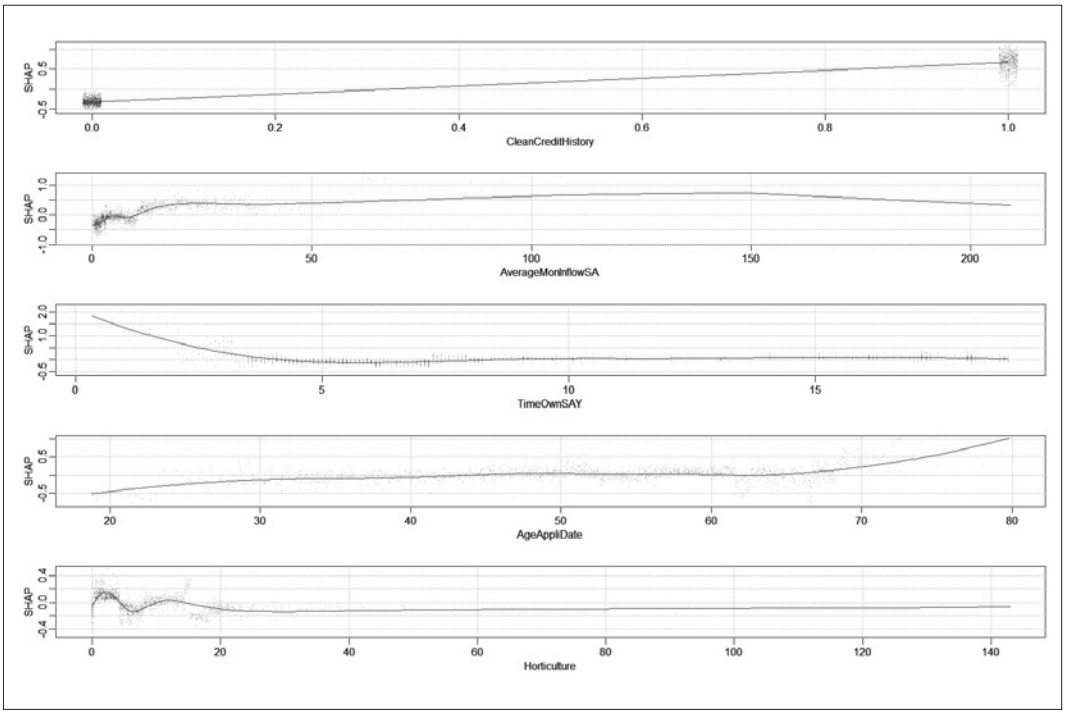
As can be seen from Table 3, the Random forests model provided even lower accuracy than the Logit model. Thus, we have further analyzed only the XGBoost model. XGBoost has predicted timely loan repayment with more than 92% accuracy. Area Under Curve (ROC AUC), an aggregate performance measure across all possible classification thresholds, is 92.23%, indicating that the model can discriminate between the two classes. XGBoost model is likely to be reliable in its predictions. An F1 score for the model is 92.78%, indicating a solid balance between precision and recall and suggesting a robust model performance.

In Figure 1, we have shown Shapley's additive explanations (SHAPS) of the XGBoost model (Nohara *et al*., 2022; Shi *et al*., 2022; Wieland *et al*., 2021) to find which of the input variables are most important for predicting timely credit repayment.

From Figure 1, the most critical factors determining timely repayment are Clean Credit History, Average Monthly Inflow to SA, Period of Owning SA, Farm Holder's Age When Applying, and Horticulture. As in the Logit model, if the client has a clean credit history, the chances are higher for the timely repayment of a loan, as evident from Figure 1. A higher average monthly inflow to SA increases the chances of repaying the loan in time. Also, as in the Logit model, if the client is older when applying for the loan, the chances to repay the loan in time are higher.

Interestingly, two variables were found to be important for prediction in XGBoost but were not significant in the final Logit model: Period of Owning SA and Horticulture. Besides timely repayment, the *Period of Owning SA* benefits the bank through long-term client loyalty and monitoring financial behaviour, enabling personalized offers and adaptive services. *Horticulture* is read from the Excerpt from the register of agricultural holdings and then distributed within the system by categories and types of products to calculate the client's creditworthiness. Calculation of creditworthiness for a working capital loan implies indebtedness concerning production costs. For example, in an investment loan, the annual repayment is related to the net profit from horticulture.

Figure 1 - XGBoost SHAPS.



Source: Produced by the author.

5. Conclusion and Policy Implications

Financial institutions in developing countries continue to grapple with the critical issue of defaults on borrowed funds, which poses a significant challenge in their efforts to broaden their services and reach a wider client base (Kassegn and Endris, 2022). This issue is closely tied to the need to improve credit literacy, enhance land productivity, diversify income sources beyond agriculture, and simplify standard credit procedures for marginal farmers and landless sharecroppers, as Das and Sharma (2023) suggested. The lack of access to financial services, particularly loans, is a critical factor in the poverty and underdevelopment of smallholder farmers. While in Serbia, only 10% of smallholder farmers use loans, in Turkey, for example, agricultural land accounts for a significant portion of the total capital, and surveys show that 78% of agricultural enterprises use agricultural loans and spend 88% of them for input purchases (Kusek *et al*., 2017). The low profit rate and the emerging individual needs increase these loan demands. The credit policy of Serbian banks is providing cheap or free loans, or 30-50% cash back for investments and WC in agriculture. The credit policy of agricultural credit cooperatives and agricultural banks in Turkey is to apply subsidies of 50% (like for livestock) to 100% (like for irrigation investments) of loan interest rates (Kusek *et al*., 2017). The same authors state that many modern agricultural enterprises have been established thanks to the projects supported by the EU and the Ministry of Agriculture, which are also the main credit supporters in Serbia.

Nevertheless, banks are deeply concerned about the high rate of loan defaults. Given these interconnected challenges, our study is dedicated to identifying the factors that contribute to the timely repayment of loans, aiming to address both the needs of the financial institutions and the borrowers effectively.

The first part of our study presents the econometric results obtained to determine the significant influence factors. Striving towards establishing and clarifying relations between independent

variables and timely loan repayment of smallholder farmers in Serbia, we developed a Logit model that provided about 64% accuracy. The model discovered 18 significant predictors out of 36 initial feasible determinants. The significant and positive influence of the variable previously credited clients indicates that pre-lending to clients can bring multiple benefits to the bank. Besides timely loan repayment, these are higher revenues, more efficient processes, and maintaining long-term client relationships. In the case of a new client, the approval process is based on available external data and risk cost based on the existing portfolio of similar client groups. The regular average monthly inflow to SA brings banks stability through timely loan repayment, as well as the opportunity to increase income and provide additional services to clients, contributing to maintaining the successful operation of the bank. Besides affecting timely repayment, a clean credit history positively affects the client's rating and allows him favourable conditions when applying for a loan. No credits in the past against the odds positively influence timely repayment. Still, it is important to understand that the absence of a previous credit history does not automatically mean an unfavourable rating. The bank will carefully analyze all available information and use alternative indicators to make an informed decision about the client's risk and rating. Regularly settling new obligations can gradually improve the client's rating over time.

The findings of our study revealed that the historical delay in CB – over five days, negatively influenced the timely loan repayment. In essence, it can significantly affect the risk assessment and conditions for approving a new loan for an agricultural holding. Banks will make loan approval decisions based on their risk management policies and practices, considering this relevant information from credit reports. Another negative influence was found with tax debt status. Cooperation between farmers, local authorities, and the state can be crucial to solve such problems effectively. In addition, providing support to farmers through education on tax obligations and proper planning can contribute to the long-term resolution of these challenges.

As opposed to expectations, our study shows that the average monthly number of SA transactions negatively affects timely repayment, even if it brings numerous benefits to the bank, including additional revenue, better understanding of clients, increased efficiency, and opportunities for personalization of services. A similar situation is for profiling in months, which provides the context for better understanding and adjusting the approach to build a long-term and mutually beneficial relationship with the client.

The client's age significantly impacts the bank's approach and strategy, and tailoring services and benefits can contribute to better client satisfaction and the bank's long-term success. Our results favour younger clients, following the findings of other authors (Kassegn and Endris, 2022). On the counter, age when applying for a loan favoured older clients. When analyzing the creditworthiness, the bank will carefully consider the age of the smallholder, taking into account all relevant factors that affect the viability of the business and the ability to meet obligations according to the agreed terms of the loan.

Some of our findings differ from the literature. For example, in our study, the number of farm members or education was not shown to influence the timely repayment of agricultural loans. At the same time, some authors found that family size and education level were significant factors (Kassegn and Endris, 2022).

The second part of our study enhances model accuracy using machine learning algorithms. XGBoost model has shown the best performance among the tested models, providing 92% accuracy. Using the SHAPS of the XGBoost model, we identified the most critical factors determining timely repayment: clean credit history, average monthly inflow to SA, period of owning SA, age when applying for a loan, and horticulture. Clean credit history, inflow, and age at the moment of application coincided with the Logit model results. Still, the XGBoost model identified two additional significant predictors: the period of owning SA and horticulture.

This paper aims to present how existing data can be used more precisely and efficiently so that banks increase the accuracy of the assessment in the approval process. Using digital tools and

data should have a double-sided positive effect. First is on the banks, a faster and cheaper process with fewer default clients. The second is on the customers, giving them faster approval times and access to funds. In general, improvement like this should also positively affect a loan price because, according to Basel II, the projected loss must be calculated in loan price like a "risk cost". If we have a lower percentage of defaults, risk cost will be decreased, and the end price – total interest rate – will be lower.

The findings of this study have several important policy implications for financial institutions and policymakers in developing countries. Firstly, the importance of a clean credit history, regular income, and specific agricultural practices like horticulture as significant factors in loan repayment suggests that banks should refine their credit assessment processes to include these variables more prominently. This could lead to more accurate risk assessments and potentially lower default rates. Secondly, the study highlights the need for financial education among smallholder farmers, particularly in managing credit and understanding tax obligations, which could be achieved through collaborations between banks, local authorities, and agricultural organizations. The findings also suggest that banks should consider the age and agricultural experience when evaluating loan applications, as these factors were shown to influence repayment. Additionally, using advanced machine learning models like XGBoost in credit scoring indicates a shift towards more sophisticated, data-driven decision-making processes in the agricultural finance sector. This could improve the accuracy of banks' analyses and help banks tailor their services more effectively to the needs of individual farmers.

Improving and extending agriculture financing remains a significant challenge and top priority for policymakers (Bharti, 2018). For a few decades, the determination of creditworthiness for small individual farmers did not improve. The most often used tool in the worldwide market is the traditional method of credit capacity determination, established in micro-credit organizations many years ago. Future directions for our research might include assessing the factors influencing loan approval. Given that the agricultural loans are specific in terms of quantity and scope of data required from smallholder farmers, we are also attentive to improving and expanding the existing methodology for loan approval, which is currently in banks.

References

Amadhila E., Ikhide S., 2016. Constraints to financing agriculture in Namibia. *African Review of Economics and Finance*, 8(2): 82-112.

Amedi M., Dumayiri M., Mohammed A.-R.S., 2019. Loan Repayment and Its Implication on Agricultural Financing in Ghana-The Case of MiDA Agriculture Program. *International Journal of Agricultural Management and Development*, 9(4): 391-408.

Balana B.B., Oyeyemi M.A., 2022. Agricultural credit constraints in smallholder farming in developing countries: Evidence from Nigeria. *World Development Sustainability*, 1: 100012. https://doi.org/10.1016/j.wds.2022.100012.

Behera A.R., Behera M., 2022. Access and Repayment of Institutional Agricultural Credit by Farmers in Tribal Areas of Odisha: Trends, Determinants and Policy Measures. *Journal of Asian and African Studies*, 59(2): 623-639. https://doi.org/10.1177/00219096221117075.

Bharti N., 2018. Evolution of agriculture finance in India: a historical perspective. *Agricultural Finance Review*, 78(3): 376-392. https://doi.org/10.1108/AFR-05-2017-0035.

CB, 2023. *Association of Serbian Banks*. Retrieved December 20, 2023, from https://www.ubs-asb.com/en/about-us/credit-bureau.

Chandio A.A., Jiang Y., Wei F., Guangshun X., 2018. Effects of agricultural credit on wheat productivity of small farms in Sindh, Pakistan. *Agricultural Finance Review*, 78(5): 592-610. https://doi.org/10.1108/AFR-02-2017-0010.

Chavas J., Rivieccio G., Di Falco S., De Luca G., Capitanio F., 2022. Agricultural diversification, productivity, and food security across time and space. *Agricultural Economics*, 53(S1): 41-58. https://doi.org/10.1111/agec.12742.

Chen J., Katchova A.L., Zhou C., 2021. Agricultural loan delinquency prediction using machine learning methods. *International Food and Agribusiness Management Review*, 24(5): 797-812. https://doi.org/10.22434/IFAMR2020.0019.

Das S., Sharma K.K., 2023. Credit Debt and Over-in-

debtedness Among the Agrarian Community: A Case Study in Krishnagar-I C.D. Block of Nadia District, West Bengal. *Contemporary Voice of Dalit*, online first. https://doi.org/10.1177/2455328X231176469.

Dey K., Mishra P.K., 2022. Mainstreaming blended finance in climate-smart agriculture: Complementarity, modality, and proximity. *Journal of Rural Studies*, 92: 342-353. https://doi.org/10.1016/j.jrurstud.2022.04.011.

Đokić D., Novaković T., Tekić D., Matkovski B., Zekić S., Milić D., 2022. Technical Efficiency of Agriculture in the European Union and Western Balkans: SFA Method. *Agriculture*, 12(12): 1992. https://doi.org/10.3390/agriculture12121992.

Elnaggar M.A., EL Azeem M.A., Maghraby F.A., 2020. *Machine Learning Model for Predicting Non-performing Agricultural Loans*, pp. 395-404. https://doi.org/10.1007/978-3-030-44289-7_37.

Endris E., Kassegn A., 2023. Analysis of growth and constraints of agricultural micro- and small-scale enterprises in North Wollo Zone, Amhara Regional State, Ethiopia. *Cogent Social Sciences*, 9(1): https://doi.org/10.1080/23311886.2023.2197291.

Eroglu A., Ozturk H., 2016. New mathematical annuity models in a skip payment loan with rhythmic skips. *The Engineering Economist*, 61(1): 70-78. https://doi.org/10.1080/0013791X.2015.1095382.

European Commission, 2024. *Common Agricultural Policy*. Retrieved May 15, 2024, from https://agriculture.ec.europa.eu/common-agricultural-policy_en#:~:text=The%20common%20agricultural%20policy%20(CAP,and%20keeps%20rural%20areas%20vibrant.

Feder G., Lau L.J., Lin J.Y., Xiaopeng L., 1989. Agricultural credit and farm performance in China. *Journal of Comparative Economics*, 13(4): 508-526. https://doi.org/10.1016/0147-5967(89)90024-3.

Formato R.A., 1992. Generalized Formula for the Periodic Payment in a Skip Payment Loan with Arbitrary Skips. *The Engineering Economist*, 37(4): 355-359. https://doi.org/10.1080/00137919208903080.

Gong M., Elahi E., 2022. A nexus between farmland rights, and access, demand, and amount of agricultural loan under the socialist system of China. *Land Use Policy*, 120: 106279. https://doi.org/10.1016/j.landusepol.2022.106279.

Hardy W.E., Weed J.B., 1980. Objective Evaluation for Agricultural Lending. *Journal of Agricultural and Applied Economics*, 12(1): 159-164. https://doi.org/10.1017/S0081305200015429.

Havemann T., Negra C., Werneck F., 2020. Blended finance for agriculture: exploring the constraints and possibilities of combining financial instruments for sustainable transitions. *Agriculture and Human Values*, 37(4): 1281-1292. https://doi.org/10.1007/s10460-020-10131-8.

Heyl K., Döring T., Garske B., Stubenrauch J., Ekardt F., 2021. The Common Agricultural Policy beyond 2020: A critical review in light of global environmental goals. *Review of European, Comparative & International Environmental Law*, 30(1): 95-106. https://doi.org/10.1111/reel.12351.

Huang J., Wang Y., 2014. Financing Sustainable Agriculture Under Climate Change. *Journal of Integrative Agriculture*, 13(4): 698-712. https://doi.org/10.1016/S2095-3119(13)60698-X.

Humbatova S.I., Hajiyev N.G.O., 2021. Investment and loaning in Azerbaijan agriculture. *Bulgarian Journal of Agricultural Science*, 26(6): 1116-1128.

Ifft J., Kuethe T.H., Lyons G., Schultz A., Zhu J.Y., 2023. Crop insurance's impact on commercial bank loan volumes: Theory and evidence. *Applied Economic Perspectives and Policy*, 46(1): 318-337. https://doi.org/10.1002/aepp.13388.

Isibor C., Nkamigbo D.C., 2019. Economic determinants of loan repayment to large and small scale farmer-beneficiaries of bank of agriculture loans from 2010-2016 in Anambra State, Nigeria. *International Journal of Agricultural Policy and Research*, 7(4): 91-99. https://doi.org/10.15739/IJAPR.19.010.

Izraelov M., Silber J., 2019. An assessment of the global food security index. *Food Security*, 11(5): 1135-1152. https://doi.org/10.1007/s12571-019-00941-y.

Jovanović O., Zubović J., 2019. The importance of subsidies for SME development in the agricultural sector of Serbia. *Western Balkan Journal of Agricultural Economics and Rural Development*, 1(1): 51-61. https://doi.org/10.22004/ag.econ.301955.

Kadanali E., Kaya E., 2020. Agricultural Loan and Agricultural Production Value in Turkey. *Alinteri Journal of Agriculture Science*, 35(1): 93-98. https://doi.org/10.28955/alinterizbd.740339.

Kassegn A., Endris E., 2022. Factors affecting loan repayment rate among smallholder farmers got loans from the Amhara Credit and Saving Institution: In the case of *Habru* District, Amhara Regional State, Ethiopia. *International Area Studies Review*, 25(1): 73-96. https://doi.org/10.1177/22338659211040993.

Kaya E., Kadanalı E., 2022. The nexus between agricultural production and agricultural loans for banking sector groups in Turkey. *Agricultural Finance Review*, 82(1): 151-168. https://doi.org/10.1108/AFR-09-2020-0149.

Kong S.T., Loubere N., 2021. Digitally Down to

the Countryside: Fintech and Rural Development in China. *The Journal of Development Studies*, 57(10): 1739-1754. https://doi.org/10.1080/00220388.2021.1919631.

Kulawik J., Wieliczko B., Soliwoda M., 2018. Is there room for financial instruments in the Common Agricultural Policy? Casus of Poland. In: *The Common Agricultural Policy of the European Union – the present and the future EU Member States point of view*, Instytut Ekonomiki Rolnictwa i Gospodarki Zywnosciowej Panstwowy Instytut Badawczy, pp. 34-42. https://doi.org/10.30858/pw/9788376587431.3.

Kusek G., Turker M., Akdemir S., Hayran S., 2017. Structural characteristics of the agricultural sector in terms of access to agricultural credits in Turkey. *New Medit*, 16(4): 66-72.

Lin C., He L., 2020. Targeted monetary policy and agriculture business loans. *The North American Journal of Economics and Finance*, 54: 101290. https://doi.org/10.1016/j.najef.2020.101290.

Ljubičić N., Popović V., Kostić M., Pajić M., Buđen M., Gligorević K., Dražić M., Bižić M., Crnojević V., 2023. Multivariate Interaction Analysis of Zea mays L. Genotypes Growth Productivity in Different Environmental Conditions. *Plants*, 12(11): 2165. https://doi.org/10.3390/plants12112165.

Maricic M., Bulajic M., Dobrota M., Jeremic V., 2016. Redesigning the Global Food Security Index: a multivariate composite I-distance indicator approach. *International Journal of Food and Agricultural Economics*, 4(1): 69-86. https://doi.org/10.22004/ag.econ.231376.

Martin S.J., Clapp J., 2015. Finance for Agriculture or Agriculture for Finance? *Journal of Agrarian Change*, 15(4): 549-559. https://doi.org/10.1111/joac.12110.

Matkovski B., Zekić S., Jurjević Ž., Đokić D., 2022. The agribusiness sector as a regional export opportunity: evidence for the Vojvodina region. *International Journal of Emerging Markets*, 17(10): 2468-2489. https://doi.org/10.1108/IJOEM-05-2020-0560.

McKenzie F.C., Williams J., 2015. Sustainable food production: constraints, challenges and choices by 2050. *Food Security*, 7(2): 221-233. https://doi.org/10.1007/s12571-015-0441-1.

Miglietta P.P., Porrini D., Fusco G., Capitanio F., 2020. Crowding out agricultural insurance and the subsidy system in Italy: empirical evidence of the charity hazard phenomenon. *Agricultural Finance Review*, 81(2): 237-249. https://doi.org/10.1108/AFR-04-2020-0061.

Miglionico A., 2022. Digital payments system and market disruption. *Law and Financial Markets Review*, 16(3): 181-196. https://doi.org/10.1080/17521440.2023.2215481.

Milić D., Mijić K., Jakšić D., 2018. Opportunistic management behavior in reporting earnings of agricultural companies. *Custos e Agronegocio*, 14(1): 125-142.

Milić D., Tekić D., Novaković T., Zekić V., Popov M., Mihajlov Z., 2022. Credit Rating of Agricultural and Food Companies in Vojvodina. *Contemporary Agriculture*, 71(1-2): 51-56. https://doi.org/10.2478/contagri-2022-0008.

Mirč D., Milić D., Tica N., 2023. Status of Agricultural Farmers in the Republic of Serbia in Relation to Banking Institutions. *Contemporary Agriculture*, 72(3): 160-164. https://doi.org/10.2478/contagri-2023-0020.

Mirović V., Bolesnikov D., 2013. Application of Asset Securitization in Financing Agriculture in Serbia. *Economics of Agriculture*, 60(3): 551-564. https://doi.org/10.22004/ag.econ.158255.

Mishra K., Gallenstein R.A., Miranda M.J., Sam A.G., Toledo P., Mulangu F., 2021. Insured Loans and Credit Access: Evidence from a Randomized Field Experiment in Northern Ghana. *American Journal of Agricultural Economics*, 103(3): 923-943. https://doi.org/10.1111/ajae.12136.

Moon I., 1994. Generalized formula for the periodic geometric gradient series payment in a skip payment loan with arbitrary skips. *The Engineering Economist*, 39(2): 177-185. https://doi.org/10.1080/00137919408903120.

Murungi K., Alhassan A.L., Zeka B., 2023. Regulation and agriculture financing in Kenya. *Agricultural Finance Review*, 83(4/5): 783-799. https://doi.org/10.1108/AFR-10-2022-0130.

Nohara Y., Matsumoto K., Soejima H., Nakashima N., 2022. Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Computer Methods and Programs in Biomedicine*, 214: 106584. https://doi.org/10.1016/j.cmpb.2021.106584.

Olagunju F.I., Adeojo R.J., Ayojimi W., Awe T.E., Oriade O.A., 2023. Causal nexus between agricultural credit rationing and repayment performance: A two-stage Tobit regression. *AIMS Agriculture and Food*, 8(1): 151-163. https://doi.org/10.3934/agrfood.2023008.

Onyiriuba L., Okoro E.U.O., Ibe G.I., 2020. Strategic government policies on agricultural financing in African emerging markets. *Agricultural Finance Review*, 80(4): 563-588. https://doi.org/10.1108/AFR-01-2020-0013.

Ozalp A., 2019. Financial Analysis of Agricultural Development Cooperatives: A Case of Western

Mediterranean Region, Turkey. *New Medit*, 18(2): 119-132. https://doi.org/10.30682/nm1902h.

Pejak B., Lugonja P., Antić A., Panić M., Pandžić M., Alexakis E., Mavrepis P. , Zhou N., Marko O., Crnojević V., 2022. Soya Yield Prediction on a Within-Field Scale Using Machine Learning Models Trained on Sentinel-2 and Soil Data. *Remote Sensing*, 14(9): 2256. https://doi.org/10.3390/rs14092256.

Peprah J.A., Koomson I., Sebu J., Bukari C., 2021. Improving productivity among smallholder farmers in Ghana: does financial inclusion matter? *Agricultural Finance Review*, 81(4): 481-502. https://doi.org/10.1108/AFR-12-2019-0132.

Pishbahar E., Ghahremanzadeh M., Ainollahi M., Ferdowsi R., 2015. Factors Influencing Agricultural Credits Repayment Performance among Farmers in East Azarbaijan Province of Iran. *Journal of Agricultural Science and Technology*, 17(5): 1095-1101.

Poliatykina L., Samoshkina I., Borisova V., 2022. Accounting and Financial Problems of Bank Lending to Agro-Industrial Enterprises in Ukraine. *Baltic Journal of Economic Studies*, 8(1): 126-133. https://doi.org/10.30525/2256-0742/2022-8-1-126-133.

Radulović M., Brdar S., Pejak B., Lugonja P., Athanasiadis I., Pajević N., Pavić D., Crnojević V., 2023. Machine learning-based detection of irrigation in Vojvodina (Serbia) using Sentinel-2 data. *GIScience & Remote Sensing*, 60(1): https://doi.org/10.1080/15481603.2023.2262010.

Rathore R., Mishra S., Kumar P., 2017. Factors Affecting Non-Repayment of Agricultural Loan: A Case Study of Rajasthan Marudhara Gramin Bank. *International Journal of Current Microbiology and Applied Sciences*, 6(4): 1052-1059. https://doi.org/10.20546/ijcmas.2017.604.130.

Ray P., Das B., 2023. Agricultural Credit Utilization and Repayment by Farm Households in Tripura. *Indian Journal of Extension Education*, 59(2): 30-35. https://doi.org/10.48165/IJEE.2023.59207.

Shi D., Zhou F., Mu W., Ling C., Mu T., Yu G., Li R., 2022. Deep insights into the viscosity of deep eutectic solvents by an XGBoost-based model plus SHapley Additive exPlanation. *Physical Chemistry Chemical Physics*, 24(42): 26029-26036. https://doi.org/10.1039/D2CP03423A.

Sileshi M., Nyikal R., Wangia S., 2012. Factors Affecting Loan Repayment Performance of Smallholder Farmers in East Hararghe, Ethiopia. *Developing Country Studies*, 2(11): 205-213.

SORS, 2023a. Statistical Office of the Republic of Serbia. Retrieved January 22, 2024, from https://www.stat.gov.rs/en-us/oblasti/poljoprivreda-sumarstvo-i-ribarstvo/.

SORS, 2023b. Statistical Office of the Republic of Serbia. Retrieved November 15, 2023, from https://data.stat.gov.rs/?caller=SDDB.

Staniszewski J., Borychowski M., 2020. The impact of the subsidies on efficiency of different sized farms. Case study of the Common Agricultural Policy of the European Union. *Agricultural Economics (Zemědělská Ekonomika)*, 66(8): 373-380. https://doi.org/10.17221/151/2020-AGRICECON.

Tang L., Sun S., 2022. Fiscal incentives, financial support for agriculture, and urban-rural inequality. *International Review of Financial Analysis*, 80: 102057. https://doi.org/10.1016/j.irfa.2022.102057.

Tekić D., Mutavdžić B., Milić D., Novković N., Zekić V., Novaković T., 2021. Credit risk assessment of agricultural enterprises in the Republic of Serbia: Logistic regression vs discriminant analysis. *Economics of Agriculture*, 68(4): 881-894. https://doi.org/10.5937/ekoPolj2104881T.

Teye E.S., Quarshie P.T., 2022. Impact of agricultural finance on technology adoption, agricultural productivity and rural household economic wellbeing in Ghana: a case study of rice farmers in Shai-Osudoku District. *South African Geographical Journal*, 104(2): 231-250. https://doi.org/10.1080/03736245.2021.1962395.

Wieland R., Lakes T., Nendel C., 2021. Using Shapley additive explanations to interpret extreme gradient boosting predictions of grassland degradation in Xilingol, China. *Geoscientific Model Development*, 14(3): 1493-1510. https://doi.org/10.5194/gmd-14-1493-2021.

World Bank, 2023, September 19. *Agriculture and Food: Overview*. Retrieved October 27, 2023, from https://www.worldbank.org/en/topic/agriculture/overview.

Xia Y., Long H., Li Z., Wang J., 2022. Farmers' Credit Risk Assessment Based on Sustainable Supply Chain Finance for Green Agriculture. *Sustainability*, 14(19): 12836. https://doi.org/10.3390/su141912836.

Zheng H., Zhang Z., 2021. Analyzing Characteristics and Implications of the Mortgage Default of Agricultural Land Management Rights in Recent China Based on 724 Court Decisions. *Land*, 10(7): 729. https://doi.org/10.3390/land10070729.

Household willingness to contribute financially to improve rural water access in Tunisia

Ons Kammarti*, Marielle Montginoul**, Hamadi Habaieb***

DOI: 10.30682/nm2404f
JEL codes: C5, R2

Abstract

Within the context of Tunisia's ongoing water sector reforms that aim to develop a long-term strategy to ensure the reliability and sustainability of supply in rural areas, this study aims to examine the willingness of households individually connected to locally managed water networks to contribute financially towards the rehabilitation of an existing water supply system and the implementation of a new water supply system managed by the national operator. The findings indicate that households are willing to pay 22% of the per-household costs of rehabilitation of an existing locally managed network, representing the equivalent of 2.2 times the interprofessional guaranteed minimum wage (SMIG), and 16% of the per-household costs of setting up a new water supply network managed by the national operator, 3.4 times the SMIG. The findings also indicate that the willingness to contribute financially towards the implementation of a given project is influenced by household characteristics and the state of the water supply within a household.

***Keywords**: Contingent valuation, Drinking water access, Tunisia, Water distribution networks, Willingness to pay*

Introduction

The UN defines 'safely managed' drinking water as water from an improved source that is accessible on the premises, available when needed, and free from fecal and priority chemical contamination. Between 2015 and 2022, there was a noticeable increase in the global coverage of safely managed drinking water. Data on indicator 6.1.1, which tracks safely managed drinking water in 142 countries representing 51% of the global population, show that global coverage rose from 69% in 2015 to 73% in 2022. The progress in rural areas is more marked, with coverage increasing from 56% to 62%, while urban areas saw a slight increase from 80% to 81% (WHO/UNICEF, 2023).

Despite these gains, the pace of progress remains insufficient to meet the Sustainable Development Goal (SDG) 6.1 target of universal access by 2030. To achieve universal coverage by 2030, the current rate of progress must be accelerated dramatically. Specifically, a six-fold increase in the current rate of improvement is required globally. This challenge is even more pronounced in low-income countries, where the

* National Agronomic Institute of Tunisia (INAT), University of Carthage, Tunis, Tunisia / INRAE UMR G-EAU, Montpellier University, Montpellier, France.
** INRAE UMR G-EAU, Montpellier University, Montpellier, France.
*** National Agronomic Institute of Tunisia (INAT), University of Carthage, Tunis, Tunisia.
Corresponding author: kammarti.ons@gmail.com

rate of progress would need to increase 20-fold to meet the targets (WHO/UNICEF, 2023).

The water sectors in various developing countries have distinct histories and challenges shaped by each country's unique policy and institutional settings. Nevertheless, they share some common pressing issues. These governments face a complex issue that requires resolution, as they are confronted with both internal social and political demands for improved water services, as well as increasing international pressure for the same. However, they have limited flexibility to take action and fulfill these demands due to the growing constraints on available water resources and finances (Favre, 2021).

Two primary factors hinder progress towards equality in access to water between densely inhabited urban areas and sparsely populated rural areas, and impede efforts to connect the unconnected population to water supply networks: the necessity to effectively manage the use of water resources that are becoming increasingly scarce, and the challenge of financing infrastructure in a sustainable manner (Favre, 2021).

Over several decades, Tunisia has made extensive investments to enhance access to drinking water in rural areas, resulting in a significant increase in individual connections to water supply systems. However, the water sector has to face a number of constraints, in particular: a physical constraint associated with the scarcity of water, the deterioration of infrastructure and difficulties in financing projects to rehabilitate and extend the hydraulic networks that remain entirely funded by the Ministry of Agriculture, Water Resources, and Fisheries (MARHP) in rural areas. These constraints are likely to affect the sustainability of drinking water supply and thus threaten to achieve the sustainable development goal of "Ensuring universal access to water and sanitation (WASH)".

Several studies have noted the lack of information regarding household preferences as an impediment to implementing improved water services in developing nations (Vásquez, 2014; Whittington *et al*., 1990). In response to the need for preference information, contingent valuation (CV) studies have been employed to elicit willingness to pay (WTP) for improved water services in developing countries (Vásquez, 2014; Whittington, 1998). CV surveys have proven useful for collecting information on preferences for public goods and services in these contexts (Whittington, 1998). Consequently, CV findings have been used to inform policy decisions in various ways, including cost-benefit analyses and setting price rates for improved services (Vásquez *et al*., 2009). An improved understanding of WTP can help assess the economic feasibility of system improvements, design price structures, and inform affordability and equity policies in order to implement efficient, sustainable, and cost-recoverable water projects (Vásquez, 2014; Vásquez & Espaillat, 2016).

Considering the evidence that households are willing to pay significant amounts for safe drinking water (Vásquez *et al*., 2009) and within the context of Tunisia's ongoing water sector reforms, the present study aims to examine the willingness of households individually connected to networks managed by Water Associations known as Agricultural Development Groups to financially contribute towards 1) the rehabilitation of an existing water supply system and 2) the implementation of a new water supply system managed by the national water utility, using the survey-based contingent valuation method. The objective of these financial contributions is to help improve damaged rural water supply systems, which largely depend on limited government and external funding. The rest of the paper is organized as follows: The next section provides an overview of the water sector in Tunisia. The subsequent section describes the study area, study design, sampling methodology, and data analysis strategy. The following section presents the survey results including willingness-to-pay (WTP) estimates for the proposed projects and examines the determinants that influenced respondents' WTP. The final sections discuss the findings, highlight some policy implications, and provide conclusions.

1. An Overview of Tunisia—s Water Sector: Past Achievements, Current Challenges, Future Directions, and Rural Realities

In recent decades, Tunisia has made considerable efforts to improve water coverage rates throughout the country (Favre, 2021). Urban ar-

eas attained a 100% coverage rate well before 2021, whereas rural areas recorded a 95% coverage rate in 2021 (MARHP, 2021). The MARHP has set a target to achieve universal access to water (SDG 6.1) by 2030. Despite significant progress, Tunisia continues to face challenges due to water scarcity, irregular supply, overexploitation, increasing demand, and suboptimal regulation mechanisms. These challenges are anticipated to worsen with the impacts of climate change, potentially hindering the goal of universal access (MARHP, 2021).

Recognizing the critical need for qualitative changes in public water policy, the Tunisian government has identified two key areas for intervention: demand management by reducing specific consumption and improving network efficiency, and supply optimization by investing in large-scale water transfer and desalination plants. Since 2019, the government has been working on a strategic plan titled "L'élaboration de la vision et de la stratégie du secteur de l'eau à l'horizon 2050 pour la Tunisie" ("Development of the Vision and Strategy for the Water Sector in Tunisia by 2050"). This forward-looking strategy aims to renovate a significant portion of the water network by 2050 and introduce smart meters to enhance system efficiency. The overarching goal is to develop a secure and efficient potable water supply system that meets international standards and ensures that rural water services are equivalent to those in urban areas (MARHP, 2021).

Additionally, the Tunisian government is adopting a water-energy-food nexus approach within this strategy to facilitate a transition towards sustainability. Studies, such as those by Scardigno *et al*. (2017), highlight that the interconnections between water and food security, particularly in the Middle East and North Africa (MENA) region, are likely to face increasing risks due to growing demand, resource constraints, and the escalating impacts of climate change. By integrating this nexus approach, Tunisia aims to mitigate these risks and promote sustainable resource management (MARHP, 2021).

In Tunisia, the rural potable water sector is characterized by the coexistence of two distinct management systems: the distribution of water in highly populated rural areas is ensured by the National Company for the Exploitation and Distribution of Water (SONEDE). Water supply in sparsely populated areas is locally managed by Water Associations called Agricultural Development Groups (GDAs) with technical support from the Regional Agricultural Development Offices (CRDAs), the regional departments of the Ministry of Agriculture, Water Resources, and Fisheries, the actor responsible for implementing rural water projects.

The SONEDE water bills include either water or water plus sewerage. All users subscribing to the water service provided by the SONEDE are subject to paying a sewer fee in areas with collective sanitation services provided by the National Sanitation Office (ONAS) and receive a unique quarterly bill based on their water consumption. The SONEDE nationally applies a binomial tariff structure composed of fixed and variable parts. The variable part is a combined Inclining Rate Tariff (IRT) whereby the tariff depends on the water consumption level and total consumption is charged at the rate of the top band and Increasing Block Tariff (IBT) whereby initial blocks of consumption are charged at a lower rate than the additional blocks of consumption for users charged for water and collective sanitation, and an IRT for users only charged for water. In addition to the variable part, both operators apply a fixed charge for every user based on the water meter diameter (the larger the diameter, the higher the charge) for the SONEDE and on the water consumption levels for the ONAS (Favre & Montginoul, 2018).

Water pricing structures vary greatly among Agricultural Development Groups due to the diverse nature of their systems. Households connected to GDA water networks do not have access to the collective sanitation service and rely on non-collective sanitation solutions.

2. Methods

2.1. *Selection of the Study Area*

This study was conducted in Mornag, a municipality located in the governorate of Ben Arous in Tunisia. The study area was selected

due to its rural nature, the range of population densities within it—from sparsely populated areas to highly populated rural areas (we used the term 'highly populated rural areas' due to the lack of a clear definition and boundaries for peri-urban areas), and the coexistence of the two management systems. The selection of the study area and the Contingent Valuation (CV) survey were preceded by a series of semi-structured interviews. Their purpose was to understand the complexity of providing water services in rural areas, to identify the diversity of actors involved, to determine the characteristics of the selected study area, and to select localities. The interviews were conducted with experts from the General Directorate of Rural Engineering and Water Use belonging to the Ministry of Agriculture, Water Resources, and Fisheries, and the Regional Agricultural Development Office of Ben Arous.

Eleven localities of varying densities were chosen. A locality refers to a territorial area served by its own water distribution network managed by a GDA. This network can be either independent or connected to another hydraulic network. The initial plan was to survey 30 households randomly selected in each locality, except in some localities that had fewer than 30 households, particularly in low-density areas. In practice, surveying ceased once repetitive responses about access to alternative water sources and coping strategies started being detected, signifying that further surveying would not yield new valuable insights.

2.2. *Questionnaire and Sample*

To achieve the objective of this study, a CV survey was conducted from July to August 2022. The questionnaire was pretested through a pilot survey administered to 20 households. The final survey questionnaire was then administered to 161 respondents who self-identified as having sufficient knowledge about the current water situation within their households.

The questionnaire was divided into two sections. The first section collected information about the socioeconomic characteristics of the respondents/households, the house condition, information related to the piped water including the water bill amount, water consumption, its uses, perception of the piped water quality, issues with water services, and access to alternative water sources. The second section described the scenarios of the CV study.

Prior to conducting face-to-face interviews that lasted between 20 and 80 minutes, informed consent was obtained from all participants.

3. Study Design

Limited government and external funding have slowed progress on improving damaged rural water supply systems. The CV scenarios are based on the assumption that, to address this issue and potentially accelerate the government's efforts to improve the rural water sector, requiring households to contribute financially towards the rehabilitation of an existing water supply system or the implementation of a new water supply system in their neighborhood may be worthwhile.

Two household-cost-sharing scenarios were initially designed to assess the willingness of households to contribute financially towards rehabilitating an existing water supply system or implementing a new water supply system in their neighborhood and to determine whether the same households would be willing to pay for the resulting service on a regular basis once the water supply system is renewed or installed, assuming that the project would be implemented with or without the households' contributions. The CV scenarios were designed to provide information about the terms and conditions of both conceptually interrelated components (Whittington, 1998), enabling respondents to make informed choices.

This approach was similar to the approaches employed by Adams and Vásquez (2019) in Ghana and Favre (2021) in Tunisia. Specifically, Favre (2021) asked respondents two independent questions. The first question examined willingness to pay for a sustainable individual connection to the SONEDE network for non-piped households. The second question assessed willingness to pay regular bills for water consumption in order to maintain service quality in

the future. Favre (2021) justified the use of this approach as the two stated WTP did not have the same occurrence, nor the same payment vehicle. By presenting the questions separately, respondents were able to distinguish between the two different commitments and provide independent WTP responses.

The initial amounts proposed to households corresponded to the per-household costs of a hydraulic project (either rehabilitation of an existing GDA network or setting up a SONEDE network) and were based on estimates provided by the Ministry of Agriculture, Water Resources and Fisheries. These estimates are presented to funders to offer a general financial appraisal of a project in the conceptual stage, and are calibrated based on the specific nature and setting of the envisaged project. The initial amounts served as a starting point for evaluating households' willingness to pay for the proposed infrastructure projects. The study proposed one or two scenarios, depending on the current condition of the network (Table 1). The amounts were converted from Tunisian dinars to US dollars using the October 2022 exchange rate of 1 TND = 0.31 USD.

Table 1 - Summary of the study design.

Target (Scenario)	*Statement*	*Elicitation format*	*Rationale*
Households connected to a network managed by a GDA presenting anomalies (Scenario 1)	1. Would you be willing to pay 4000 TND to participate in the renewal of the existing network managed by the GDA in order to benefit from an improved service consisting of a continuous supply of good quality water?	Using a descending bidding game method with increments of 500 TND (157 USD)	To investigate willingness to financially contribute towards the rehabilitation of the existing infrastructure allowing the improvement of the GDA's performance
	2. Once the service improved, would you be willing to pay 2 TND/m^3 consumed? *(Abandoned question)*	Using a descending bidding game method with increments of 0.1 TND (0.03 USD) until reaching the current applied charge per cubic meter	
Households connected to a network managed by a GDA (Scenario 2)	1. Would you be willing to pay 8000 TND to participate in the implementation of a new network managed by the SONEDE.	Using a descending bidding game method with increments of 500 TND (157 USD)	To investigate willingness to financially contribute towards the implementation of a new water supply system managed by the SONEDE.
	2. Once connected to the SONEDE's network, would you be willing to pay 22 TND per quarter for a water consumption of 20 m^3? *(Abandoned question)*	Using an ascending bidding game method with increments of 5 TND (1.57 USD)	

3.1. *Scenario proposed to households connected to a GDA network presenting anomalies (Scenario 1)*

The first scenario targeted households connected to a network managed by a GDA, presenting anomalies resulting from poor infrastructure that led to frequent interruptions, or caused by its dependency on the SONEDE for its water supply in cases where the GDA receives its supply from the SONEDE. The sample size for this scenario was 111 households connected to 8 different GDA networks. The initial scenario consisted of two questions. Households would have been first asked if they would be willing to pay 4000 TND (1256 USD) to participate in the renewal of the existing network managed by the GDA in order to benefit from an improved service consisting of a continuous supply of good quality water. This question was presented in a bidding game format. Households would then have been asked if they would be willing to pay 2 TND (0.63 USD) per cubic meter consumed once the service is improved/the project is completed. The second question was also presented in a bidding game format. These two questions were preceded by an introductory sentence developed to help households understand the scenario. The introductory sentence mentioned that the government needs funding in order to accelerate efforts to improve the water sector and repair damaged systems, and that financial contributions from households could help (Table 1).

Upon further reflection, only the first question was kept because it would be difficult to accurately determine the real costs per cubic meter for each locality before the projects are actually completed, due to the multiple location-specific variables that would impact those costs, including: the current costs of water per cubic meter, hydrological factors like reliance on groundwater and associated energy costs, and geographical factors like spatial distribution. Additionally, we knew that households familiar with the service are conscious that it requires payment on a regular basis and that improvements imply higher charges.

The rationale for studying this scenario was to investigate the willingness to financially contribute towards the rehabilitation of the existing network while maintaining the same management system.

3.2. *Scenario proposed to all households connected to a GDA network (Scenario 2)*

The second scenario targeted all households connected to a network managed by a GDA, regardless of whether the current network presented anomalies or was well managed. The sample size for this scenario was 161 households connected to 11 different GDA networks. The initial scenario included two questions presented in bidding game format. Households would have been first asked if they would be willing to pay 8000 TND (2512 USD) to participate in the implementation of a new network managed by the SONEDE. They would then have been asked if once connected to the SONEDE's network, they would be willing to pay 22 TND (6.91 USD) per quarter, a fee based on a projection made by the utility, allowing the costs to be recovered over the next three years for a water consumption of 20 m^3. This volume corresponds to a consumption equal to 50 liters/person/day, a volume allowing reaching the intermediate access level, to meet the needs related to direct consumption, personal hygiene and basic domestic uses and to maintain health problems at a low level (Howard & Bartram, 2003). These two questions were preceded by an introductory sentence developed to help households understand the scenario. The introductory sentence mentioned that the government needs funding to accelerate efforts to improve the water sector and set up a new water supply system allowing future intervention by the SONEDE in the neighborhood, and that financial contributions from households could help (Table 1).

The second question was eliminated because we expected that households familiar with simpler payment modes would have difficulties relating charges to the volume of water consumed.

The rationale for studying this scenario was to investigate the willingness to financially contribute towards the implementation of a new water supply system managed by the SONEDE.

4. Dealing with biases

This section outlines the potential sources of bias and ways in which we sought to mitigate these biases.

4.1. *Hypothetical bias*

Hypothetical bias refers to the tendency of respondents in contingent valuation surveys to overestimate their willingness to pay for public goods (Carson, 2012). Several steps were taken to mitigate this bias in respondents' willingness to pay estimates. First, the survey scenarios were not presented as purely hypothetical. Avoiding telling respondents repeatedly that the scenarios are hypothetical helps avoid inflating their willingness to pay, as studies have found that scenarios presented as purely hypothetical tend to overestimate willingness to pay (Carson, 2012). Making the scenarios seem more realistic by not emphasizing their hypothetical nature encourages respondents to think more carefully about their actual willingness to pay (Carson, 2012). Furthermore, we assumed that for public services in developing countries, the issue of hypothetical bias tends to be less significant as respondents are more familiar with the service characteristics and less likely to misunderstand the scenario (Whittington *et al*., 1990), and that households are familiar with the water service provided, based on the fact that they are already connected to the network. Familiarity with the good or service has been shown to reduce the hypothetical bias in CV studies. Studies show that the greater the familiarity of respondents with a good or service, the less the hypothetical bias in their willingness to pay estimates (Mitchell and Carson, 1989). WTP estimates for public goods tend to be free from hypothetical bias (Whittington *et al*., 1990).

4.2. *Strategic bias*

Strategic bias is a type of respondent bias that can be problematic in contingent valuation (CV) studies. It occurs when respondents deliberately misstate their true willingness to pay (WTP) for strategic reasons. There are two forms of strategic behavior: free-riding and over-pledging (Mitchell and Carson, 1989).

Free-riding occurs when an individual understates their true WTP for a public good on the expectation that others will pay enough for that good; therefore, they do not have to pay. On the other hand, over-pledging occurs when an individual overstates their WTP. They do this, assuming their stated WTP will influence the provision of the good, under the belief that it will not actually be used for future pricing policy (Venkatachalam, 2004).

Adopting recommendations (Mitchell and Carson, 1989) to mitigate strategic bias could improve the likelihood of useful WTP estimates. Telling respondents that a larger number of people are being interviewed can give them the impression that their stated WTP will not meaningfully influence the outcome. Giving respondents the impression that the good may not be provided if WTP is understated may discourage the understatement of true willingness to pay. Using payment vehicles that remind respondents of their budget constraints can help discourage the overestimation of the true WTP.

4.3. *Starting-Point Bias*

We assumed that if we choose an open-ended elicitation format to ask households about their potential contribution without mentioning the real costs of the projects, households will tend to answer by giving a willingness to pay value that could be significantly lower than the real costs of the projects. Starting point bias arises when the initial bid (posited by the interviewer) influences respondents' final bids (Boyle *et al*., 1985). The choice of the descending bidding game approach allowed us to mention the estimated project costs as a starting point, thus avoiding this type of bias.

4.4. *Scope Test*

A scope test examines whether respondents are willing to pay more for a good that is larger in scope, either in terms of quality or quantity (Carson *et al*., 2001). This study did not conduct scope tests to examine how willingness to pay

estimates varied with different water volumes. There were three key reasons for not considering the scope effects. First, statistical scope insensitivity is insufficient to fully invalidate the results of the contingent valuation study. Second, with the small sample size, the study would lack adequate statistical power to detect a scope effect if one was present. Third, it is challenging, if not impossible, to statistically observe a scope effect when eliciting WTP for high levels of environmental goods (Lopes & Kipperberg, 2020). The effect of increased volumes of water on willingness to pay estimates was not considered by fixing the volume parameter, corresponding to one cubic meter, and 20 cubic meters per quarter for the first and second scenarios, respectively.

4.5. *Metric bias*

Metric bias occurs when respondents assign a value to the good based on a different metric or scale, typically one that is less precise than the metric intended by the researcher designing the contingent valuation study (Mitchell and Carson, 1989). To anticipate potential metric bias, we compared 1 cubic meter to a 1000-liter water container size. By comparing 1 cubic meter to the volume of a container commonly used for water storage in the study area, we aimed to help respondents conceptualize the volume amount so that their valuation would be based on the defined measurement scale.

4.6. *Sequencing*

The sequencing effect, also known as question order bias, may occur when the order in which questions are asked can influence respondents' willingness to pay values (Cummings *et al*., 1986; Mitchell and Carson, 1989). We expected that this type of bias may occur for the group of households who received the first and second scenarios. We hypothesized that for the second scenario described, respondents may use the starting value provided for the first scenario as an anchoring point, even though the costs for the implementation of a new water supply system managed by the SONEDE are actually higher.

To investigate this potential sequencing effect and its derived anchoring effect, we planned to split respondents into two equal groups and change the order of the scenarios, asking about one scenario first for one group and the other scenario first for the second group. Comparing the values between the groups would indicate if respondents anchor on the initial starting value provided. This test was not conducted. Instead, we clearly communicated to respondents the differences in costs and standards between the two projects described in an attempt to minimize this type of bias.

5. Dealing with protest responses

In contingent valuation studies, some respondents indicate that they are unwilling to pay any monetary amount for a public good due to disagreements with the procedural elements of the contingent valuation method, such as the perceived unfairness of having to pay more. These types of responses are problematic for benefit-cost analysis, as they may reflect that a person values the good in question but is unwilling to pay for it. From an economic perspective, such responses, considered either suspect or inappropriate, cannot be included in the benefit-cost analysis because they do not represent the "true" economic values. Responses indicating an unwillingness to pay any monetary amount for a public good due to certain circumstances, such as an inability to afford it, are not treated as a protest response because it is accepted in economics that money is not a perfect indicator of utility since levels of wealth vary among individuals (Jorgensen *et al*., 1999). The authors acknowledged that it was difficult to develop an effective strategy that would convince respondents to put aside their impressions of the role of government in providing public goods, issues about equity between social groups, considerations about their own efficacy in influencing outcomes related to public good decision-making, or their evaluations of funding public good changes through additional household contributions. For the present study, all responses connected to elements of the hypothetical scenario were regarded as protests and were eliminated from the analysis.

6. Models

For each scenario, a multiple regression model was constructed with a set of independent variables related to household characteristics and variables related to piped water and its use. Regression analysis tests whether the willingness of households to contribute financially towards rehabilitating an existing water supply system or implementing a new water supply system in their neighborhood can be explained by variables suggested by economic theory, such as income and perception of water quality, as suggested by Whittington *et al*. (1990). The statistical analyses were conducted using the STATA15 software.

7. Variables

A transformation to the dependent variable representing bid amounts was applied prior to statistical modeling, with positive bids log-transformed while zero bid values remained unchanged. Independent variables were divided into two catego-

Table 2 - Independent variables.

Variable	*Description*	*Expected sign*	*Rationale*
related to household characteristics			
Minimum monthly household income	= ln (Minimum monthly household income)	+	The literature discusses various methods for addressing missing income data in CV surveys, including the use of proxies. For example, due to many respondents not reporting their actual income in a survey examining the public's willingness to pay for changes in national forest management conducted by Czajkowski *et al*. (2017), the variable used as an income proxy was the question "How would you rate the financial situation of your household?" measured on 5-point Likert scale where 1 represented a "Bad" financial situation and 5 represented a "Good" financial situation. This approach produced interpretable results. The present study examined incomes of households using a self-reported measure of minimum monthly income required to meet basic needs. Respondents were asked to provide an estimate of the income level needed to support a basic standard of living, rather than reporting their actual income. This approach, used as an alternative to collect income data from respondents, was needed because the length of the survey questionnaire did not allow the inclusion of detailed income questions. It is assumed that households with higher minimum monthly incomes will tend to have higher actual monthly incomes. As a consequence, households with higher minimum monthly household income levels may have greater willingness to contribute to the costs of the proposed project.
Household size		+	Larger households are likely to have higher demand and thus may express higher willingness to pay for the project being studied. A study by Vásquez (2014) found a positive empirical relationship between household size and willingness to pay for improved water services. However, economic theory suggests the effect of household size can be negative because freely disposable income decreases with the number of household members (Ahlheim & Schneider, 2013).

Variable	*Description*	*Expected sign*	*Rationale*
related to piped water, its use, and coping practices with service unreliability			
Using alternative sources requiring fetching water on a regular basis	=1 Yes; =0 No	The sign for this variable cannot be determined with certainty a priori.	A regular reliance on alternative sources due to issues with irregular supply and/or poor tap water quality for households connected to a piped water network may influence WTP in two different ways. Possibility 1: Households relying on alternative water sources will have lower WTP for the proposed project, because these alternative sources are freely available and potentially reduce water bills and expenditures on water in general. Possibility 2: Households relying on alternative water sources will have higher WTP for the proposed project, because it is significantly time-consuming to fetch water from these sources and could impact their ability to generate income through other activities.
Subjective perception of the current piped water quality	=1 Good or rather good; = 0 Otherwise	-	Households may pay less to improve the service if current water quality is good.
Storing piped water (summer 2022)	=1 Yes ; =0 No	+	Households using small plastic containers with a total volume not exceeding 1000 liters may be more impacted by irregular supply. This may reflect a response mechanism to cope with frequent or long interruptions, and thus they may be more willing to pay for improvements. The survey questions clearly differentiated between and distinctly asked about both piped water storage and rainwater storage. This expectation is based on this finding "Households that adopt coping measures to deal with unreliable water services, like treating and storing piped water at home, are also likely to support a project that aimed to ensure reliable supplies of safe drinking water because an improved system would reduce the need for implementing those coping measures" (Vásquez & Espaillat, 2016).
related to piped water, its use, and coping practices with service unreliability			

ries (Table 2): those related to household and individual characteristics, and those related to piped water and its use. Table 2 presents the rationale for including each variable.

8. Results

8.1. *Semi-structured interviews: Key findings*

The semi-structured interviews allowed us to gather information about the per-inhabitant costs of a hydraulic project, either for the rehabilitation of an existing GDA network or the establishment of a SONEDE network. According to the experts we consulted, the SONEDE is responsible for expanding an existing network within a limited geographical area, while the MARHP is responsible for setting up new water supply networks that will be managed by local communities or connecting an entire locality to the SONEDE network. The experts also informed us that although the GDAs operate under the same regulatory framework, they

Variable	*Description*	*Expected sign*	*Rationale*
Household water supply situation (summer 2022)	1 = Continuous supply 2 = More than 12 hours of continuous supply per day 3 = Less than 12 hours of continuous supply per day 4 = Few hours to one full day of continuous supply each 2 to 3 days 5 = Few hours to one full day of continuous supply each 3 to 4 days 6 = Few hours to one full day of continuous supply each 4 to 5 days 7 = Few hours to one full day of continuous supply each 5 to 6 days 8 = Few hours to one full day of continuous supply per week 9 = Few hours to one full day of continuous supply each period greater than 7 days	+	The questionnaire included a question about water interruptions over the past 1-2 months (summer 2022). Responses were classified on an ordered scale reflecting the severity and frequency of reported interruptions, with higher values indicating more severe lack of water supply. The response classification provided a measure of the intermittent nature of the water supply by evaluating the duration of supply periods and intervals between them. Households reporting more days without water would express greater willingness to pay for the proposed project. This rationale is based on the results of a study conducted by Vásquez and Espaillat (2016), which demonstrated that WTP varied depending on the level of system reliability, measured by the number of days in a week with service interruptions. Households who experienced service interruptions every day of the week were willing to pay more than those without service interruptions.

differ significantly in terms of the number of connected households, water sources (either connected to the national operator's network, independently managing groundwater/surface water sources, or relying on a combination of sources), and pricing structures. After explaining the study's purpose, the CRDA provided us with a list of localities within the study area that would be suitable targets for the survey. Eleven localities with varying densities were chosen from a total of approximately 25 to capture the full range of diversity among the GDAs. Among these, 5 are connected to the national operator network, 3 rely on locally-sourced groundwater, 1 is connected to another GDA network, 1 relies on a combination of a connection to the national operator network and a locally-based groundwater source, and 1 relies on surface water. The number of households connected to these networks varies significantly, ranging from approximately 10 to roughly 600, and the length of the networks ranges from 1.5 to 20 kilometers.

8.2. *Protest responses and non-responses*

The protest responses are shown in Table 3, with a distinction between Scenarios 1 and 2.

For Scenario 1, 21 households indicated that they were not willing to pay for the project, justifying their zero WTP response with one or two of the reasons mentioned in Table 3. These responses were regarded as protest responses. The most frequently cited reason was demanding the intervention of the SONEDE in the neighborhood (Table 3). The two responses indicating gender-related issues (inability to make a household decision as a woman) were regarded as non-responses and were eliminated from the analysis.

For Scenario 2, 32 households indicated that they were not willing to pay for the project, justifying their zero WTP response with one or two of the reasons mentioned in Table 3. The most frequently cited reason was a preference for GDA and willing to contribute to its improvement (Table 3). 4 responses indicating gender-related issues (inability to make a household decision as a woman) were

Table 3 - Protest responses.

Reason/Pair of reasons	*Number of responses*	*Percentage*
Scenario 1		
Demanding the intervention of the SONEDE in the neighborhood	11	52%
Inadequate management mode of the GDA	5	24%
Demanding the intervention of the SONEDE in the neighborhood *AND* inadequate management mode of the GDA	5	24%
	21	*100%*
Proportion of non-protest responses	*19%*	
Scenario 2		
Lack of trust in the reliability of the project	1	3%
The service is a fundamental right and must be guaranteed to all individuals	1	3%
Preference for GDA and willing to contribute to its improvement	20	63%
Future service provided by the SONEDE perceived to be expensive or of poor quality AND preference for GDA and willing to contribute to its improvement	3	9%
Lack of trust in the reliability of the project AND future service provided by the SONEDE perceived to be expensive or of poor quality	1	3%
Refusal of the intervention of the SONEDE in the neighborhood AND Preference for GDA and willing to contribute to its improvement	5	16%
Lack of trust in the reliability of the project AND preference for GDA and willing to contribute to its improvement	1	3%
	32	*100%*
Proportion of non-protest responses	*20%*	

regarded as non-responses and eliminated from the analysis. In their study investigating WTP for the improvement of tap water quality in Greece, Polyzou *et al*. (2011) considered the belief that there is no need to improve water quality as a true-zero response. Based on this consideration, we classified the response "Being satisfied with the management by the GDA" and the belief that there are no issues with the service provided (24 responses) as true zeros. The reason "A preference for GDA but unwilling to contribute to its improvement" (1 response) was also regarded as a true zero response. The appropriate handling of protest zeros depends on the specific survey context being examined. Our study classified responses indicating a "lack of trust in the reliability of the project" as protest zeros, whereas Vásquez and Espaillat (2016) included the answer to "if the respondent believed that the proposed project could be implemented" as a predictor in their statistical models. Their results showed that those believing implementation was feasible were more likely to vote in favor of the proposed project.

The proportion of non-protest responses was 19% for Scenario 1 and 20% for Scenario 2 (Table 3). These percentages are probably due to the fact that implementing infrastructure has long been, and continues to be, the responsibility of public institutions. Consequently, the rural community has come to take this for granted.

All protest responses for scenario 1 were directly related to the management mode itself, with refusal to contribute indicating a rejection of the service provider. This suggests that intrinsic factors, such as a lack of trust in the management community of a GDA, influence the respondents' decisions. Additionally, it may indicate that water prices per cubic meter set by the GDAs are perceived to be higher than those set by the SONEDE. We are not asserting that GDA services are more expensive, as it is difficult to compare due to the tariff structure applied by the SONEDE.

97% of protest responses for scenario 2 (excluding the response that the service is a fundamental right and must be guaranteed to all

Table 4 - Willingness-to-pay estimates.

	Observations	*Mean WTP*	*Standard deviation*
Scenario 1	88	864	1240
Scenario 2	125	1318	2380

individuals) were directly related to the service provider. This indicates a refusal of the national operator's services or reveals that these respondents have greater trust in the GDA. In the context of a reform where the government is actively seeking an alternative to the GDA management system, the preference for the service provider itself must be considered.

After removing the protest responses and non-responses from the dataset, the mean amount that households connected to a GDA network presenting anomalies were willing to financially contribute towards the rehabilitation of the existing infrastructure was 864 TND (268 USD), representing 22% of the per-household costs of the project. The mean amount that households connected to a GDA network were willing to financially contribute towards the implementation of a new water supply system managed by the SONEDE was 1318 TND (409 USD), representing 16% of the per-household costs of the project. The contribution for the per-household costs of rehabilitating an existing locally managed network represented the equivalent of 2.2 times the inter-professional guaranteed minimum wage (SMIG) fixed under the terms of Decree n° 2022-769 of 19 October 2022 at 390.7 TND (122.7 USD) (40-hour work week regime), and the contribution for the per-household costs of setting up a new water supply network managed by the national operator represented 3.4 times the SMIG.

8.3. *Estimation results*

As expected, the coefficients of *Minimum monthly household income* were positive and statistically significant at the 1% level in Model 1 referring to the scenario proposed to households connected to a GDA network presenting anomalies, and at the 10% level in Model 2 referring to the scenario proposed to all households connected to a GDA network (Table 5). Our findings are in line with the economic theory. Significant income effects on willingness to pay for improved water services are consistently found in the literature. Vásquez and Espaillat (2016) showed that willingness to pay for reliable supplies of safe potable water increases with monthly household income levels. The coefficients of *Household size* were statistically insignificant in both models (Table 5). A study by Favre (2021) also found the coefficient for household size to be statistically insignificant in the statistical model assessing willingness to pay for sustainable individual connections to the SONEDE network among non-piped households.

Households storing tap water to cope with sudden interruptions in tap water supply or to prevent such disruptions use a variety of small plastic containers to store water at home (Figure 1). Only a relatively small percentage of the surveyed households used devices with large storage capacities. We hypothesized that households storing piped water in plastic containers, with a total volume not exceeding 1000 liters, due to unreliable water supply, would report a higher willingness to pay for the proposed projects. However, the empirical analysis found the coefficients of *Storing piped water* to be negative and nonsignificant (Table 5). Vásquez and Espaillat (2016) also reported insignificant effects of water storage on households' willingness to pay for reliable supplies of safe potable water in Guatemala.

The results of Model 2 require careful consideration. A household's willingness to pay for the implementation of a new water supply system managed by the national water utility may depend on whether they are currently experiencing water interruptions or perceive the water quality as poor. The sample of households surveyed included two groups - those connected to GDAs that were well-managed with no reported interruptions, and those connected to GDAs that experienced frequent interruptions.

Figure 1 - Two examples of a set of small plastic containers used to store piped water at home.

(Source: Authors, July 2022).

This heterogeneity could explain the significantly higher mean value of *Subjective perception of the current piped water quality* (t= 2.3520, p = 0.0196) and the significantly lower mean value of *Household water supply situation* (t= -3.7520, p = 0.0002) compared to the group of the first scenario (Table 5).

The coefficients of *Household water supply situation* were positive and statistically significant at the 10% level in both Model 1 and Model 2. This finding suggests that respondents reporting more days without water express greater willingness to contribute financially to the rehabilitation of the existing infrastructure/the implementation of a new water supply system managed by the SONEDE.

In Model 1, the coefficient of *Subjective perception of the current piped water quality* meets the expected negative sign but does not reach statistical significance. The absence of a statistically significant effect may be partially explained by 75% of the post-protest sample abstaining entirely from drinking tap water. In Model 2, the coefficient of this variable was statistically significant at the 10% level. Respondents reporting good quality of tap water were less willing to contribute financially to the implementation of a new water supply system managed by the SONEDE.

For Scenario 2, the research sought to examine location-specific effects by using dummy variables for each locality/GDA, as per the method of Basani *et al*. (2008). However, the sample size prevented this intended methodology. Instead, the "*Current status of the GDA*" variable was generated to represent whether interruptions were occurring or not at the locality level. This variable considered the overall system performance rather than solely the surveyed households' responses within that system (the status quo service provision within a locality). Even if some households connected to a given GDA did not report interruptions, issues could still persist across the entire infrastructure.

For scenario 2, a simple linear regression was conducted using this "*Current status of the GDA*" variable as a predictor. It achieved statistical significance in predicting the dependent variable (β = 1.900, t = 2.79, p < 0.01). A t-test was used to compare the mean willingness to pay for the implementation of a new water supply system managed by the SONEDE between two groups of households: those connected to well-managed GDAs with no reported interruptions and those connected to GDAs experiencing frequent interruptions. The results revealed that households belonging to the second group had a significantly higher mean willingness to pay for the implementation of a new water supply system managed by the SONEDE (t= -2.1010, p = 0.0377). Households currently not experiencing interruptions or water quality issues were still willing to pay for this project. This suggests that individuals perceive value in a centralized system as a more reliable long-term option, even if the GDA to which they are connected is currently well managed.

Table 5 - Linear regression: variables influencing households' WTP.

Variables	*Model 1 (Scenario1)*	*Mean variable*	*Model 2 (Scenario2)*	*Mean variable*
Ln (Minimum monthly household income)	*2.134***	7.11	*1.147**	7.02
	(2.81)		(1.74)	
Household size	0.197	4.33	0.0280	4.69
	(0.81)		(0.15)	
Using alternatives sources requiring fetching water on a regular basis	-1.322	0.11	0.982	0.20
	(-1.23)		(1.16)	
Subjective perception of the current piped water quality	-0.921	0.20	-1.484*	0.35
	(-1.04)		(-1.91)	
Household water supply situation (summer 2022)	*0.333**	4.44	*0.285**	3.33
	(1.75)		(1.81)	
Storing piped water (summer 2022)	-0.736	0.67	-0.672	0.59
	(-0.99)		(-1.00)	
(constant)	*-12.43***	-	-4.559	-
	(-2.34)	-	(-0.97)	-
R^2	0.1959	-	0.1431	-
Observations	85	-	121	-

** p <0.1, ** p <0.05, *** p <0.01. Numbers in parentheses are corresponding t-statistics*

9. Discussion

The following discussion presents an analysis of the strengths and limitations of the present study, shedding light on its contributions to the field while acknowledging areas that require additional investigation.

The strengths lie in the robustness of the research design. In order to come up with effective recommendations, the scenarios were designed to reflect current realities and issues in the existing management systems and were aligned with the Tunisian government's future reform strategy and goals for the long-term development of the water sector in rural Tunisia. The per-household costs of a hydraulic project (either rehabilitating an existing GDA network or setting up a SONEDE network) were not arbitrarily assigned or theoretical. They were based on an understanding of the Tunisian rural context and estimates provided by the Ministry of Agriculture, Water Resources and Fisheries. Furthermore, the survey methodology was carefully designed to minimize potential sources of bias that could influence the results. We addressed hypothetical, strategic, question order, starting-point, and metric biases by leveraging respondents' familiarity with the service, clearly communicating about the study, providing detailed scenario descriptions, carefully selecting the elicitation format, and simplifying complex metrics into more understandable terms. We believe that our approach to addressing these biases was effective.

Moreover, our rigorous approach to identifying and excluding protest responses and non-responses not only yielded more accurate and precise mean willingness-to-pay results for each scenario but also offered valuable insights into respondents' preferences regarding the service provider or management model. Nearly all protest responses pertained specifically to the management model, indicating that respondents' decisions are influenced by factors like: 1) intrinsic factors, such as distrust in the management community of a GDA when declining to contribute to rehabilitating an existing network, or strong trust in the local management system when rejecting potential intervention by the national operator; and 2) perceptions about water prices.

While actual income data, which require the use

of multiple indices and a great knowledge from households of the real value of household income, would have strengthened the findings, the data of minimum monthly income required to meet basic needs proved useful as it provided readily usable financial information. Using minimum monthly household income as an indicator rather than actual income, though less precise, yielded significant results in the regression analyses.

Additionally, the variables for the regression analysis were carefully selected. For instance, the variable "Frequency of purchasing water from distributing vendors" that explains an informal activity consisting of selling raw water door-to-door using pickups, initially appeared to be a promising predictor of WTP. A positive relationship was expected between reporting the purchase of water from distributing vendors and higher WTP for improvements. However, further research revealed that distributing vendors served some areas but not others. The decision of a household to buy or not buy water from distributing vendors is influenced by both preferences and the existence of informal activities within the area where the surveyed household is located. The variable was then removed from the analysis. Using an alternative water source once throughout the entire summer, after a long interruption, for example, does not reflect the same level of water insecurity as relying on it regularly. To define regular usage, the criterion was using an alternative source at least once a week during the summer of 2022. Thus, households using an alternative source at least once a week were deemed to rely on it regularly, indicating a higher degree of water insecurity.

The results could guide and inform the government's strategy by providing insight into households' potential to make financial contributions for either rehabilitating an existing GDA network or setting up a SONEDE network. This could help make projects less dependent on conventional funding channels and, thus, rely less on government budgets and external donors.

One limitation of our study was that the two household-cost-sharing scenarios designed to assess the willingness of households to contribute financially towards a hydraulic project and determine their willingness to pay for the resulting service were weakened by: 1) the difficulty in accurately determining the real costs per cubic meter for each locality before the projects are actually completed, due to the multiple-location specific variables that impact those costs, including: the current costs of water per cubic meter, hydrological factors like reliance on groundwater and associated energy costs, and geographical factors like spatial distribution; 2) the fact that the future tariff rates applied by the SONEDE are centrally predefined and applied nationally and the difficulties for households to attribute a value for the future service due to their familiarity with simpler payment modes and understand the relationship between the charges and the volume of water consumed.

We intended to divide the group of households that received both the first and second scenarios into two equal groups. The purpose was to examine the sequencing effect and its resulting anchoring effect by altering the order of the scenarios. However, we were unable to conduct this test as the sample size we had planned initially was insufficient for this purpose.

We should have considered the current situations of the GDAs as an important factor in determining the households' WTP, for example, households connected to a well-managed GDA network are likely to be less willing to pay for the second project than households connected to a GDA presenting anomalies or households living close to an area connected to the SONEDE network are likely willing to pay more than those connected to a GDA network that is completely separate from the SONEDE network. When developing our study, we should have accounted for additional factors beyond what was included. Specifically, considering the concept of social capital, as defined and discussed by Polyzou *et al*. (2011), could have strengthened our analysis. Considering these parameters has been shown to help researchers improve the hypothetical scenarios they present and further understand the reasons leading to citizens' monetary valuations (Polyzou *et al*., 2011).

In abandoning the GDA management system, the government should consider several points, including the current situations of the GDAs (e.g., the GDA being connected to the SONEDE

or relying on groundwater), the feasibility of a project in a given area that depends on several factors like the proximity to the SONEDE network and the costs of implementing a network that depend on geography, and the (in) availability and (in) reliability of groundwater.

10. Conclusion

Tunisia is currently facing several challenges related to water, including scarcity, overexploitation, and an increase in demand, which are exacerbated by inadequate regulatory measures. These difficulties are projected to worsen because of climate change and deteriorating infrastructure, making it difficult to achieve universal access to water by 2030. Despite this, little attention has been paid to exploring households' willingness to financially contribute towards the rehabilitation of an existing water supply system or the implementation of a new water supply project of households connected to locally managed networks, which could help to improve infrastructure conditions and, in turn, increase access to water. These two aspects were studied using a contingent valuation study, which is commonly employed in developing countries to assess the value of water services and has proven to be a valuable source of information regarding the benefits associated with improved access to potable water (Van Houtven *et al*., 2017; Whittington, 2010). While the main objective was to study the willingness to contribute to the realization of a hydraulic project and its determinants, the reasons for unwillingness to contribute emerged as a valuable outcome, warranting further investigation. The presence of non-protest responses is likely due to: 1) the fact that developing water systems has historically been, and remains, the responsibility of governmental authorities. Respondents expressed unwillingness to financially contribute to these projects, viewing the provision and maintenance of such infrastructure as the government's role rather than that of local residents; and 2) a preference for or against a particular service provider. The results showed that households are nearly willing to contribute to 22% of the per-household costs of rehabilitation of an existing GDA network and 16% of the per-household costs of setting up a SONEDE network. These average amounts appear to be lower than the estimated per-household costs typically presented to funders for a general financial appraisal of a project at the conceptual stage. However, it is important to note that these estimations often overlook the unique aspects of each project. In some cases, these percentages of contributions may be sufficient to fully cover the real per-household costs once all relevant factors specific to a given project are properly considered. In the context of reforms aiming, among other goals, to develop a secure and efficient potable water supply system that meets international standards and to ensure that rural water services are equivalent to those in urban areas, the exclusive responsibility of governmental authorities for infrastructure works should be revised, giving serious thought to the contribution of households.

References

Adams E.A., Vásquez W.F., 2019. Do the urban poor want household taps? Community preferences and willingness to pay for household taps in Accra, Ghana. *Journal of Environmental Management*, 247: 570-579. https://doi.org/10.1016/j.jenvman.2019.06.113.

Ahlheim M., Schneider F., 2013. Considering household size in contingent valuation studies. *Environmental Economics*, 4(1).

Basani M., Isham J., Reilly B., 2008. The determinants of water connection and water consumption: Empirical evidence from a Cambodian household survey. *World development*, 36(5): 953-968.

Boyle K.J., Bishop R.C., Welsh M.P., 1985. Board of Regents of the University of Wisconsin System Starting Point Bias in Contingent Valuation Bidding Games. *Source: Land Economics* 61(2). http://www.jstor.orgURL:http://www.jstor.org/stable/3145811.

Carson R.T., 2012. Contingent valuation: A practical alternative when prices aren't available. *Journal of Economic Perspectives*, 26(4): 27-42. https://doi.org/10.1257/jep.26.4.27.

Carson R.T., Flores N.E., Meade N.F., 2001. Contingent Valuation: Controversies and Evidence. *Environmental and Resource Economics*, 19.

Czajkowski M., Budziński W., Campbell D., Giergiczny M., Hanley N., 2017. Spatial Heterogeneity

of Willingness to Pay for Forest Management. *Environmental and Resource Economics*, 68(3): 705-727. https://doi.org/10.1007/s10640-016-0044-0.

Favre M., 2021. Are Households Willing to Finance the Cost of Individual Water Supply? Case Study in Central Tunisia. *Water Economics and Policy*, 7(4). https://doi.org/10.1142/S2382624X21500168

Favre M., Montginoul M., 2018. Water pricing in Tunisia: can an original rate structure achieve multiple objectives? *Utilities Policy*, 55: 209-223. https://doi.org/10.1016/j.jup.2018.06.004.

Howard G., Bartram J., 2003. *Domestic Water Quantity, Service, Level and Health*. World Health Organization.

Jorgensen B.S., Syme G.J., Bishop B.J., Nancarrow B.E., 1999. Protest Responses in Contingent Valuation. *Environmental and Resource Economics*, 14.

Lopes A.F., Kipperberg G., 2020. Diagnosing Insensitivity to Scope in Contingent Valuation. *Environmental and Resource Economics*, 77(1): 191-216. https://doi.org/10.1007/s10640-020-00470-9.

Ministère de l'Agriculture, des Ressources Hydrauliques et de la Pêche. Bureau de la Planification et des Équilibres Hydrauliques, 2021. *Rapport national du secteur de l'eau Année 2021*.

Mitchell R.C., Carson R.T., 1989. *Using surveys to value public goods: the contingent valuation method*. Washington, DC: Resource for the Future.

Polyzou E., Jones N., Evangelinos K.I., Halvadakis C.P., 2011. Willingness to pay for drinking water quality improvement and the influence of social capital. *Journal of Socio-Economics*, 40(1): 74-80. https://doi.org/10.1016/j.socec.2010.06.010.

Scardigno A., Capone R., El Bilali H., Cardone G., 2017. Water-food security nexus in Middle East and North Africa Region: an exploratory assessment. *New Medit*, 16(4): 31-38. https://www.researchgate.net/publication/322007516.

Van Houtven G.L., Pattanayak S.K., Usmani F., Yang J.C., 2017. What are Households Willing to Pay for Improved Water Access? Results from a Meta-Analysis. *Ecological Economics*, 136: 126-135. https://doi.org/10.1016/j.ecolecon.2017.01.023.

Vásquez W.F., 2014. Willingness to pay and willingness to work for improvements of municipal and community-managed water services. *Water Resources Research*, 50(10): 8002-8014. https://doi.org/10.1002/2014WR015913.

Vásquez W.F., Espaillat R., 2016. Willingness to pay for reliable supplies of safe drinking water in Guatemala: A referendum contingent valuation study. *Urban Water Journal*, 13(3): 284-292. https://doi.org/10.1080/1573062X.2014.991741.

Vásquez W.F., Mozumder P., Hernández-Arce J., Berrens R.P., 2009. Willingness to pay for safe drinking water: Evidence from Parral, Mexico. *Journal of Environmental Management*, 90(11): 3391-3400. https://doi.org/10.1016/j.jenvman.2009.05.009.

Venkatachalam L., 2004. The contingent valuation method: A review. *Environmental Impact Assessment Review*, 24(1): 89-124. https://doi.org/10.1016/S0195-9255 (03) 00138-0.

Whittington D., 1998. Administering Contingent Valuation Surveys in Developing Countries. *World Development*, 26(1).

Whittington D., 2010. What have we learned from 20 years of stated preference research in less-developed countries? *Annual Review of Resource Economics*, 2: 209-236. https://doi.org/10.1146/annurev.resource.012809.103908

Whittington D., Briscoe J., Mu X., Barron W., 1990. Estimating the willingness to pay for water services in developing countries: a case study of the contingent valuation in Southern Haiti. *Economic Development and Cultural Change*, 38: 293-312.

World Health Organization & United Nations Children's Fund, 2023. *Progress on household drinking water, sanitation and hygiene 2000-2022: special focus on gender*.

The determinants of US olive oil imports

A. Malek Hammami*, John Beghin**

DOI: 10.30682/nm2404g
JEL Codes: F14, Q17

Abstract

We investigate the determinants of US bilateral imports of olive oil and their dynamics from shocks in foreign supplies and changes in US olive oil demand, using an augmented sectoral gravity framework leading to equilibrium bilateral trade flows from olive oil exporters to the US market. The empirical specification uses a panel dataset at the HS-6 disaggregation level and two estimation techniques (PPML and Heckman), that account for zero trade flows, the extensive margin of trade and the potential censored distribution of exports with zero trade flows. We run Reset and HPC tests to qualify our results. On the supply side, exporters' capacity to export, multilateral trade resistance, and immigrants' networks into the US are strong determinants of the bilateral trade flows for both aggregate olive oil exports and for virgin olive oil exports, On the consumer side, US GDP, import unit value, and immigrant network effects are robust determinants of bilateral flows as well for aggregate and virgin olive oil trade flows. Migrants' stock, exporters' GDP and population, and total exports revenues increase the probability of an exporter entering the US market. Beyond the immigrant network effects, we could not find robust evidence of consumer behavior being influenced by popular press measures of the emergence of Mediterranean diet and olive oil, or measures of cultural globalization of US consumers.

***Keywords**: Olive oil, Trade, Gravity equation, Migrant network.*

1. Introduction

Olive oil exports to the US have been increasing considerably for several decades. Trade flows have quadrupled in the last three decades (see Table 1). Numerous factors might explain this strong growth of olive oil trade between the US and the rest of the world, from factors influencing import demand and export supply of olive oil. On the demand side, beyond demographic changes, income growth, and price effects, olive oil is known for its health benefits, and it is part of the increasingly popular Mediterranean diet. The presence of large immigrant populations of Mediterranean origin in the US may have popularized the use of olive oil in that cuisine. The composition of olive oil imports has also evolved over time, towards higher quality imports of virgin olive oil and away from pomace, and more diversified sources. New export suppliers have emerged and entered the growing US market. Network effects may have helped establishing olive oil business networks through Mediterranean migrants as it has happened in other industries (Combes *et al*., 2005; Rauch, 1999). The extensive margin of trade (new exporters) is another interesting aspect to investigate to explain the rapid evolution of these US imports of olive oil.

* University of Florida, Food Resource Economics Department, Tropical Research and Education Center, Gainesville, Florida, USA.
** University of Nebraska-Lincoln, Agricultural Economics Department, C. Yeutter Institute of International Trade and Finance, Lincoln, USA.
Corresponding author: ahammami@ufl.edu

Despite its economic importance (nearly $1.3 billion in 2019), international trade of olive oil between the US and the rest of the world has received limited attention. Xiong *et al*. (2014) estimated US demand for olive oil including the role of popular diet, distinguishing three olive oil types. Ronen (2017) investigated global aggregate olive oil trade looking at the impact of nontariff measures, using a gravity-like framework. Hammami and Beghin (2021) analyze the impact of recent US retaliatory tariffs imposed on olive oil imports sourced in Spain. Our study contributes to the existing literature on olive oil trade, focusing on the US, the largest importer of olive oil, with imports predominantly sourced from the Mediterranean basin. It examines the threefold increase in import volumes from the early 1990s to 2019. There patterns precede important recent shocks but appear to have persisted despite of them. Recent shocks include the US retaliatory tariff imposed on individually packaged Spanish olive oil imports in 2019-2021, trade disruptions caused by COVID-19, and recent geopolitical instabilities.

We investigate US imports of olive oil considering demand and supply determinants and elements of extensive margins using an augmented gravity-equation equilibrium framework. The framework incorporates usual demand determinants (prices, demographics, and income), the evolving sophistication of US diet, bilateral and multilateral trade costs, and supply elements explaining the intensive and extensive margins of trade from various sourcing countries, into an equilibrium framework at the sectoral level (Yotov *et al*., 2016). The framework leads to an empirical specification applied to a panel dataset at a disaggregated HS-4 and 6 levels for olive oil products, and with two estimation techniques, which account for the large number of zeros, the extensive margin of trade, and the potentially censored distribution of bilateral trade flows.

On the supply side, we find that exporters' capacity to export, multilateral trade resistance, and immigrants' networks from olive-oil exporting countries into the US are strong determinants of the bilateral trade flows for both aggregate olive oil exports and for virgin olive oil exports. On the demand side, we find that US aggregate income, the import unit value, and immigrant network effects from olive-oil exporting countries are robust determinants of bilateral flows as well for aggregate and virgin olive oil trade flows. Regarding the extensive margin of trade, migrants' stock, exporters' GDP and population, and total exports revenues increase the probability of an exporter entering the US market. Beyond the important result on migrant networks, we did not find robust evidence of systematic influences on US consumer behavior by variables proxy-ing for the popularity of Mediterranean diet, or increasing popularity of olive oil, or measuring cultural globalization of US consumers.

The following sections provide some background information on the olive oil sector in the US, and then describe the key elements of the conceptual equilibrium framework of the gravity equation with the relevant specifics of the investigation. Estimation techniques and data description follow. Findings are presented in the last section before conclusions. A Review Appendix available from the authors provides further figures, econometric runs and tests.

2. Background on the US olive oil market and olive oil exporters

2.1. *The evolution of the US olive oil market*

Table 1 shows US olive oil production supply and disappearance and documents the phenomenal growth of the market. Olive oil consumption has quadrupled since 1990. As a result, the US has become the world largest importer of olive oil (roughly 10% of world production). More than 90% of its domestic consumption is imported.

Cultural elements may have influenced the consumption of olive oil by US consumers. First, interest in and knowledge about the benefits from olive oil, Mediterranean diet, and healthy diet have been continuously increasing among Americans (Pubmed.gov, 2020). Cultural globalization may also have facilitated the move away from the Anglo-Saxon diet to a more Mediterranean one. Deeper influences may have come through cultural network effects with rising populations of immigrants from Mediterra-

Table 1 - US olive oil Production, Supply and Disappearance in 1000 tons.

Attribute	*1990/ 1991*	*1995/ 1996*	*2000/ 2001*	*2005/ 2006*	*2010/ 2011*	*2015/ 2016*	*2016/ 2017*	*2017/ 2018*	*2018/ 2019*	*2019/ 2020*
Production	1	1	0	2	5	14	15	16	16	16
Imports	100	114	212	242	290	330	316	322	355	390
Total Supply	101	115	212	244	295	344	331	338	371	406
Exports	4	11	4	9	4	8	13	12	7	6
Consumption	97	104	208	235	291	336	318	326	364	400
Distribution	101	115	212	244	295	344	331	338	371	406

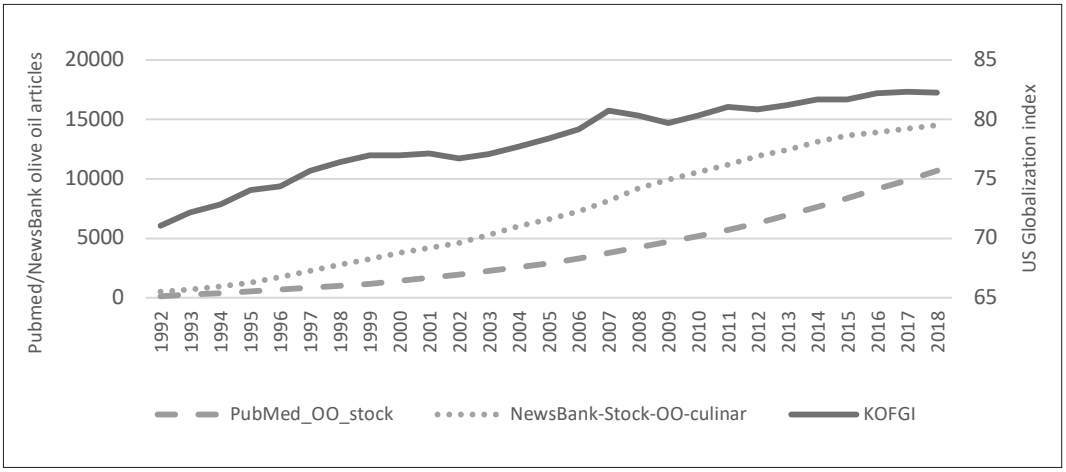
Source: USDA PS&D, 2020.

nean countries in which olive oil is paramount in the diet. For example, the increasing emigration out of Spain towards the US has been documented (Bermudez and Brey, 2017). These populations can both influence US consumers' preferences and facilitate business links back home to export to the US. We hypothesize that migrant networks may have had influenced the adoption, level of consumption, and availability of olive oil in the US and its sourcing.

To capture the growing stock of health knowledge on olive oil, we rely on Pubmed.gov to compute an index of the number of published academic refereed articles in medical journals mentioning key search terms (olive oil). This index allows for a longer and less biased series than those based on internet data[1]. In Figure 1, olive oil-related indices in the US exhibit growth patterns, as the number of medical and news articles mentioning olive oil, as well as the globalization index in the US, consistently increase since 1990.

We conjecture that the popularity of olive oil could come from cultural influence of migrants from olive-oil producing countries. Figure 2 shows the stock of migrants from olive-oil exporting countries, along with the imports from the same countries. The stock of migrants suggests a strong correlation with the increasing olive oil imports. The bilateral nature of the

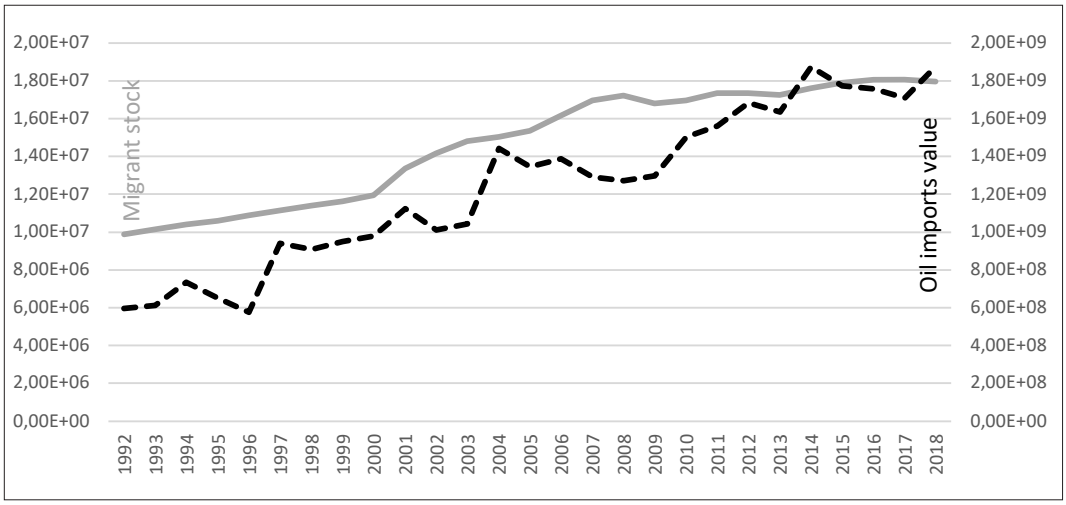
Figure 1 - Evolution of olive oil-related health, culinary and cultural indices.



Source: Pubmed.gov, 2020; kof.ethz.ch, 2020; Gygli et al., 2019; NewsBank, 2020.

[1] In the econometric estimation, we also use a related index reflecting the stock of popular press articles on Mediterranean diet using https://www.newsbank.com/ as in Xiong *et al.*

Figure 2 - Migrant stock from olive oil exporters & olive oil imports in the US.



Source: US Trade Census, 2020; Office of Immigration Statistics' (OIS), 2020.

migrant panel data provides more variation than the number-of-articles variable which only varies over time.

2.2. *Patterns of US imports of olive oil*

The olive oil market in the US can be differentiated into two main categories: virgin and non-virgin oils. Virgin oil is a higher quality product, since during this process olives have been simply pressed with no heat or chemicals involved. Virgin oils have two sub-categories defined by free acidity: extra virgin defined by free acidity of less than 0.8% and virgin olive oil with free acidity between 0.8% and 2%. These oils are pure and not refined.

All other olive oils, heat or chemically treated, are considered non-virgin and can be sometimes mixed with some virgin oil and simply called olive oil with free acidity above 2%. Olive oil extracted by chemical process is called pomace and is the lowest quality product. There are other subcategories of virgin olive oil (first cold pressed, cold pressed, and organic) (IOC, 2020).

The "Olive Oil & Its Fractions" category imported to the USA under the HS code 1509, has the largest average share of consumption of more than 90% of all olive oils imported and consumed in the US. The remaining share is the "Olive-residue Oil & Blends" category under HS code 1510 (edible and non-edible). The "Olive Oil & Its Fractions" category (HS code 1509) divides into "Virgin olive oil/fractions" category (HS code 150910 and "Refined olive oil / fractions" (HS code 150990) category[2]. The virgin olive oil (HS 150910) includes extra-virgin, labelled and organic of a superior quality than the refined one. Since the early 1990's, the share of virgin olive oil has been increasing from 35% to reach 80% of olive oil imports. This increase reflects both the rising consumption of olive oil and the progression towards higher-quality olive oil consumed in the US. The global economic crisis of 2008-09 temporally reset the clock on this evolution as shown in Figures 3.a and 3.b; trends are clear.

Figure 4 shows the evolution through time of US imports for virgin olive oil by import source. EU sources dominate (Spain and Italy), The

[2] HS 1509: Olive oil and its fractions, whether or not refined, but not chemically modified; HS 150910: Olive Oil and Its Fractions, Virgin, Not Chemically Modified; HS 150990: Olive Oil and Its Fractions, Refined But Not Chemically Modified. HS 1510: Other Oils and their fractions, obtained solely from olives, whether or not refined, but not chemically modified, including blends of these oils or fractions with oils or fractions of heading 1509.

Figure 3.a - Evolution of import shares: Olive oil and Residual (1992-2019).

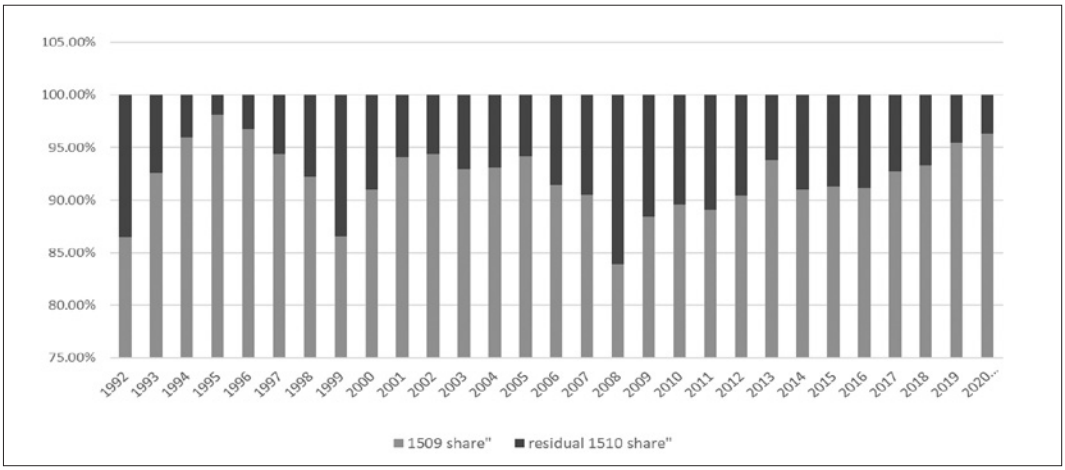
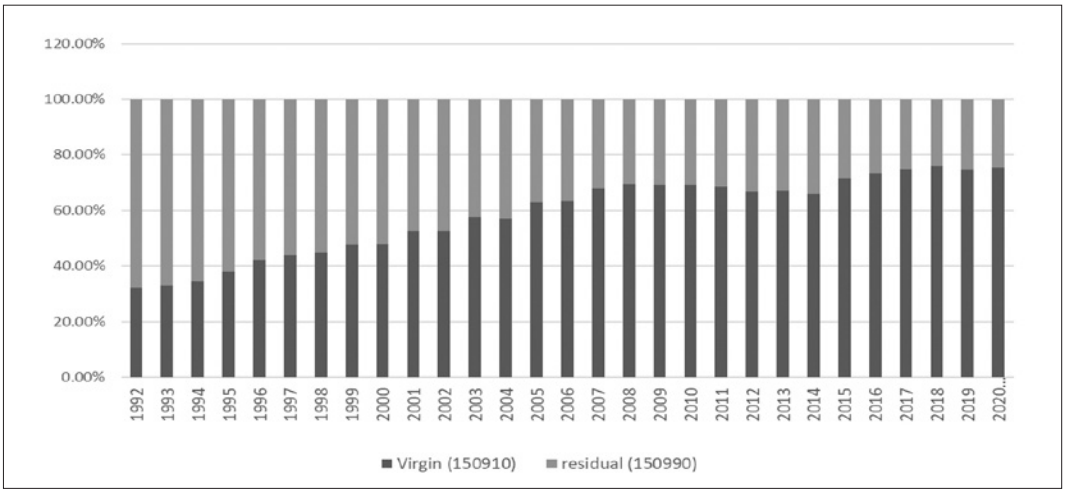


Figure 3.b - Evolution of import shares: Virgin and Refined (1992-2020).



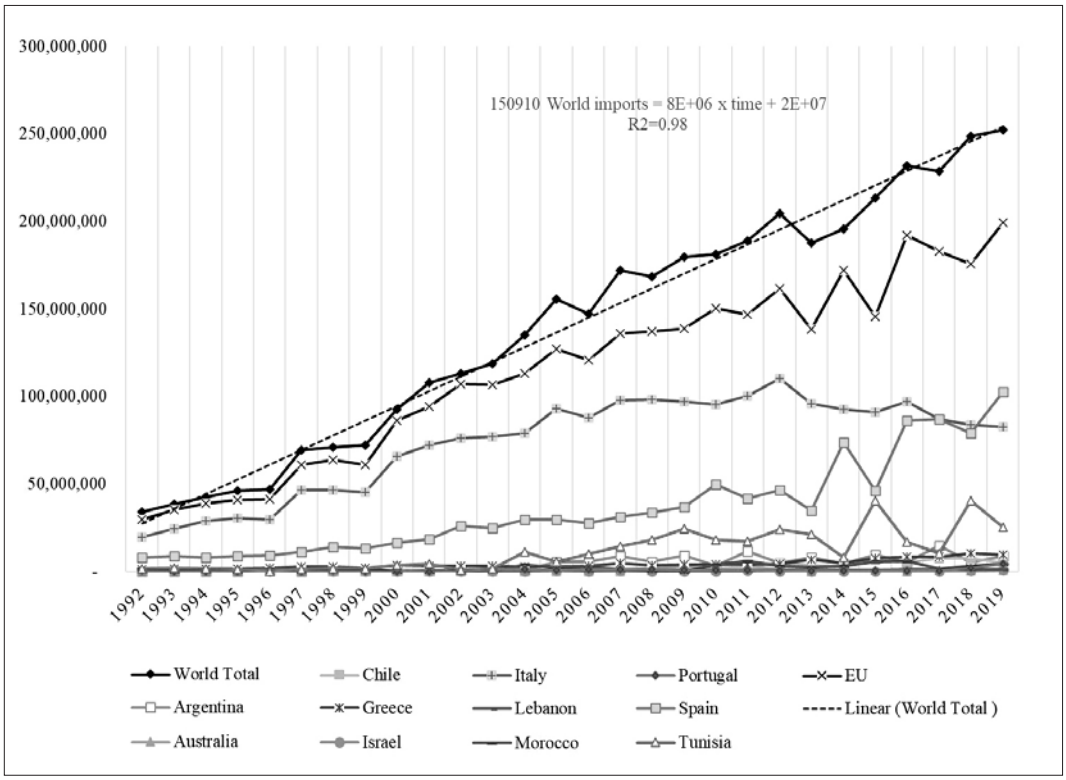
Source: US Trade Census, 2020.

importance of non-EU Mediterranean sources (Tunisia, Morocco) is rising, and a competitive fringe exists with Argentina, Israel, Lebanon, Chile, and Australia.

Most of the virgin olive oil imported to the US is from Italy and Spain. Spain overtook Italy in 2018. The rest of the countries exporting to the US have remained small exporters but with growing quantities exported. They provide a competitive fringe to the established exporters. Tunisia has increased its exports to the US the most, since 2004, approaching Spain exports in 2015, and remains the largest exporter within the fringe.

Argentina has had a noticeable increase since 2005 and is now the 4th largest virgin olive oil exporter to the US. Olive production for olive oil exhibits stochastic yields with "good and bad" years resulting in annual variations of production even for established exporters. Inventories partially mitigate these variations. Variations in export supply are also reflected by significant variations in import unit values. On average, import unit values have been rising slightly above inflation rates. Import unit values vary among exporters, indicating differences in quality and competitiveness. Italian, Israeli, and Argentini-

Figure 4 - Evolution of US virgin olive oil (150910) imports by country of origin (1992-2019).



Source: US Trade Census, 2020.

an unit values have increased more rapidly than others (see Review Appendix Figure 1.A).

US imports of refined olive oil have been on a decreasing trend for most exporters (not shown) as consumers upgraded to virgin olive oil. Italy and Spain are major exporters of refined oil, ahead of the others (fringe). The fringe of other exporters still has Tunisia as the third largest source competing with Turkey and Morocco. The unit value of refined olive oil is also increasing. However, the dispersion across sources is smaller compared to the normalized unit values of virgin olive oil (see Review Appendix Figure 2.A).

2.3. *Supply shocks in World olive oil markets*

Various supply shocks and changes interact with US demand of olive oil. Producing countries, endowed with a specific Mediterranean climate, compete to supply the world market, including the US. Profit-maximizing firms in these countries' supply chains compete and adapt to changing market conditions. New entries and production techniques have put pressure on average unit values. Spain, as an example, invested hugely on reforming olive oil production and opted for an intensive production since the 1960's. Nowadays, Spain supplies almost half of world production (46%) (Guerrero, 2014).

Italy and Greece relied both on their historical reputations and their authentic ancient know-how to signal their quality. Italy imports large amounts of Spanish oil, which allegedly find their way back on the world market. However, there has been a history of olive oil fraud concurrently, wherein false labels—whether related to quality or product source—are employed to generate increased profits (Da Silveira *et al*., 2017; Bimbo *et al*., 2019; Yan *et al*., 2020; Casadei *et al*., 2021). With the development of digital marketing strategies and globalization, olive

oil producers are going up market and non-traditional producers are entering the international market. The resulting glut in the world supply of olive oil pushes producers to differentiate their product for a higher quality (IOC, 2020; Milli, 2006; USDA PS&D, 2020). The increasing and now dominant share of virgin olive oil in US imports reflects this fact.

Finally, olive oil production has a stochastic yield due to environmental and agronomic shocks. Weather, pathology, and physiological state of the trees impact its yearly production. Olive trees are biennial trees that have alternate yearly production. One year above the average production and one year below. We later investigate this potential variability of yield, although unit values of traded olive oil reflect that variability to a great extent.

2.4. *Evidence on the evolution and sophistication of consumer demand*

Consumers around the world, including in the US, have been increasingly concerned about the quality of food they purchase (IFIC, 2018). Many studies have investigated the relation between health and nutrition information and food demand. Early research about health and nutrition factors has found evidence of diversion of US demand from food containing cholesterol and heavy fats (Brown & Schrader, 1990; and Chern *et al*., 1995). Other studies have approached health information and demand from the experimental and behavioral perspective (e.g., Hilger *et al*., 2011).

Diet trends have been emerging with various news coverage and success, such as keto, vegan, and Mediterranean diet, among others. The Mediterranean diet has spread in the industrialized world from its origin in the Mediterranean basin. Alexandratos (2006) and Regmi *et al*. (2004) investigated this rise of the Mediterranean diet and related it to globalization and income growth. Trends of global and US food consumption determinants are examined. Regmi *et al*. (2004) analyzed trade data to determine changing diets phenomenon's effect on Mediterranean diet products' trade.

Studies of olive oil demand in North America have been scarce (Del Giudice *et al*., 2015). Xiong *et al*. (2014) estimated US demand for olive-oil differentiated products using the AIDS model and accounting for the impact of information on Mediterranean diet. They find that both the stock and number of press articles discussing olive oil and health are positively related to the level of olive oil imports. They aggregate all olive oils imports into three aggregate categories. Menapace *et al*. (2011) study olive demand in Canada through a survey that demonstrated the significance of geographical indication and certification of origin Label.

Main studies on olive oil focus on Europe. Many are surveys and experiments (Karipidis *et al*., 2005; Kalogeras *et al*., 2009; Bernabéu & M. Díaz, 2016; Cacchiarelli *et al*., 2016; Carbone *et al*., 2018; and Scarpa & Del Giudice, 2004). Relevant to our analysis, in the context of a net-importing country, Kavallari *et al*. (2011) investigate the structure of import demand of Germany and UK for olive oil from southern European producers. That study itself has been based on Vlontzos and Duquenne (2008) on Greek olive oil potential in the international market. Finally, Garcia Álvarez-Coque and Martí Selva (2006) use a gravity model to estimate euro-Mediterranean fruits and vegetable trade flows.

3. Model

The gravity equation approach to bilateral trade is widely popular among trade economists, despite its drawbacks being overly structured with symmetric trade costs, fixed endowment approach to supply and its normalization of price limited to cross-section data (Yotov *et al*., 2016; Baldwin and Taglioni, 2006). These assumptions are relaxed here.

Methods of estimation have evolved greatly to address the presence of zero bilateral trade flow, the extensive margin to trade, and typical estimation mistakes biasing results. Baldwin and Talglioni (2006) identified three principal mistakes of gravity model's applications in the literature, when multilateral trade resistance terms are omitted and are correlated with error terms, leading to bias; when bilateral trade flow volumes are wrongly deflated by a common de-

flator, and when computing the wrong log-averaging of the bilateral trade flow volumes. The authors proposed several time-varying and unvarying dummies as an attempt to adjust gravity regression mistakes.

More complex issues with gravity models such as the extensive margin of trade have been addressed as well. Several authors such as (Melitz, 2003) have elaborated the extensive margin of trade, beyond a simple Heckman sample selection process explaining the decision to trade or not between two countries. Firms are heterogeneous and the most productive firms enter new markets at the extensive margin. Gould (1994), Rauch (1999), and Combes *et al*. (2005) focus on network effects on trade for differentiated products. Our approach relies on a simple sample-selection approach to the extensive margin and incorporates network effects in both margins.

3.1. *The sectoral gravity model*

We consider only a single destination *j* (*j*=*us*) as in Kavallari *et al*. (2010). Products are differentiated by country of origin. The demand in the destination country (here the US indicated by the subscript us), is obtained from maximizing a CES-utility function, with utility derived from consuming products differentiated by origin (all exporters of olive oil). The setup extends to a sectoral approach from which we abstract here to simplify the presentation. The

$$\left\{\sum_i \; a_i^{\frac{1-\sigma}{\sigma}} c_{ius}^{\frac{\sigma-1}{\sigma}}\right\}^{\frac{\sigma}{\sigma-1}},$$

subject to the following budget constraint for a set expenditure *E,*

$$\sum_i p_{ius} c_{ius} = E_{us},$$

where, $\sigma>1$ is the elasticity of substitution between goods. The exogenous taste parameter a_i is the CES preference parameter, which will be instrumental later to incorporate the impact of diet information.

The consumption of varieties from country *i* is given by c_{ius}. Total expenditure (E_{us}) measured at delivered prices ($p_{ius} = pi\ t_{ius}$) defined as a function of factory-gate prices in country of origin (p_i) and bilateral trade costs markup (1+trade cost) from *i* to ($t_{ius}>1$).

Utility maximization under a budget constraint leads to the demand for variety *i*:

$$x_{ius} = \left(\frac{a_i p_i t_{ius}}{P_{us}}\right)^{(1-\sigma)} E_{us}, \qquad (1)$$

where x_{ius} is the trade flow in value from origin *i* to the US, and P_{us} denotes a CES consumer price index or "inward multilateral resistance", capturing the ease of market access into the US:

$$P_{us} = [\sum_i \; (a_i p_i t_{ius})^{1-\sigma}]^{\frac{1}{1-\sigma}}. \qquad (2)$$

Next, a market clearing condition for each exported good equates production to the sum of demands in all destination markets, including domestic demand. The equilibrium condition leads to exporters' multi-resistance terms. The equilibrium condition assumes that the shipped quantities "melt" on the way to their destinations *j* by an amount equivalent to the trade cost:

$$Y_i = p_i Q_i = \sum_j \; x_{ij} = \sum_j \; \left(\frac{a_i p_i t_{ij}}{P_j}\right)^{1-\sigma} E_j. \quad (3)$$

It is equal to factory-gate prices p_i multiplied by Q_i supply of the given product in exporter *i* before melting of the iceberg occurs (representing trade cost). Bilateral trade cost is an iceberg cost. For US consumers to receive x_{ius}, exporters have to send $x_{ius}\, t_{ius}$ with $t_{ius} > 1$. Buyers bear that extra-cost of x_{ius} (t_{ius}-1). Equation (3) presents the determinants of the equilibrium trade flow between countries *i* and *j*. The term

$$(a_i p_i)^{1-\sigma} = \frac{Y_i}{\Pi_i},$$

is used to eliminate the factory price in the equilibrium condition. The price index Π_i is given by:

$$\Pi_i \equiv \sum_j \; \left(t_{ij}/P_j\right)^{1-\sigma} E_j, \qquad (4)$$

which is the "outward multilateral resistance" that shows exporter *i*'s ease of market access into all *j* countries (Baldwin and Taglioni, 2006). Substituting (3) and (4) in (1) yields:

$$x_{ius} = \frac{Y_i E_{us}}{\Pi_i} \left(\frac{t_{ius}}{P_{us}}\right)^{1-\sigma}, \qquad (5)$$

where $Y_i E_{jus}$ illustrates the "size term," and

$$\frac{1}{\Pi_i} \left(\frac{t_{ius}}{P_{us}}\right)^{1-\sigma}$$

is the "trade cost term." In our sectoral applica-

tion, we assume that Y_i represents the capacity to export olive oil by country i. Then for any exporter i the ratio

$$\frac{1}{\Pi_i}\left(\frac{t_{ius}}{P_{us}}\right)^{1-\sigma}$$

represents the trade cost of the US market relative to all destination markets served by that exporter. It can be shown that this ratio varies monotonically with the ratio of the US import unit value for that exporter and the average real import unit value of that exporter to all destinations. We use this characteristic in our empirical strategy.

3.2. *Empirical strategy (estimation approach and empirical specification)*

As shown in section 2, the set of countries exporting olive oil to the US has been changing over time. Some countries entered the market several years ago, such as Australia, Brazil, Algeria, Peru, Slovenia, among others. When a competitor has not yet entered or chooses to exit the US market for a given year, its trade volume will be taking the value of zero. In general, to accommodate zeroes, many investigations use PPML, which allows to include zeroes as part of the intensive margin of trade by taking the exponent of equation (5) and the logarithm of continuous variables on the right-hand-side of the equation.

One can add an extensive margin to this which in PPML is confounded with the intensive margin and does not address potential selection into exporting to a destination market and the potential censoring in zero observations. PPML estimates are consistent under heteroskedasticity and the approach provides a natural solution to mechanically handle zero values of the dependent variables and provides a robust covariance matrix estimator (Santos Silva and Tenreyro, 2006).

We address the extensive margin of trade using Heckman sample selection, which accounts for the censoring of the latent variable representing the decision to trade or not through its first-stage Probit. We do not account for the second element of extensive margin of heterogeneous firms noted above. For completeness we also provide truncated OLS estimates in the Review Appendix, using the strictly positive observations in the dataset. We also run a Probit of the probability to export by countries over time, to gauge the fit of the selection equation in the Heckman model.

In alignment with Kavallari *et al*. (2010), our gravity equation focuses on a sole importer, specifically the US, utilizing an imbalanced panel. This choice involves a trade-off, resulting in reduced variation in the destination country (solely over time). However, it facilitates the acquisition of results specific to the US, as opposed to estimates reflecting a diverse panel of countries.

For Statistical tests, first, we run a Ramsey RESET test to check if the models exhibit evidence of misspecification. Second, we use the HPC test of Santos Silva *et al*. (2015) to choose which of PPML or Heckman fits our olive oil data the best. The HPC test discriminates between two competing models to fit data with many zeros. The test compares "two-process models" like Heckman to a simpler 1-process model accommodating zeros without explicit extensive margin like the PPML approach.

Specification and variable proxies

Since we have panel data (1992-2018) for 21 exporters, we add a time subscript to our specification. Each olive oil product k imported by the *US* at year t from exporter i, has the following trade flow equation:

$$X_{iUStk} = \frac{Y_{it}E_{US,t}}{\Pi_{ikt}}\left(\frac{t_{iUStk}}{P_{USkt}}\right)^{1-\sigma}. \qquad (6)$$

We also make use of the ratio indicated in footnote 3. Since both the US import unit value and the exporter's average import unit value are time varying, our combined variable

$$\frac{1}{\pi_{it}}\left(\frac{p_{it}\, t_{iust}}{P_{ust}}\right)^{1-\sigma}$$

is obviously time varying. For product k, it is proxy-ed by the ratio

$$\frac{UV_{i,US,t,k}}{avgUV_{i,t,k}}$$

where $avgUV_{i,t,k}$ is the average unit value of product k over all destinations served by exporter i. It plays the role of outward trade resistance measure since it includes all the t_{ij} trade costs for all j markets served by exporter i.

Further, we use the time-varying capacity to export olive oil by exporters to approximate output y_{it}. Variable $Tot\_exports_{i,t,k}$ is exporter i's total exports of olive oil k to the world in year t. It is a solid proxy for exporter i's production and capacity to export olive oil k. We choose this variable as olive production data from FAO are incomplete for several countries and do not disaggregate olive oil types. The smoothing role of inventories is implicit in total exports, without having to model it explicitly. We also address potential endogeneity of supply determinants using a series of exporter fixed effects to address some of the potential endogeneity issues coming from omitted variables. This is not implementable on the importer side since we only have a single destination in the trade flow data (j=USA).

On the demand side of equation (6), we already have the import unit value $UV_{i,t,k}$ and then we use 2010-price constant US GDP, $realGDP_{US,t}$ as a proxy for consumers' income. The latter income measure includes demographic change affecting US GDP.

Regarding trade-flow determinants related to diet adoption, cultural influences, and market penetration, we first rely on immigration networks (stock of migrants from olive oil exporting countries). Immigration and tourism are important factors having network effects that can affect food trade flows (Kavallari *et al*., 2011). Rauch and Trindade (2002) found that the share of ethnic Chinese populations as immigrants in a country affects bilateral trade of that country with China. This is consistent with earlier investigations (Gould (1994) and Rauch (1999)). Rauch (1999) used proximity variables such as distances –this was prior to the CEPII database availability on distance. We investigate this geographic proximity relying on the geodist database of CEPII. Following Gould (1994) and Combes *et al*. (2005), we look at migrant networks coming to the US. The immigration data are collected from the Office of Immigration Statistics' (OIS) yearbooks of immigration statistics 1996-2019. We use tables of Persons Obtaining Lawful Permanent Resident Status by Region and Country of Birth (Office of Immigration Statistics, 2021). In the Heckman model with the extensive margin, we further investigate the role of migrants into the extensive margin. The latter would capture business network influences, rather than their cultural influence, of migrant networks from olive-oil producing countries.

We rely on established empirical strategies to incorporate the impact of health information and information on popular diets in demand systems. Variables capturing US consumer's demand sophistication are incorporated into a CES framework through preference parameter X in equation (1). Indices reflecting the sophistication of demand are as follows. First, we use the KOF index of cultural globalization which measures the degree of globalization of 122 countries over time based on economic, social, and political criteria on a 100-score scale for each country. The index was first introduced by Dreher (2006) at the Konjunkturforschungsstelle at ETH Zurich, Switzerland from which it takes the name "KOF" (kof.ethz.ch, 2020; Gygli *et al*., 2019). We use the index value for the US. It varies over time. Next, we use a PubMed index which counts the number of published scientific refereed articles about the searched terms (olive oil, healthy diets, etc.). This index only counts the health and medical publications. We use both annual flow and accumulated stock of publications (see pubmed.gov, 2020). We also develop a popularity index reflecting the stock of news articles on Mediterranean diet and health benefits associated with olive oil based on the NewsBank database as in Xiong *et al*. (2014).

In some additional runs (see Review Appendix), we investigate adding yield to capture technical change elements in olive oil production not captured by the unit value. Yield data is incomplete and leads to a decrease in observations. We also look at the impact of regional trade agreements on olive oil trade. Tariffs on olive oil are typically low and do not vary much over time, especially in the context of a single destination. However, deeper market access may have an impact on trade flows via trade facilitation measures. We also look at lead and lag effects of RTAs on olive oil trade flows. These results did not exhibit any robustness once exporter fixed effects are introduced (see Review Appendix table A.1). The limited number of RTAs for the US may also explain the lack of significant results.

Trade flow data are based on US Trade Cen-

sus. Observations were restricted to the 21 top exporters (Algeria, Argentina, Australia, Brazil, Chile, Croatia, Cyprus, Egypt, France, Greece, Israel, Italy, Jordan, Lebanon, Morocco, Portugal, Slovenia, Spain, Tunisia, Turkey, and Peru) representing more than 99.8% (for 2019) of imports to the US. Excluded countries from the initial selection are Albania, Canada, China, Japan, Mexico, Montenegro, New Zealand, Palestine, South Africa, and Syria. These countries were excluded because several data series were missing and because of market disruptions, such as in Syria. The panel dataset is available for total exports, imported quantities and prices (import unit values). It extends from 1992 to 2018 consisting of olive oil (HS:1509) and its two major components: virgin oil (HS: 150910) and refined oil (HS:150910). Pomace or residual olive oil (HS:1510) is dropped from the estimation because of many missing observations and its limited use in food consumption. We use physical quantity trade data (variable c_{iUStk} rather than value X_{iUStk}) since we have both physical unit data and import values in US Trade Census. In sum, our panel has 26 years (1992-2018), and 21 countries, with 546 observations. For each olive oil type, there is a considerable number of zeros, which we keep, and some missing unit values which we addressed below.

We have the final empirical specification of trade flows as follows:

$$c_{iUStk} = f\left(realGDP_{US,t}, \frac{UV_{i,US,t,k}}{avgUV_{i,t,k}}, Tot_{exports_{itk}}, fixed\ effect_i, migrant\ stock_{it}, taste\ shifter_{ust}\right) \quad (7)$$

Equation (7) is run in the typical exponential form of the log of continuous variables and fixed effects for the PPML estimation. For the Heckman and truncated OLS specifications, a double log specification of (7) is run. For the Heckman specification, the inverse Mill's Ratio coming from the covariance between selection and level equations is also included as an explanatory variable. We specify the selection equation as follows:

$$L_{iUStk} = f\left(realGDP_{US,t}, \frac{UV_{i,US,t,k}}{avgUV_{i,t,k}}, Tot_{exports_{itk}}, fixed\ effect_i, migrant\ stock_{it}, taste\ shifter_{ust}, pop_{i,t}, GDP_{i,t}, Cons_{US,t}, Exprev_{i,t}, EU_{i,t}\right) \quad (8)$$

where, L_{iUStk} is the trade latent variable. It takes the following values:

$$\begin{cases} L_{iUStk} = 1 \text{ when } c_{iUStk} > 0 \\ L_{iUStk} = 0 \text{ when } c_{iUStk} = 0 \end{cases}.$$

Variable $pop_{i,t}$ is the exporter's population, $GDP_{i,t}$ is the exporter's GDP, $Exprev_{i,t}$ is the exporter's total olive oil export revenues, and $EU_{i,t}$ is a dummy that indicates whether the exporter is an EU member, all expressed at a time period *t*. We also included the US consumption of olive oil at a time *t*, $Cons_{US,t}$, to capture the growth of the market over time. The part-one Probit (or the selection equation) has the same variables as the main Heckman (part-two) equation, with additional variables explaining the decision to export or not and which are excluded from the trade level equation for proper identification. Appendix Table 1 provides a list of the variables, with their definition, units and sources and descriptive statistics.

Dealing with missing observations

For several observations, the import unit values $UV_{i,US,t,k}$ (or in a few cases $avgUV_{i,t,k}$ values were missing which would have resulted in omitting these observations from regressions. This is the case for instance for the zero-trade flow observations which do not have an observed unit value to the US market. We use two instruments to replace missing values of $UV_{i,US,t,k}$, which represent plausible expectations of potential exporters in country *i*. First, we use the average import unit value of the two geographically closest countries in the sample that are exporting to the US for that year. Distance data from CEPII allow us to compare country pairs distance-wise. For example, Cyprus has the following countries from the sample ordered by increasing distance: Lebanon, Israel, Jordan, Egypt, Turkey, and Greece. We compute its missing unit values by averaging the unit values of Lebanon and Israel, assuming they exist. Second, we assume that the missing *UV* to the US is expected to be equal to the average unit value for that country and year ($UV_{i,US,t} = avgUV_{i,t}$), which means the "expected" import unit value of *i*'s olive oil for the US equals the average unit value for its worldwide shipments for that year. The choice of instrument has virtually no impact on the results.

Table 2 - Comparative Gravity estimations PPML-PROBIT & HECKMAN.

HS code	*HS 1509*			*HS 150910*			*HS 150990*		
Variables	*PPML*	*PROBIT*	*HECKMAN*	*PPML*	*PROBIT*	*HECKMAN*	*PPML*	*PROBIT*	*HECKMAN*
ln_gdp_USA	0.828***	-1.273	1.093***	2.350***	0.231	2.299***	0.931**	-0.029	0.397
	(0.004)	(0.709)	(0.007)	(0.000)	(0.949)	(0.000)	(0.023)	(0.992)	(0.710)
ln_IUV	-1.052***	-0.037	-1.337***	-1.295***	-0.158	-1.224***	0.516*	0.039	-0.924***
	(0.000)	(0.841)	(0.000)	(0.000)	(0.546)	(0.000)	(0.052)	(0.825)	(0.000)
ln_Exporter's total_Exports	0.826***	-0.512	0.824***	0.722***	0.014	0.682***	0.379**	0.237	0.414***
	(0.000)	(0.104)	(0.000)	(0.000)	(0.963)	(0.000)	(0.042)	(0.238)	(0.000)
ln_Stock_mig	1.145***	0.761***	1.029***	0.973***	0.685***	0.732***	1.341***	0.325***	1.764***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.008)	(0.000)	(0.003)	(0.001)
ln_ KOFCuGIdf	0.382	-2.492	-0.113	-0.308	0.195	-1.116	-1.560***	-0.51	0.291
	(0.304)	(0.269)	(0.880)	(0.221)	(0.926)	(0.227)	(0.003)	(0.753)	(0.883)
ln_pop_ Exporter		-0.92***			-0.97***			-0.631**	
		(0.002)			(0.004)			(0.012)	
ln_gdp_ Exporter		0.736**			1.090**			0.748***	
		(0.033)			(0.011)			(0.003)	
ln_US_ consumption		1.323			0.545			0.655	
		(0.403)			(0.737)			(0.629)	
ln_tot_exp_ rev		0.900***			0.261			-0.049	
		(0.006)			(0.395)			(0.808)	
member_eu_ Exporter		0.020			0.023			0.244	
		(0.972)			(0.966)			(0.460)	
Exporters' Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mill's ratio			0.767***			1.008***			1.930**
			(0.004)			(0.000)			(0.022)
_cons	-40.69***	12.907	-45.84***	-80.26***	-45.724	-72.43***	-33.81***	-23.25	-30.92
	(0.000)	(0.887)	(0.004)	(0.000)	(0.637)	(0.000)	(0.005)	(0.768)	(0.258)
N	542	534	534	515	512	512	484	456	480
R-sq (pseudo)	0.981	0.610	-	0.984	0.632	-	0.956	0.498	-

*P-values in parenthesis. *p=<0.1; **p<0.05; ***p<0.01. Country specific fixed effects of this regression were deleted from this table for simplicity reasons*

4. Results

Table 2 presents the estimation results for the chosen estimations methods (PPML and Heckman sample selection (with its Probit selection equation)). Truncated OLS was also estimated (see Review Appendix Table A.3). All specifications include exporter fixed effects, which are not reported to save space. Country fixed effects are significant, except for France, Croatia, Jordan, and Portugal. Peru is omitted and in the intercept. Results are available upon request. The three oil categories are run as separate regres-

sions. In separate runs with the disaggregated categories (HS 150910 and 150990) we included cross-price effects which were not significant.

All models explain well the variation of the import level (R^2 for olive oil HS1509 and its subcategories >0.9). The unit value ratio as a price proxy for trade cost and exporter trade resistance has the expected negative sign, except for the refined lower quality oils HS:150990 that has an insignificant positive sign. Similarly, the GDP as proxy for income has a systematically significant and positive sign, which confirms the aggregate income effect for olive oil products in the US. The refined olive oil HS:150990 with its lower quality has an insignificant income coefficient.

Among non-price determinants, the candidate variables were tried with mixed success. First, network effects appear to be very robust in all equations, the stock of migrants from exporting countries living in the US explain much of the variation of olive oil imports over time and across export sources. For the shift of taste parameters, we tried several combinations of the KOF index and PubMed and NewsBank indices, both in stock and flow form. However, we could not find robust results, the estimations were plagued by sign reversals, and loss of significance. In addition, the inclusion of these variables in the specification tends to dilute the significance of the income variable. Hence, we are not confident that we can capture the alleged taste changes with these indices, beyond the cultural influence captured by the migrant network variable. Detailed results are reported in the Review Appendix. Table 2 below illustrates the type of insignificant results which we obtained (here with the KOF index).

A comparative horizontal reading of Table 2 gives a cross quality comparison for oils. The higher quality virgin olive oil (HS:150910) has generally a higher price and income response than its lower quality analogue (HS:150990). Migrants' stock has almost the same effect on both subcategories. Olive oil (HS:1509) that regroups both subcategories, shows effects that are close to the average of the effects on two subcomponents at the HS-6 level.

The Heckman sample selection specification is also shown in Table 2. Its first-stage Probit explains the extensive margin for exporters who decide to export or not. The Probit explains about 60% of the variation of the dependent variable (R^2 of 0.6) in most runs. It shows that migrants' stock, exporters' GDP, exporters' population, and total exports revenue are among significant factors that influence decisions to initiate exports of olive oil to the US. This result also provides some insight on the business-network of migrant to create new trade (the extensive margin) as opposed to their influence on the intensive margin through their expanding cultural influence. Heckman's second stage (trade levels) provides results comparable to PPML and truncated OLS results (see Review Appendix Table A.3). Coefficients' signs and robustness are similar to PPML, with stronger price and income response for the higher quality virgin olive oil HS:150910. Heckman's Mills' ratio significance suggests that sample selection is present.

Table 3 presents RESET tests for the models

Table 3 - RESET test.

HS / Test results	*1509*				*150910*				*150990*			
	OLS	PPML	PROBIT	*HECK-MAN*	OLS	PPML	PROBIT	*HECK-MAN*	OLS	PPML	PROBIT	*HECK-MAN*
RESET test P-Value	0.792	0.0684	0.203	0.652	0.000	0.0017	0.0238	0.0000	0.0000	0.0000	0.0237	0.8333
Chi2/F-value	0.35	3.32	1.62	0.20	8.22	9.81	5.11	18.71	9.71	32.95	5.12	0.04
N	449	542	534	534	413	515	512	512	332	484	456	480
R-sq (pseudo)	0.937	0.981	0.61	-	0.921	0.984	0.63	-	0.874	0.956	0.498	-

Table 4 - HPC test.

HS / Test results	*1509*		*150910*		*150990*	
HPC test	PPML	HECKMAN	PPML	HECKMAN	PPML	HECKMAN
P-Value	0.000	0.133	0.000	0.181	0.000	0.535
T-value	9.280	1.114	14.694	0.913	9.663	-0.087
N	534	534	511	512	477	480
R-sq (pseudo)	0.981	-	0.984	-	0.956	-

shown in Table 2. For the aggregate olive oil (HS:1509), the RESET tests suggest that the Heckman trade volume equation is acceptable. The Reset test for PPML is borderline significant at 10% but not at 5% or less. We read it as inconclusive. For the disaggregated regressions (150910 and 150990), the RESET tests indicate evidence of misspecification for the chosen specifications. The Heckman trade volume equation fails the reset test for 150910 but passes the test for 150990. Table 4 presents the HPC tests in a comparison between PPML and Heckman sample selection models to discriminate the best model (one-part versus two-part models) for our data. The HPC test uses predicted trade levels but based on a logarithmic estimation for the Heckman and in trade levels for PPML. The Heckman estimated values are then reconverted in levels for the test, following Santos Silva *et al*. (2015). The test clearly rejects PPML, while not rejecting Heckman for both aggregate and disaggregated trade flows.

We conclude that both RESET and HPC tests suggest the Heckman sample-selection model provides the best fit for our data to address the US import demand determinants for olive oil, compared to PPML. The Inverse Mills Ratio is also significant which suggests that the error terms of the selection equation and the trade equation are correlated. The extensive margin contributes to explaining why exporting countries enter or not the US market for olive oil. The selection equation shows the positive role of migrant networks on the extensive margin of trade.

Beyond the specifications presented in Table 2, multiple variations in specifications were tried. We added yield, to capture technical change. However, several countries had missing yield data, leading to a decrease in observations. Results were not robust with sign reversal and various level of significance. Regional trade agreements (RTA) between the US and its partners were also introduced, including lag and lead effects to attempt to capture potential effects of deeper trade integration. (See Review Appendix table A.1 for an example of RTA specification). Results again were not robust. Both were removed from our preferred specification because exporters' fixed effects capture much of these two variables, except their time variation.

In addition, bilateral distance which, often used to proxy bilateral trade cost, was tried and found not to be robust. It exhibited sign reversal (significant positive sign), the inverse of what the theory suggests. Three elements contribute to this non result. The lack of variation in destinations combined with the geography of olive production is the first and second ones. Most of olive oil exporters are located in the Mediterranean basin (say, as opposed to Mexico and Canada for other commodities) and quality differs from an exporter to another with new-world exporters (Chile, Argentina) just emerging relative to further-distant established EU exporters. Last, our unit cost ratio variable incorporates the cost of shipping the various olive oils. It is a solid proxy for several trade costs including the cost of overcoming distance.

Finally, for both subcategories of olive oil (virgin and refined), we have thought of testing cross-price effects in the specifications. It appeared that cross-prices were not significant with no effect on the explanatory variables nor the R^2.

5. Discussion

This paper investigated the determinants of US import demand and export supply of olive oil in a partial equilibrium framework accounting for both demand and supply determinants. We used an augmented gravity equation equilibrium framework applied to the single destination US market. the most novel and interesting results pertain to immigrant network effects. On the demand side, we find that migrant networks exhibit a systematic positive influence on demand via a cultural element in the intensive margin of trade reflecting new preferences for olive oil. These migrant communities and networks propagate new culinary habits and food consumption, in this case, the Mediterranean diet and olive oil. Related to the supply side, we found strong network effects at the extensive margin, created by immigrants from exporting countries. These networks facilitate new trade flows, through their influence in business networks lowering the cost of entry into the US market for new exporters.

Related to conventional determinants of demand, we found that income and relative prices are important determinants of olive oil demand with high elasticities in the case of virgin olive oil. Beyond this strong result related to migrant networks, and despite using several alternative proxies to measure the influence of information and cultural influences on US olive oil demand, we could not find any additional robust association, somewhat to our surprise. This result is in contrast to Xiong *et al*. We use a much finer disaggregation of olive oil by origin relative to Xiong *et al*. who abstracted from the influence of migrant network in their context. The strong aggregation in Xiong *et al*. smooths out much of the variability we have in the bilateral flows and allows for identifying an aggregate trend, too difficult to identify in our disaggregated bilateral flows.

On the supply side, exporters' specific fixed effects were used to absorb the cross-sectional variation among olive oil exporters to the US market. The variability of export supply over time and geography was captured by the time-varying value of exports to all destinations for each exporter. In addition, the relative trade costs of destination markets were captured by the ratio of the trade unit value to the US relative to those of other destinations, and this for each exporter.

In the Heckman selection model, besides the network effect, other variables influence the extensive margin, including the exporter's GDP, population, and total exports revenues, as suggested by the Probit results. From an empirical point of view, the Heckman sample selection specification was the best model to fit the data and did not exhibit evidence of misspecification as suggested by the RESET test. The HPC test suggests that the Heckman sample selection with its explicit extensive margin fits the data better than the PPML approach does. PPML provides a legitimate way to incorporate zero but lacks the extensive margin component.

In future investigations and using shorter time series, one could further disaggregate olive oil at the HS-8 or HS-10 levels. In addition, a decomposition of bulk/non-bulk packaging, and a consideration of organic and labelled subcategories of olive oil could help provide deeper insights on the quality upgrade and market segmentation which took place over time in the US market.

6. Conclusions

The main conclusions are first that migrant networks from exporting countries matter a lot to rationalize the growth of the US olive oil market driving culinary changes in consumption and business network facilitating the entry of exporters. Secondly, price and income responses are large, especially for virgin olive oil, linking US economic prosperity and the growth of these markets while maintaining a competitive pricing environment.

These findings offer valuable insights for US olive oil traders, growers, trade policymakers, and stakeholders, enabling them to understand the reasons behind the growing interest in and spending on olive oil by US consumers within the context of a dynamic global olive oil market. This understanding empowers established and new stakeholders in the US olive oil market to

monitor emerging trends and devise appropriate strategies, including pricing and market segmentation in virgin olive oil to optimize their respective positions and overall well-being.

References

Alexandratos N., 2006. The Mediterranean diet in a world context. *Public Health Nutrition*, 9(1a): 111-117.

Baldwin R., Taglioni D., 2006. *Gravity for dummies and dummies for gravity equations* (No. w12516). National Bureau of Economic Research.

Beghin J.C., Schweizer H., 2021. Agricultural Trade Costs. *Applied Economic Perspectives and Policy*, 43(2): 500-530.

Bermudez A., Brey E., 2017. Is Spain becoming a country of emigration again? Data evidence and public responses. In: *South-North migration of EU citizens in times of crisis*. Cham: Springer, 83-98.

Bernabéu R., Díaz M., 2016. Preference for olive oil consumption in the Spanish local market. *Spanish Journal of Agricultural Research*, 14(4): e0108.

Bimbo F., Bonanno A., Viscecchia R., 2019. An empirical framework to study food labelling fraud: an application to the Italian extra-virgin olive oil market. *Australian Journal of Agricultural and Resource Economics*, 63(4): 701-725.

Brown D.J., Schrader L.F., 1990. Cholesterol information and shell egg consumption. *American Journal of Agricultural Economics*, 72(3): 548-555.

Cacchiarelli L., Carbone A., Laureti T., Sorrentino A., 2016. The Value of different Quality Clues in the Italian Olive Oil Market. *Italian Review of Agricultural Economics*, 71(1): 372-379.

Carbone A., Cacchiarelli L., Sabbatini V., 2018. Exploring quality and its value in the Italian olive oil market: a panel data analysis. *Agricultural and Food Economics*, 6(1): 6.

Casadei E., Valli E., Panni F., Donarski J., Gubern J.F., Lucci, P., Conte L., Lacoste F., Maquet A., Brereton P., Bendini A., Gallina Toschi T., 2021. Emerging trends in olive oil fraud and possible countermeasures. *Food Control*, 124: 107902.

CEPII, 2020. Database available online at: http://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele.asp. Retrieved March 2021.

Chern W.S., Loehman E.T., Yen S.T., 1995. Information, health risk beliefs, and the demand for fats and oils. *The Review of Economics and Statistics*, 77(3): 555-564.

Combes P.-P., Lafourcade M., Mayer T., 2005. The trade-creating effects of business and social networks: evidence from France. *Journal of international Economics*, 66.1: 1-29.

Da Silveira R., Vágula J.M., de Lima Figueiredo I., Claus T., Galuch M.B., Junior O.O.S., Visentainer J.V., 2017. Rapid methodology via mass spectrometry to quantify addition of soybean oil in extra virgin olive oil: A comparison with traditional methods adopted by food industry to identify fraud. *Food Research International*, 102: 43-50.

Del Giudice T., Cavallo C., Caracciolo F., Cicia G., 2015. What attributes of extra virgin olive oil are really important for consumers: a meta-analysis of consumers' stated preferences. *Agricultural and Food Economics*, 3(1): 20.

Dreher A., 2006. Does Globalization Affect Growth? Evidence from a new Index of Globalization, *Applied Economics*, 38(10): 1091-1110.

García Alvarez-Coque J.M., Martí Selva M., 2006. *A Gravity Approach to Assess the Effects of Association Agreements on Euromediterranean Trade of Fruits and Vegetables*, Working Paper 06/15, TRADEAG - Agricultural Trade Agreements.

Gould D.M., 1994. Immigrant links to the home country: empirical implications for US bilateral trade flows. *The Review of Economics and Statistics*, 76(2): 302-316.

Guerrero M., 2014. *Olive Oil Production in Spain Set to Rebound*. USDA/FAS. GAIN Report, number: SP1402. https://apps.fas.usda.gov. Retrieved: 14/10/2020.

Gygli S., Haelg F., Potrafke N., Sturm J.E., 2019. The KOF Globalisation Index – Revisited, *Review of International Organizations*, 14(3): 543-574. https://doi.org/10.1007/s11558-019-09344-2.

Hammami A.M., Beghin J.C., 2021. The Trade and Welfare Impacts of the US Retaliatory Tariff on EU Olive Oil. *Agricultural Economics*, 52(5): 807-818.

Hilger J., Rafert G., Villas-Boas S., 2011. Expert opinion and the demand for experience goods: an experimental approach in the retail wine market. *Review of Economics and Statistics*, 93(4): 1289-1296.

IFIC, 2018. *2018 Food and Health Survey*. https://foodinsight.org/2018-food-and-health-survey/. Retrieved: 27/09/2020.

International Olive Council (IOC), 2020. *Designations and definitions of olive oils*. https://www.internationaloliveoil.org/olive-world/olive-oil/. Retrieved: 10/02/2020.

Kalogeras N., Valchovska S., Baourakis G., Kalaitzis P., 2009. Dutch consumers' willingness to pay for organic olive oil. *Journal of International Food & Agribusiness Marketing*, 21(4): 286-311.

Karipidis P.I., Tsakiridou E., Tabakis N.M., Mattas K., 2005. Hedonic analysis of retail egg prices. *Journal of Food Distribution Research*, 36: 856-2016-56442: 68-73.

Kavallari A., Maas S., Schmitz P.M., 2010. Evidence on Euromediterranean trade integration: The case of German olive oil imports. *German Journal of Agricultural Economics*, 59(670-2016-45880): 40-46.

Kavallari A., Maas S., Schmitz P.M., 2011. Examining the determinants of olive oil demand in nonproducing countries: Evidence from Germany and the UK. *Journal of Food Products Marketing*, 17(2-3): 355-372.

kof.ethz.ch, 2020. *KOF Globalisation Index*. https://kof.ethz.ch/en/forecasts-and-indicators/indicators/kof-globalisation-index.html. Retrieved: 10/04/2020.

Melitz M.J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6): 1695-1725.

Menapace L., Colson G., Grebitus C., Facendola M., 2011. Consumers' preferences for geographical origin labels: evidence from the Canadian olive oil market. *European Review of Agricultural Economics*, 38(2): 193-212.

Milli S., 2006. Olive Oil Marketing on Non-Traditional Markets: Prospects and Strategies. *New Medit*, 5(1): 27-37.

National Library of Medicine, Pubmed.gov, 2020. https://pubmed.ncbi.nlm.nih.gov/. Retrieved: 10/02/2020.

NewsBank, 2020. https://www.newsbank.com/. Retrieved: 10/02/2020.

Office of Immigration Statistics (OIS), 2021. *Yearbook of Immigration Statistics*. https://www.dhs.gov/immigration-statistics/yearbook. Retrieved: 04/04/2021.

Ramsey J.B., 1969. Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society*: Series B (Methodological), 31(2): 350-371.

Rauch J.E., 1999. Networks versus markets in international trade. *Journal of International Economics*, 48(1): 7-35.

Rauch J.E., Trindade V., 2002. Ethnic Chinese networks in international trade. *Review of Economics and Statistics*, 84(1): 116-130.

Regmi A., Ballenger N., Putnam J., 2004. Globalisation and income growth promote the Mediterranean diet. *Public Health Nutrition*, 7(7): 977-983.

Ronen E., 2017. The Trade-Enhancing Effects of Non-Tariff Measures on Virgin Olive Oil. *International Journal of Food and Agricultural Economics*, 5(3): 9-26.

Santos Silva J.M.C., Tenreyro S., 2006. The log of gravity. *The Review of Economics and Statistics*, 88(4): 641-658.

Santos Silva J.M.C., Tenreyro S., Windmeijer F., 2015. Testing competing models for non-negative data with many zeros. *Journal of Econometric Methods*, 4(1): 29-46.

Scarpa R., Del Giudice T., 2004. Market Segmentation via Mixed Logit: Extra-Virgin Olive Oil in Urban Italy. *Journal of agricultural & food industrial organization*, 2(1): 20.

USDA-PS&D, 2020. *United States Department of Agriculture- Foreign Agricultural Service-Production, Supply and Distribution*. https://apps.fas.usda.gov/psdonline/app/index.html#/app/advQuery. Retrieved in March 2021.

US Trade Census, 2020. *USA Trade Online*. https://usatrade.census.gov/. Retrieved in March 2021.

Vance C., Ritter N., 2014. Is peace a missing value or a zero? On selection models in political science. *Journal of Peace Research*, 51(4): 528-540.

Vlontzos G.N., Duquenne M.N., 2008. Greek Olive Oil: How can its International Market potential be realized? *Estey Journal of International Law and Trade Policy*, 9(1753-2016-141197): 32-47.

Yan J., Erasmus S.W., Toro M.A., Huang H., van Ruth S.M., 2020. Food fraud: Assessing fraud vulnerability in the extra virgin olive oil supply chain. *Food Control*, 111: 107081.

Yotov Y.V., Piermartini R., Monteiro J.A., Larch M., 2016. *An advanced guide to trade policy analysis: The structural gravity model*. Geneva: World Trade Organization.

Xiong B., Sumner D., Matthews W., 2014. A new market for an old food: the US demand for olive oil. *Agricultural Economics*, 45(S1): 107-118.

Appendix

Summary Table of Regression Variables

Variable	*definition*	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*	*Unit*	*Source*
FL_1509	Bilateral trade volume HS 1509	567	1.06E+07	3.00E+07	0	1.60E+08	Kg	US Trade Census, 2020
FL_150910	Bilateral trade volume HS 150910	567	6487873	1.86E+07	0	1.10E+08	Kg	US Trade Census, 2020
FL_150990	Bilateral trade volume HS 150990	567	4078652	1.27E+07	0	8.25E+07	Kg	US Trade Census, 2020
GDP_USA	Aggregate US income	567	1.38E+13	2.43E+12	9.38E+12	1.79E+13	USD	CEPII database, 2020
IUV _1509	Relative bilateral unit value HS 1509	554	1.255093	2.937924	0.021978	68.55963	Ratio, index	US Trade Census, 2020
IUV _150910	Relative bilateral unit value HS 150910	557	1.714746	10.25325	0.2003365	217.1838	Ratio, index	US Trade Census, 2020
IUV _150990	Relative bilateral unit value HS 150990	550	1.397808	3.715903	0.1140869	78.16504	Ratio, index	US Trade Census, 2020
STOCK_MIG	Stock of migrants from country *i* in the US	567	119556.6	112751.2	8	505831	Person	OIS, 2021
KOFCUGIDF	KOF index of globalization	567	81.56524	7.051316	71.24317	91.72095	index	kof.ethz.ch, 2020; Gygli *et al.*, 2019
POP_EXPORTER	Exporter's population	567	34.22425	41.31275	0.6106	209.4693	Person	CEPII database, 2020
GDP_EXPORTER	Exporter's aggregate income	558	4.85E+11	6.40E+11	1.10E+10	2.57E+12	USD	CEPII database, 2020
US_ CONSUMPTION	US olive oil consumption	567	236.4815	75.76036	104	364	1000 tons	USDA-FAS, PS&D, 2020
TOT_ EXPORTS~1509	Total exports 1509 by country *i*	562	2.09E+08	5.43E+08	0	4.16E+09	Tonnes	FAOStat, 2021
TOT_ EXPORTS~150910	Total exports 150910 by country *i*	521	1.61E+08	4.05E+08	65	3.15E+09	Tonnes	FAOStat, 2021
TOT_ EXPORTS~150990	Total exports 150990 by country *i*	495	4.05E+07	1.06E+08	0	8.84E+08	Tonnes	FAOStat, 2021
MEMBER_EU_O	Membership in the EU	567	0.3015873	0.4593523	0	1	Dummy	CEPII database, 2020

The sustainability of sugarcane production: A systematic review of sustainability indicators

Gabriel Chico Viegas*, Ana Marta-Costa**,
Rui Fragoso***, Edgar Cambaza*

DOI: 10.30682/nm2404h
JEL codes: Q01, Q13

Abstract

This review aims to identify indicators to assess the sustainability of sugarcane production. Preferred Reporting Items for Systematic Reviews and Meta-analyses reviewed 44 eligible articles from Scopus and Web of Science databases. The results show that only 34% of the identified studies focus on the three sustainability dimensions: environmental, economic and social. The social dimension continues to be under-studied. From the 119 indicators identified, 16 common indicators used to assess the sustainability of sugarcane were identified, and the analysis of frequency shows that the most commonly suggested indicators were Greenhouse gas emissions, Water use, Water quality, Employment generation, and Initiatives to promote the Local community's welfare, Profit, and Distance to sugar mills. Further, research combining the three dimensions of sustainability and those that separately evaluate the sustainability of sugarcane in the production stage on the field and the factory is recommended.

Keywords*: Sustainability indicators, Sustainability dimensions, Sugarcane production systems, PRISMA framework.*

1. Introduction

Sugarcane plantations are one of the world's fastest-growing agricultural systems, contributing around 21% of global crop production between 2000 and 2020 and generating significant revenue for many nations. Although sugarcane production has positively impacted the economy, it has led to increased greenhouse gas emissions, biodiversity loss, soil erosion, nutrient loss, soil pollution, expropriation of land and resources, displacement of local populations, and land tenure disputes (Hall *et al*., 2017; Souto *et al*., 2024). In this context, assessing sugarcane production sustainability is crucial for agribusiness survival and market competitiveness, ensuring a sustainable future through environmentally, economically, and socially responsible farming practices (TraceX Technologies, 2023).

They are the operational representation of a system's attributes and allow for measuring and monitoring changes that are relevant to human and environmental well-being (Gamboa *et al*., 2016; Waas *et al*., 2014).

Although numerous approaches and indicators have been created and examined to assess the sustainability of agricultural systems over the years, finding a consensus model for efficient

* ISCED Open University, Beira, Sofala, Mozambique.
** University of Trás-os-Montes and Alto Douro, Centre for Transdisciplinary Development Studies (CETRAD), Vila Real, Portugal.
*** University of Évora, Center for Advanced Studies in Management and Economics (CEFAGE), Évora, Portugal.
Corresponding author: gabrielviegas19@gmail.com

assessment of the sustainability of sugarcane production has proven challenging (Latruffe *et al*., 2016; Rasmussen *et al*., 2017). Therefore, the systematic literature review will be the first to combine different indicators to assess the sustainability of sugarcane, creating groups of indicators that can be used as a reference to assess the sustainability of this crop in the future. Combining indicators from the main pillars of sustainability can help us understand how sugarcane production systems affect the environment, economy, and society. The results will guide sugarcane farmers, academics, and policymakers on which indicators to prioritize in production, strategies, and policy.

The review has six parts: (1) Introduction, which contextualizes the study; (2) Literature Review, which discusses sustainability and sustainability indicators; (3) Methodology, which describes the article's construction; (4) Results, which presents the findings; (5) Discussion, which analyzes and contrasts the findings to other studies with a similar research focus; and (6) Conclusion, which answers the objectives and identifies the article's contribution, limitations, and suggestions for future studies.

2. Literature review

2.1. *Sustainable Development and Sustainability*

According to the World Commission on Environment and Development (1987), sustainable development is economically viable, environmentally sound, and socially acceptable development that meets current demands without compromising future needs. However, Pham and Smith (2014) argue that this concept is ambiguous and open to interpretation. Although sustainable development is often used as a synonym for sustainability, they are two different concepts with different meanings. Cruz *et al*. (2018) add that sustainability is the period that a system can maintain itself through a set of management strategies, while Garcia-Bustamante *et al*. (2018) emphasize that sustainability is characteristic of dynamic systems that allow them to keep themselves through time with no discernible endpoint. In summary, sustainability incorporates environmental, economic, and social factors, whereas sustainable development improves economic well-being and quality of life over time without compromising future generations' demands. Ensuring sustainability in different activity systems is the way to achieve sustainable development, which has become a global imperative in recent years (Depetris-Chauvin *et al*., 2023).

In the context of agricultural systems, different authors have deduced the application of sustainability to balance the environmental, economic and social dimensions (Cruz *et al*., 2018; Latruffe *et al*., 2016; Marta-Costa *et al*., 2022; Mokrani *et al*., 2022). Corvo *et al*. (2021) refer to the environmental dimension as focusing on the environmentally sustainable behavior of human activities. Karvonen *et al*. (2017) associate the economic dimension with how agricultural systems improve people's quality of life. Unfortunately, receiving less attention than other dimensions (Massuça *et al*., 2023), Janker and Mann (2020) examine the social dimension, focusing on how agricultural systems influenced by environmental and economic factors contribute to human well-being and social justice.

According to Gayatri *et al*. (2016), focusing on just one dimension of sustainability to the detriment of others can be risky. The same authors emphasize that agricultural production systems cannot be sustainable if they cannot produce enough food or pay farmers sufficiently, even if they maintain environmental quality. Similarly, high-productivity agricultural systems that use more inputs to offset yields or cause environmental damage are not sustainable.

Therefore, it is essential to combine the three dimensions to obtain results that truly reflect the genuine concept of sustainability.

Growing demand for sugar and biofuels has led to increased sugarcane cultivation, raising a myriad of negative impacts on sugarcane production. Garcia-Bustamante *et al*. (2018) identified soil degradation, high water consumption, land use change, atmospheric pollution, inequity in the rural sector, low salaries, and even laborer exploitation as some of the challenges of sugarcane production. Sustainability assessment plays a crucial role in assessing sugarcane pro-

duction's environmental, social, and economic impacts and identifying areas for improvement.

2.2. *Methodologies for assessing sustainability in agriculture*

According to Cruz *et al*. (2018), two main approaches are used to assess sustainability in agriculture, the "bottom-up" and "top-down". The first approach consists of the joint selection of sustainable development indicators and subsequent construction of the global framework in a participatory manner. The "bottom-up" approach is inclusive and can obtain better results by incorporating stakeholders in framework design. On the other hand, the "top-down" approach begins by defining the global sustainability assessment structure, which is further disaggregated into a set of indicators.

Methodologies such as Life Cycle Assessment (LCA), Sustainability Assessment Frameworks (SAFs), and Multi-Criteria Analysis (MCA) have been developed to assess agricultural sustainability (Lacovidou & Voulvoulis, 2018). LCA has gained wider acceptance and is used by most professionals worldwide (Kralisch *et al*., 2015). It quantifies the environmental impacts of sugarcane production throughout its life cycle, from cultivation to processing and consumption (Silalertruksa *et al*., 2017).

For the sugar sector, sustainability assessment has been carried out mainly for sugarcane biorefineries with the integration of sugar, ethanol and the production of by-products. Duarte *et al*. (2013) and Silalertruksa *et al*. (2017) used LCA to assess the environmental footprints of sugarcane ethanol production in Brazil and Thailand, and they identified greenhouse gas emissions, water use, and land use changes as having significant environmental impacts on sugarcane production. Coutinho *et al*. (2017) used a participatory SAF to assess the sustainability of sugarcane expansion in Brazil, involving sugarcane producers, government officials, and environmental organizations, and identify and prioritize sustainability indicators.

Gnansounou *et al*. (2017) employed MCA to assess the sustainability of selected sugarcane biorefinery-centered systems in Brazil, considering environmental, social and economic criteria and comparing their sustainability performance. Based on the same dimensions, Turetta *et al*. (2017) developed a framework to assess the sustainability of the sugarcane sector in Brazil.

In the expectation of improving the sustainability assessment process, the composite sustainability index has been developed for sugar farming through an Analytic Hierarchy Process (AHP) and available ecological, technological and socio-economic data (Aguilar-Rivera, 2019). With the global pressure and appeal for sustainability in all activities, constructing more inclusive, transparent and replicable evaluation methods to assess sustainability became necessary. The SustenAgro Support System, developed by Embrapa in Brazil, has been used to analyze the impacts of different management practices on the sustainability of sugarcane production. It is based on a set of indicators and criteria that cover the three dimensions of sustainability that evaluate and compare systems individually (De Jesus *et al*., 2019). Besides being a current method, SustenAgro was developed by sugarcane experts, includes the three dimensions of sustainability, and can be applied to individual production units for precise, easy-to-interpret results.

However, for sustainability to be assessed, the corresponding indicators must be available (Prasara-A & Gheewala, 2021; Prasara-A *et al*., 2019). According to Cruz *et al*. (2018), sustainability indicators are tools that can evaluate the effects of management changes. They can be used to monitor the trends of a particular condition and help identify challenges that may require additional resources (De Jesus *et al*., 2019). They can also compare sustainability performances between farms, regions and countries. The indicators must satisfy specific requirements, which Shukor and Ng (2022) summarize as follows: (i) be comparable; (ii) show the true nature of the process or function they represent; (iii) be understandable, more reliable and accessible; (iv) be able to build the same criteria and compare them using time series and units; (v) be obtained on a regular enough basis of time to guarantee that the firms take occasional action; and (vi) be understandable for the user and match the user's information needs.

In agriculture, methods such as Delphi, *Indicateurs de Durabilité des Exploitations Agricoles* (Agricultural Sustainability Indicators, IDEA), *Marco para la Evaluación de Sistemas de Manejo incorporando Indicadores de Sustentabilidad* (Management Systems Assessment Framework Incorporating Sustainability Indicators, MESMIS), Monitoring Tool for Integrated Farm Sustainability (MOTIFS), Response-Inducing Sustainability Evaluation (RISE), or Sustainability Assessment of Farming and the Environment Framework (SAFE) have been developed to evaluate sustainability in farming using composite indicators of agricultural sustainability.

The practicality in the procedures and guarantee of the inclusion of the participants of the study area in the process of building, selecting, and validating the indicators has become Delphi well adapted in agriculture (Naisola-Ruiter, 2022; Zhao *et al*., 2023). Its main features are its anonymity and the ability to reduce the influence of one expert to another, contributing to the objectivity of the results (Bélanger *et al*., 2012; Cruz *et al*., 2018).

Sustainability indicators are grouped into economic, environmental, and social (Gani *et al*., 2021), and depending on these aspects, there are different indicators for assessing sustainability in sugarcane systems. According to Garcia-Bustamante *et al*. (2018), Aguilar-Rivera (2022) and Joglekar *et al*. (2022), some examples of indicators are Greenhouse gas emissions, Water use, Biodiversity and Workers' rights.

There are some initiatives to advance the development and application of a universal set of indicators for sugarcane, such as the Bonsucro Production Standard, Bioindicators for the Sustainability of the Sugar Agro-Industry and the Sustainable Sugarcane Initiative (Aguilar-Rivera, 2019). However, these indicators are not universally applicable across contexts and scales (Sawaengsak *et al*., 2019), due to the local conditions, such as soil, climate, social development and others (Aguilar-Rivera, 2022), needing to continually carry out research in different realities to identify common indicators specific to each context.

3. Methodology

The systematic review comprised identifying and selecting publications for analysis using the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flowchart (see Figure 1). PRISMA is a framework that guarantees that systematic review reports are more transparent, clear and comprehensive (Li *et al*., 2020). Credible Scopus and Web of Science databases were used in selecting the publications, which have comprehensive global and regional coverage of scientific journals, conference proceedings and books such as pointed out by Baas *et al*. (2020) and Birkle *et al*. (2020). These databases ensure that the highest quality data are indexed through rigorous content selection and re-evaluation by an independent Content Selection and Advisory Board.

Forty-four (44) articles containing the keywords "sugarcane" and "sustainability" in their titles, published between 2013 to 2022, and written in English, Portuguese, and Spanish, were identified using the mentioned databases. The past ten years were selected because important sustainability and climate change events like the Paris Agreement, the UN Sustainable Development Summit, COP26, and COP27, whose decisions impacted agricultural activities, have been highlighted. The inclusion criteria for the systematic review were studies that focused on sugarcane, addressed sugarcane sustainability, and provided the full text of the article. In addition to these criteria, for discussion and achievement of the study objectives, articles that identified or mentioned sustainability indicators received more attention. Documents such as books, book chapters, reviews and conference papers were excluded from the analysis. These items are excluded because books and book chapters are not primary sources and conference articles are not peer-reviewed. The search of articles was open to all subject areas and the data was extracted on July 18, 2023.

To identify the common indicators, the authors followed the adapted approach from Nadaraja *et al*. (2021), which is considered a common indicator - identified in at least two studies. However, it should be noted that this approach can lead to a false impression of the high frequency of an indi-

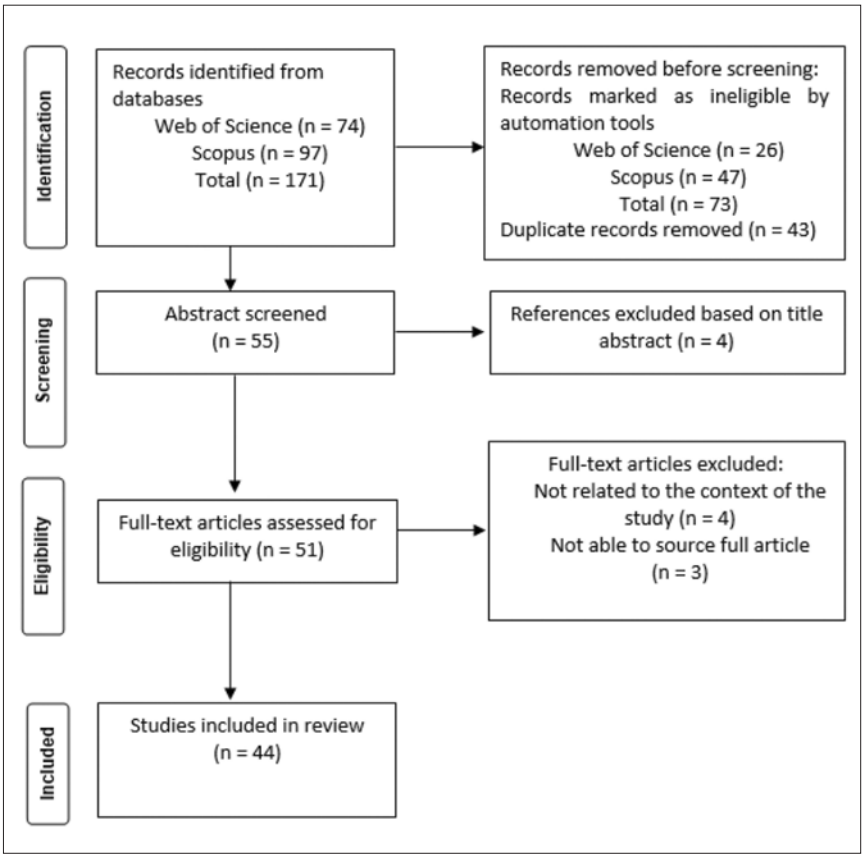


Figure 1 - Outline of the study selection flow diagram.

cator (common and very relevant) because it uses as a reference the number of studies where the indicator is present and not the number of times an indicator has been quoted/researched by different authors. Thus, this situation opens space for situations where the same indicator is present in other studies of the same author. To address this limitation, the present review included indicators indicated in at least two studies by different authors. However, indicators that were frequent in studies by the same author were also included because it is believed that their continued presence in the studies, although these are from the same authors, shows the indicator's relevance.

Information about the dimensions of sustainability addressed in each article and the sustainability indicators identified was extracted and evaluated from the selected articles. Finally, the common indicators used in assessing sugarcane sustainability were analyzed by calculating their frequency in the sample.

4. Results

4.1. *Characterization of the selected database*

The year 2022 had the most publications (10) about the sustainability of sugarcane production out of the 44 total articles reviewed, with 2018 coming in second with seven publications. The least number of publications, two (2), were recorded in 2013 and 2015 (see Figure 2). The increased trend in the number of publications reflects the concern of the government, researchers, and society about sustainable development.

Brazil, the world's largest sugarcane producer, has been the leading country where published studies were conducted, accounting for 41% (see Figure 3). Following Brazil, Thailand and India contributed 18% and 14% of all publications, respectively. The remaining studies were conducted in Mexico, China, Colombia, Cuba, Ecuador, Indonesia, Iran, Jamaica, Portugal, and South Africa.

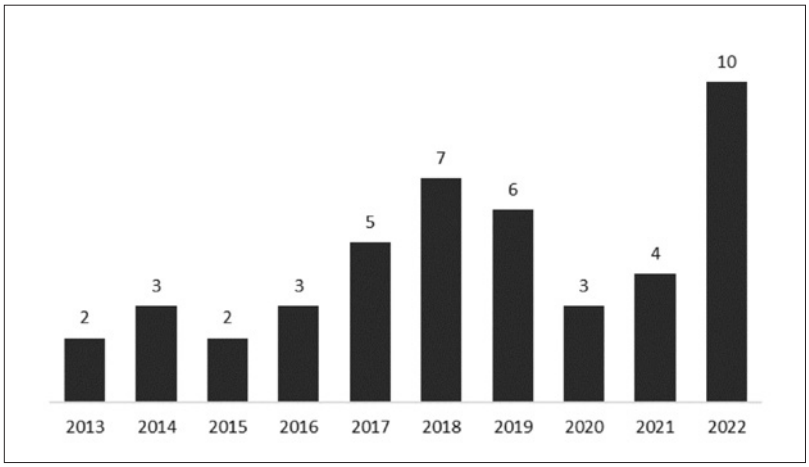
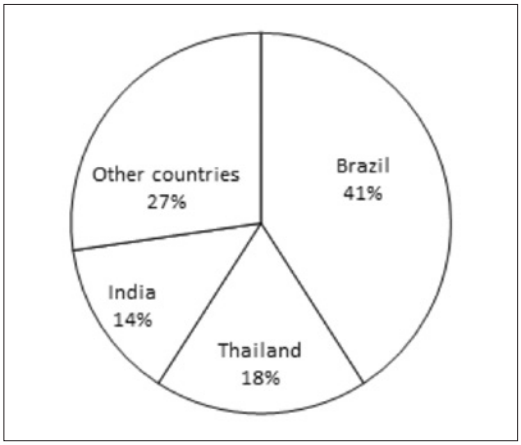


Figure 2 - Number of publications per year between the period of 2013 and 2022.

Table 1 shows the distribution of studies by sustainability dimension, according to the indicators evaluated in each work. During the period analyzed, the largest number of studies (15) focused on the environmental, economic, and social dimensions together. Seven studies focused on the environmental and economic dimensions, four on the environmental and social dimensions, and three on the economic and social dimensions. Additionally, nine studies focused exclusively on the environmental dimension and one on the social dimension. Five studies addressed sugarcane sustainability without emphasizing any specific dimension of sustainability. No studies were found focusing solely on the economic dimension of sustainability.

Figure 3 - Publications per countries where the studies were conducted between 2022 and 2023.



4.2. *Sugarcane sustainability indicators identification*

From sampled studies, 119 indicators were found in all three dimensions of sustainability. The economic dimension covered the highest number of indicators (38.7%), followed by environmental (31.1%) and social (30.3%).

4.2.1. *Environmental dimension*

The environmental dimension had 37 indicators in the Atmosphere and Water, Land and Biodiversity, Materials, and Energy themes. Of the Atmosphere and Water theme's indicators, Greenhouse gas emissions (GHG), Water use and Water quality were the most common (see

Table 1 - Frequency of studies by sustainability dimension addressed.

Dimensions (n = 44 articles)	*Frequency*	*Percentage*
Environmental + Economic + Social	15	34.1
Environmental + Economic	7	15.9
Environmental + Social	4	9.1
Economic + Social	3	6.8
Environmental	9	20.5
Social	1	2.3
None	5	11.4

Table 2 - Indicators under the environmental dimension.

Themes	*Indicators*		
	Designation	*Frequency*	*Sources*
Atmosphere and water	GHG emissions	7	De Jesus *et al*. (2019); Leal and Nogueira (2014); Lopez-Ortega *et al*. (2021); Poli *et al*. (2022); Prasara-A *et al*. (2019); Silalertruksa *et al*. (2015); Yani *et al*. (2022)
	Water use	2	Lopez-Ortega *et al*. (2021); Prasara-A *et al*. (2019)
	Water quality	3	Leal and Nogueira (2014); Yani *et al*. (2022); Lopez-Ortega *et al*. (2021)

Table 3 - Indicators under the social dimension.

Themes	*Indicators*		
	Designation	*Frequency*	*Source*
Equity and Decent Livelihood	Employment generation	6	Prasara and Gheewala (2016); Prasara-A *et al*. (2019); Prasara-A and Gheewala (2021); Leal and Nogueira (2014); Lopez-Ortega *et al*. (2021); Poli *et al*. (2022)
	Initiatives to promote the local community's welfare	2	De Jesus *et al*. (2019); Yani *et al*. (2022)

Table 2). The higher frequency of GHG emissions in the study sample indicates that researchers value it when measuring the sustainability of sugarcane production.

4.2.2. *Social Dimension*

The social dimension included 36 indicators under Labor Rights, Safety and Health, and Equity and Decent Livelihood themes. The most common indicators in this dimension were Employment generation and Initiatives to promote the local community's welfare, from the Equity and Decent Livelihoods themes. Six papers cited "Employment generation", reflecting researchers' strong belief in its impact on communities (see Table 3).

Six (6) additional social indicators were identified, which we believe are relevant and deserve to be highlighted in this study due to the authors' interest in them. These are Conflict with land rights, Wages, and Wage satisfaction from Equity and Decent Livelihoods, Forced labor in the field, Freedom of collective bargaining, and Accidents in the previous year from Labor Rights and Safety and Health.

4.3. *Economic Dimension*

The economic dimension had 46 indicators. The most common indicators in this dimension were Cane yield, Sugarcane productivity, Production and Profit, from the investment theme, and Distance to sugar mills from the Local Economy and Product Quality (see Table 4). Cane yield was the indicator most cited (3 times).

Therefore, according to the results, 16 common indicators were found. However, only five (5) indicators were measured. Water use, Cane yield, and Distance to the sugar mill were measured through interviews with sugarcane farm owners, while Employment generation was measured through interviews with workers, and management and Wage satisfaction were measured through interviews with farm workers. Most indicators (68.8%) were presented without measurement (see Table 5).

5. Discussions

The article reviewed the literature on the sustainability of sugarcane production and identi-

Table 4 - Indicators under the economic dimension.

Themes	*Indicators*		
	Designation	*Frequency*	*Sources*
Investment	Cane yield	3	Aguilar-Rivera (2019); Lopez-Ortega *et al*. (2021); Prasara and Gheewala (2016)
	Sugarcane productivity	2	De Jesus *et al*. (2019); Leal and Nogueira (2014)
	Sugarcane production	2	Aguilar-Rivera (2019); Yani *et al*. (2022)
	Profit	2	Lopez-Ortega *et al*. (2021); Yani *et al*. (2022)
Local Economy and Product Quality	Distance to sugar mills	2	Aguilar-Rivera (2019); Prasara and Gheewala (2016)

fied the most common indicators for assessing the sustainability of the sugarcane production systems. The results of Figure 2 show that more attention on this topic has increased after 2016, and this could be attributed to two important events, namely, the UN Sustainable Development Summit in 2015 and the entry of the Paris Agreement into force in 2016. These milestones have influenced governments, researchers, and societies to emphasize the importance of sustainable practices in sugarcane production systems. In practical terms, most studies (54.6%) present apparent gaps due to the non-inclusion of the three main dimensions in the same research, thus not reflecting the real effect of sugarcane production on the three main aspects of sustainability. Kristensen and Mosgaard (2020) and Rajeev *et al*. (2017) argue that for a production system to be considered sustainable, there must be a joint integration of the environmental, economic and social aspects of the value chain within the same framework. Therefore, not including the three main dimensions in studies may pose a risk to the conclusions being drawn about the sustainability of the system under assessment.

Climate change is now the target of global attention. It is reflected in several initiatives and agreements that aim to reduce its adverse impacts on agriculture, health, education, and other areas, contributing to greater awareness in the community. Therefore, the presence of 79.5% of articles discussing the environmental dimension of sustainability is a reflection of this concern with environmental issues, specifically in the cultivation of sugarcane, which contributes approximately 11% of the world's agricultural waste and 400 million tons of CO2 equivalent annually (BONSUCRO, 2023; Bordonal *et al*., 2018; De Figueiredo *et al*., 2010). On the other hand, the social dimension receives less attention. The weak presence of the social dimension in studies on sugarcane sustainability is in line with Massuça *et al*. (2023) findings. This situation is associated with the fact that it addresses aspects related to the well-being of communities, which is often challenging to define consensual methods of measurement and interpretation, leaving subjectivity and a lack of clarity in the definition and measurement of the corresponding indicators (Baffoe & Mutisya, 2015; Bubicz *et al*., 2019; Hale *et al*., 2019). Boström (2012) adds that the meaning of social sustainability remains unclear and does not have an evident scientific basis for measuring it. Similarly, Eizenberg and Jabareen (2017), Gaviglio *et al*. (2016), Janker *et al*. (2019), and Sidhoum (2018) emphasize that lack of cohesion in the perception of the definition of the social dimension creates confusion in determining the means of operationalization.

Although the environmental dimension receives the most attention, economic indicators were more prevalent in this study than environmental indicators. This scenario can be attributed to the fact that economic indicators are usual-

Table 5 - Description of common indicators.

Indicator	*Description*	*Measurement*
GHG emissions	Emissions of CO2, CH4, and N2O at each stage of the supply chain (van Eijck *et al*., 2014)	Not mentioned
Water use	Quantity of water required to irrigate and process crops within the plantation systems (Garcia-Bustamante *et al*., 2018; Prasara-A *et al*., 2019).	Interviews with sugarcane farm owners
Water quality	Chemical, physical and biological characteristics of water based on patterns of its use (Chapman, 2021).	Not mentioned
Employment generation	Number of annual full-time equivalent jobs generated by tons of sugarcane produced and processed (Lopez-Ortega *et al*., 2021; Prasara-A *et al*., 2019)	Interviews with workers and management
Initiatives to promote the local community's welfare	Initiatives carried out in communities to mitigate the negative impacts that certain activities may have on the community (Ahmad & Nomani, 2015).	Not measured
Conflict with land rights	Acts against farmers' rights in their land (Boone, 2019).	Not measured
Wages	Key agricultural activities' daily wages without a meal (Hassan & Kornher, 2022; Prasara-A *et al*., 2019).	Not mentioned
Wage Satisfaction	Workers' feelings towards their wages (Stander *et al*., 2019).	Interviews with farm workers
Forced labor in the field	Work required of anyone in a non-voluntary way (Machado *et al*., 2017).	Not mentioned
Freedom of collective bargaining of workers	Negotiation between organizations and employees to determine the terms and conditions of employment (Nwokocha, 2015).	Not mentioned
Accident in the previous year	Refers to the unexpected incident that usually causes harm or injury.	Not mentioned
Cane yield	Quantity of cane produced within a given period per unit of land area harvested (Fischer, 2015).	Interviews with sugarcane farm owners
Sugarcane productivity	Amount produced by farm given a set of resources and inputs (FAO, 2017).	Not mentioned
Sugarcane production	Quantity of the product sold or consumed and normally recorded in tons (FAO, 2013)	Not mentioned
Profit	Financial gain made when the income exceeds expenses, costs, and taxes (Nadaraja *et al*., 2021)	Not mentioned
Distance to sugar mills	Refers to the distance from the sugarcane field production to the factory.	Interview with sugarcane farm owners

ly quantitative and, therefore, apparently easier to measure in the researchers' view compared to environmental indicators, which are more complex (Cruz *et al*., 2018). However, the results found in the present study are contrary to those obtained by some studies, which address sustainability indicators in agricultural production systems despite not mentioning specific crops. Nadaraja *et al*. (2021) found the highest number of environmental indicators, followed by social and economic indicators. Unlike this, Bathaei and Štreimikienė (2023) found the highest number of indicators in the environmental dimension, followed by economic and social. These findings show that the selection of indicators per dimension to be studied by the different authors does not follow a specific rule but rather depends on the researcher's interest, the research objectives and the study context.

Of the 16 common indicators identified, GHG emissions, Employment generation, Water quality and Sugarcane yield had the highest frequen-

cy in the studies. The high frequency of these indicators is associated with the fact that sugarcane is a crop focused on income and, therefore, requires large areas of land, a large workforce, and the use of large quantities of agrochemicals. Due to the intense application of fertilizers, pesticides, herbicides, and the recurrent practice of burning in the fields, this crop contributes to the emission of large quantities of greenhouse gases and negative effects on water quality. This situation becomes a dilemma for this sector, which, on the one hand, struggles to increase the yield of sugarcane to meet the demands of the international market and, on the other, to reduce the emission of GHGs into the atmosphere to comply with global requirements. Nadaraja *et al*. (2021) and Bathaei and Štreimikienė (2023) also noted these indicators as common and vital to agricultural sustainability evaluation.

Most selected indicators (68.8%) were not measured in the studied sample. Bélanger *et al*. (2012) emphasize that indicators are used to simplify, quantify and communicate efficiently. In this context, if the indicators are not quantified, they do not fulfill the function for which they were developed. Some of the challenges in measuring indicators in sugarcane production are related to the high cost of equipment, the subjectivity of information due to dependence on data provided by the employer itself, and the variations of the payment mode in this sector (Lopez-Ortega *et al*., 2021; Nadaraja *et al*., 2021; Prasara-A *et al*., 2019). To fix this problem and ensure that available, measurable and applicable indicators are chosen, they must be selected using a bottom-up approach that includes all the people with a stake in the value chain. This way, users can also easily adopt and accept the indicators.

6. Conclusions

Although the environmental dimension is addressed in most studies, indicators of the economic dimension were frequent due to their ease of identification and measurement. The social dimension continues to receive the least attention from authors in studies evaluating the sustainability of sugarcane, despite the significant pressure that the sugar sector has faced in recent years to pay more attention to the conditions of workers and the communities where production takes place.

The selection of sustainability indicators and dimensions for study does not follow a specific rule but rather depends on the researcher's interest, research objectives and the context of the study itself. Overall, 16 common indicators used to assess the sustainability of sugarcane were identified and should receive special attention in future studies on this crop. For instance, indicators such as GHG emissions; water use, water quality, employment generation, and Initiatives to promote the local community welfare, crop yield, sugarcane productivity, sugarcane production, profit and distance to sugar mills are considered to have the highest impact on sustainability in sugarcane production. These results indicate that greater attention should be paid to these indicators in the sugarcane production process to enable more efficient management and ensure sustainability across different production systems. Given the nature of cash crops primarily aimed at commerce, utilizing large land areas and requiring substantial labor, these indicators can also be applied in analyzing the sustainability of other crops such as soybeans, tobacco, sesame, and others, while acknowledging the specific characteristics of each crop. However, the indicators identified may not be universally applicable due to various factors associated with sugarcane production, such as geographical location, cultural differences, and political influences. One of the major challenges encountered in the selected studies was the difficulty in measuring a significant portion of the identified indicators due to their complexity, which could limit their utility for different stakeholders. Therefore, it is essential to identify and analyze the most common indicators used in sugarcane across different geographic areas and conditions, ensuring that these indicators are adapted to the specific context under study.

Due to its focus, it is important to employ bottom-up approaches for indicator selection in further research. This approach allows for the incorporation of perspectives from various stakeholders and considers the diverse realities of sugarcane production worldwide, as exemplified by the SustenAgro Support System.

Finally, it is important to conduct studies that evaluate the sustainability of sugarcane in the production process in the field and during the transformation phase in the factory. This approach will provide insights into sustainability and its influencing factors across different contexts and perspectives. It will also help identify the most relevant indicators for each phase of sugarcane production.

Acknowledgments

This research was supported by FCT - Fundação para a Ciência e Tecnologia under the projects UIDB/04011/2020 (https://doi.org/10.54499/UIDB/04011/2020) and UIDB/04007/2020 (https://doi.org/10.54499/UIDB/04007/2020).

References

Aguilar-Rivera N., 2019. A framework for the analysis of socioeconomic and geographic sugarcane agro industry sustainability. *Socio-Economic Planning Sciences*, 66: 149-160. https://doi.org/10.1016/j.seps.2018.07.006.

Aguilar-Rivera N., 2022. Bioindicators for the sustainability of sugar agro-industry. *Sugar Tech*, 24(3): 651-661. https://doi.org/10.1007/s12355-021-01105-z.

Ahmad S., Nomani A., 2015. Problems of corporate social responsibility in Sugarcane value chain in Utter Pradesh: an analysis through integrated model of CCSR. *Management Studies and Economic Systems*, 1(4): 207-212.

Baas J., Schotten M., Plume A., Côté G., Karimi R., 2020. Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative science studies*, 1(1): 377-386. https://doi.org/10.1162/qss_a_0001.

Baffoe G., Mutisya E., 2015. Social sustainability: A review of indicators and empirical application. *Environmental Management and Sustainable Development*, 4(2): 242-262. https://doi.org/10.5296/emsd.v4i2.8399.

Bathaei A., Štreimikienė D., 2023. A Systematic Review of Agricultural Sustainability Indicators. *Agriculture*, 13(2): 241. https://doi.org/10.3390/agriculture13020241.

Bélanger V., Vanasse A., Parent D., Allard G., Pellerin D., 2012. Development of agri-environmental indicators to assess dairy farm sustainability in Quebec, Eastern Canada. *Ecological indicators*, 23: 421-430. https://doi.org/10.1016/j.ecolind.2012.04.027.

Birkle C., Pendlebury D., Schnell J., Adams J., 2020. Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, 1(1): 363-376. https://doi.org/10.1162/qss_a_00018.

BONSUCRO, 2023. *Where do GHG emissions in sugarcane come from?* Retrieved from https://bonsucro.com/ghg-emissions-in-sugarcane/ (acessed at 21.02.2024).

Boone C., 2019. Legal empowerment of the poor through property rights reform: Tensions and trade-offs of land registration and titling in sub-Saharan Africa. *The Journal of Development Studies*, 55*(*3): 384-400. https://doi.org/10.1080/00220388.2018.1451633.

Bordonal R.O., Carvalho J.L.N., Lal R., Figueiredo E.B., Oliveira B.G., La Scala N., 2018. Sustainability of sugarcane production in Brazil. A review. *Agronomy for Sustainable Development*, 38: 1-23. https://doi.org/10.1007/s13593-018-0490-x.

Boström M., 2012. A missing pillar? Challenges in theorizing and practicing social sustainability: introduction to the special issue. *Sustainability: Science, practice and policy*, 8(1): 3-14. https://doi.org/10.1080/15487733.2012.11908080.

Bubicz M.E., Barbosa-Póvoa A.P.F.D., Carvalho A., 2019. Incorporating social aspects in sustainable supply chains: Trends and future directions. *Journal of Cleaner Production*, 237. https://doi.org/10.1016/j.jclepro.2019.06.331.

Chapman D.V., 2021. *Water quality assessments: a guide to the use of biota, sediments and water in environmental monitoring*. London: CRC Press, 656 pp. https://doi.org/10.1201/9781003062103.

Corvo L., Pastore L., Antonelli A., Petruzzella D., 2021. Social impact and sustainability in short food supply chains: An experimental assessment tool. *New Medit*, 20(3): 175-189. https://doi.org/10.30682/nm2103l.

Coutinho H.L.C., Turetta A.P.D., Monteiro J.M.G., Castro S.S., Pietrafesa J.P., 2017. Participatory sustainability assessment for sugarcane expansion in Goiás, Brazil. *Sustainability*, 9*(*9): 1573. https://doi.org/10.3390/su9091573.

Cruz J.F., Mena Y., Rodríguez-Estévez V., 2018. Methodologies for assessing sustainability in farming systems. In: Gokten S., Gokten P.O. (eds.), *Sustainability Assessment and Reporting,* IntechOpen, pp. 33-58. https://doi.org/10.5772/intechopen.79220.

De Figueiredo E.B., Panosso A.R., Romão R., La Scala N., 2010. Greenhouse gas emission associ-

ated with sugar production in southern Brazil. *Carbon balance and management*, 5: 1-7. https://doi.org/10.1186/1750-0680-5-3.

De Jesus K.R.E.D., Torquato S.A., Machado P.G., Brumatti Zorzo C.R., Cardoso B.O., Leal M.R.L.V., Picoli M.C.A., Ramos R.C., Dalmago G.A., Capitani D.H.D., Duft D.G., Suárez J.G., Pierozzi J.I., Trevelin L.C., Moreira D.A., 2019. Sustainability assessment of sugarcane production systems: SustenAgro Decision Support System. *Environmental Development*, 32. https://doi.org/10.1016/j.envdev.2019.05.003.

Depetris-Chauvin N., Fernandez Olmos M., Hu W., Malorgio G., 2023. Costs and benefits of sustainability-oriented innovation in the agri-food industry: A review. *New Medit*, 22(3): 23-45.

Duarte C.G., Gaudreau K., Gibson R.B., Malheiros T.F., 2013. Sustainability assessment of sugarcane-ethanol production in Brazil: A case study of a sugarcane mill in São Paulo state. *Ecological indicators*, 30: 119-129. https://doi.org/10.1016/j.ecolind.2013.02.011.

Eizenberg E., Jabareen Y., 2017. Social sustainability: A new conceptual framework. *Sustainability*, 9(1): 68.

FAO, 2013. *Sustainability Assessment of Food and Agricultural Systems (SAFA) Guidelines*. Rome: FAO.

FAO, 2017. *Productivity and Efficiency Measurement in Agriculture*. Rome: FAO.

Fischer R., 2015. Definitions and determination of crop yield, yield gaps, and of rates of change. *Field Crops Research*, 182: 9-18.

Gamboa G., Kovacic Z., Di Masso M., Mingorría S., Gomiero T., Rivera-Ferré M., Giampietro M., 2016. The complexity of food systems: Defining relevant attributes and indicators for the evaluation of food supply chains in Spain. *Sustainability*, 8(6): 515.

Gani A., Asjad M., Talib F., Khan Z.A., Siddiquee A.N., 2021. Identification, ranking and prioritisation of vital environmental sustainability indicators in manufacturing sector using pareto analysis cum best-worst method. *International Journal of Sustainable Engineering*, 14(3): 226-244.

Garcia-Bustamante C.A., Aguilar-Rivera N., Zepeda-Pirron M., Armendariz-Arnez C., 2018. Development of indicators for the sustainability of the sugar industry. *Environmental & Socio-Economic Studies*, 6(4): 22-38. https://doi.org/10.2478/environ-2018-0025.

Gaviglio A., Bertocchi M., Marescotti M. E., Demartini E., Pirani A., 2016. The social pillar of sustainability: a quantitative approach at the farm level. *Agricultural and Food Economics*, 4(1): 1-19.

Gayatri S., Gasso-tortajada V., Vaarst M., 2016. Assessing sustainability of smallholder beef cattle farming in Indonesia: A case study using the FAO SAFA framework. *Journal of Sustainable Development*, 9(3).

Gnansounou E., Alves C.M., Pachón E.R., Vaskan P., 2017. Comparative assessment of selected sugarcane biorefinery-centered systems in Brazil: A multi-criteria method based on sustainability indicators. *Bioresource technology*, 243: 600-610. https://doi.org/10.1016/j.biortech.2017.07.004.

Hale J., Legun K., Campbell H., Carolan M., 2019. Social sustainability indicators as performance. *Geoforum*, 103: 47-55.

Hall R., Scoones I., Tsikata D., 2017. Plantations, outgrowers and commercial farming in Africa: agricultural commercialisation and implications for agrarian change. *The Journal of Peasant Studies*, 44(3): 515-537.

Hassan M.F., Kornher L., 2022. Farm wage and Rice Price dynamics in Bangladesh. *Food Security*, 14(1): 127-146.

Janker J., Mann S., 2020. Understanding the social dimension of sustainability in agriculture: a critical review of sustainability assessment tools. *Environment, Development and Sustainability*, 22(3): 1671-1691.

Janker J., Mann S., Rist S., 2019. Social sustainability in agriculture–A system-based framework. *Journal of Rural Studies*, 65: 32-42.

Joglekar S.N., Dalwankar G., Qureshi N., Mandavgane S.A., 2022. Sugarcane valorization: selection of process routes based on sustainability index. *Environmental Science and Pollution Research*: 1-14.

Karvonen J., Halder P., Kangas J., Leskinen P., 2017. Indicators and tools for assessing sustainability impacts of the forest bioeconomy. *Forest ecosystems*, 4(1): 1-20.

Kralisch D., Ott D., Gericke D., 2015. Rules and benefits of life cycle assessment in green chemical process and synthesis design: a tutorial review. *Green Chemistry*, 17(1): 123-145.

Kristensen H.S., Mosgaard M.A., 2020. A review of micro level indicators for a circular economy–moving away from the three dimensions of sustainability? *Journal of Cleaner Production*, 243: 118531.

Lacovidou E., Voulvoulis N., 2018. A multi-criteria sustainability assessment framework: development and application in comparing two food waste management options using a UK region as a case study. *Environmental Science and Pollution Research*, 25: 35821-35834.

Latruffe L., Diazabakana A., Bockstaller C., Desjeux

Y., Finn J., Kelly E., Uthes S., 2016. Measurement of sustainability in agriculture: a review of indicators. *Studies in Agricultural Economics*, 118(3): 123-130.

Leal M.R.L.V., Nogueira L.A.H., 2014. The sustainability of sugarcane ethanol: The impacts of the production model. *Chemical Engineering Transactions*, 37: 835-840. https://doi.org/10.3303/CET1437140.

Li T., Hua F., Dan S., Zhong Y., Levey C., Song Y., 2020. Reporting quality of systematic review abstracts in operative dentistry: An assessment using the PRISMA for Abstracts guidelines. *Journal of Dentistry*, 102.

Lopez-Ortega M.G., Guadalajara Y., Junqueira T.L., Sampaio I.L.M., Bonomi A., Sanchez A., 2021. Sustainability analysis of bioethanol production in Mexico by a retrofitted sugarcane industry based on the Brazilian expertise. *Energy*, 232(12). https://doi.org/10.1016/j.energy.2021.121056.

Machado P.G., Rampazo N.A.M., Picoli M.C.A., Miranda C.G., Duft D.G., de Jesus K.R.E., 2017. Analysis of socioeconomic and environmental sensitivity of sugarcane cultivation using a Geographic Information System. *Land Use Policy*, 69: 64-74.

Marta-Costa A., Trigo A., Costa J.M., Fragoso R., 2022. Standards and indicators to assess sustainability: the relevance of metrics and inventories. In: Costa J.M., Catarino S., Escalona J.M., Comuzzo P. (eds), *Improving Sustainable Viticulture and Winemaking Practices*. Amsterdam: Elsevier, pp. 391-414.

Massuça J., Marta-Costa A., Lucas M.R., 2023. Social dimension of sustainability: assessment in the agribusiness context. *New Medit*, 22(2): 63-80.

Mokrani A., Ben Nasr J., Sai M.B., Bachta M.S., 2022. Sustainability assessment of Tunisian olive growing systems. *New Medit*, 21(4): 3-13.

Nadaraja D., Lu C., Islam M.M., 2021. The sustainability assessment of plantation agriculture-a systematic review of sustainability indicators. *Sustainable Production and Consumption*, 26: 892-910.

Naisola-Ruiter V., 2022. The Delphi technique: a tutorial. *Research in Hospitality Management*, 12(1): 91-97.

Nwokocha I., 2015. Employers and the enhancement of industrial harmony in private sector organizations in Nigeria. *IOSR Journal of Humanities and Social Science*, 20(5): 28-35.

Pham L.V., Smith C., 2014. Drivers of agricultural sustainability in developing countries: A review. *Environment Systems and Decisions*, 34: 326-341.

Poli F.M.L., Islas-Samperio J.M., Bustamante C.G.A., Rivero J.S.C., Grande-Acosta G.K., Gallardo-Alvarez R.M., Escobedo C.A., 2022. Sustainability Assessment of Solid Biofuels from Agro-Industrial Residues Case of Sugarcane Bagasse in a Mexican Sugar Mill. *Sustainability*, 14(3): 25. https://doi.org/10.3390/su14031711

Prasara-A J., Gheewala S.H., 2016. Sustainability of sugarcane cultivation: case study of selected sites in north-eastern Thailand. *Journal of Cleaner Production*, 134: 613-622. https://doi.org/10.1016/j.jclepro.2015.09.029.

Prasara-A J., Gheewala S.H., 2021. An assessment of social sustainability of sugarcane and cassava cultivation in Thailand. *Sustainable Production and Consumption*, 27: 372-382. https://doi.org/10.1016/j.spc.2020.11.009.

Prasara-A J., Gheewala S.H., Silalertruksa T., Pongpat P., Sawaengsak W., 2019. Environmental and social life cycle assessment to enhance sustainability of sugarcane-based products in Thailand. *Clean Technologies and Environmental Policy*, 21(7): 1447-1458. https://doi.org/10.1007/s10098-019-01715-y.

Rajeev A., Pati R.K., Padhi S.S., Govindan K., 2017. Evolution of sustainability in supply chain management: A literature review. *Journal of Cleaner Production*, 162: 299-314.

Rasmussen L.V., Bierbaum R., Oldekop J.A., Agrawal A., 2017. Bridging the practitioner-researcher divide: Indicators to track environmental, economic, and sociocultural sustainability of agricultural commodity production. *Global environmental change*, 42: 33-46.

Sawaengsak W., Olsen S.I., Hauschild M.Z., Gheewala S.H., 2019. Development of a social impact assessment method and application to a case study of sugarcane, sugar, and ethanol in Thailand. *International Journal of Life Cycle Assessment*, 24(11): 2054-2072. https://doi.org/10.1007/s11367-019-01624-8.

Shi L., Han L., Yang F., Gao L., 2019. The evolution of sustainable development theory: Types, goals, and research prospects. *Sustainability*, 11(24): 7158.

Shukor S.A., Ng G.K., 2022. Environmental indicators for sustainability assessment in edible oil processing industry based on Delphi Method. *Cleaner Engineering and Technology*, 10: 100558.

Sidhoum A.A., 2018. Valuing social sustainability in agriculture: An approach based on social outputs' shadow prices. *Journal of Cleaner Production*, 203: 273-286.

Silalertruksa T., Gheewala S.H., Pongpat P., 2015. Sustainability assessment of sugarcane biorefinery and molasses ethanol production in Thailand using

eco-efficiency indicator. *Applied Energy*, 160: 603-609. https://doi.org/10.1016/j.apenergy.2015.08.087.

Silalertruksa T., Pongpat P., Gheewala S.H., 2017. Life cycle assessment for enhancing environmental sustainability of sugarcane biorefinery in Thailand. *Journal of Cleaner Production*, 140: 906-913. https://doi.org/10.1016/j.jclepro.2016.06.010.

Souto S., Bezerra R., Gonçalves S., Santos M., Gurgel M., 2024. Conflicts and socioenvironmental injustice in the Acaú-Goiana Extractive Reserve. *Ambiente & Sociedade*, 27: e02253.

Stander M.W., De Coning J.A., Rothmann S., 2019. Do wage and wage satisfaction compensate for the effects of a dissatisfying job on life satisfaction? *SA - Journal of Industrial Psychology*, 45(1): 1-11.

TraceX Technologies, 2023. *Measuring farm sustainability metrics for better agriculture*. TraceX Technologies. Retrieved from https://tracextech.com/measuring-farm-sustainability-metrics/ (acessed at 31.01.2024).

Turetta A.P.D., Kuyper T., Malheiros T.F., Coutinho H.L.D.C., 2017. A framework proposal for sustainability assessment of sugarcane in Brazil. *Land Use Policy*, 68: 597-603. https://doi.org/10.1016/j.landusepol.2017.08.011.

Van Eijck J., Romijn H., Smeets E., Bailis R., Rooijakkers M., Hooijkaas N., Faaij A., 2014. Comparative analysis of key socio-economic and environmental impacts of smallholder and plantation based jatropha biofuel production systems in Tanzania. *Biomass and Bioenergy*, 61: 25-45.

Waas T., Hugé J., Block T., Wright T., Benitez-Capistros F., Verbruggen A., 2014. Sustainability assessment and indicators: Tools in a decision-making strategy for sustainable development. *Sustainability*, 6(9): 5512-5534.

World Commission on Environment and Development, 1987. *Our Common Future*. Oxford: Oxford University Press, pp. 372.

Yani M., Machfud S., Asrol M., Hambali E., Papilo P., Mursidah S., Marimin M., 2022. An Adaptive Fuzzy Multi-Criteria Model for Sustainability Assessment of Sugarcane Agroindustry Supply Chain. *Ieee Access*, 10: 5497-5517. https://doi.org/10.1109/access.2022.3140519.

Zhao P., Ali Z.M., Ahmad Y., 2023. Developing indicators for sustainable urban regeneration in historic urban areas: Delphi method and analytic hierarchy process (AHP). *Sustainable Cities and Society*, 99: 104990.

Cooperative branded food purchasing behaviour: Understand the role of consumer social responsibility, trust, and agri-rural ties through Theory of Planned Behaviour

Gülçin Koç*, Cennet Oğuz**

DOI: 10.30682/nm2404i
JEL codes: D12, Q13

Abstract

The extended version of "Theory of Planned Behaviour"-TPB have extensively used in food consumer research during last two-decade, but extension the TPB with consumers' Social Responsibility (SR) and its use on cooperatives' branded food (collective brand) is missing in the literature, therefore this study aims to fill these gaps in the literature and to determine whether consumers' SR, trust and agri-rural interest play role in purchasing the collective brand. The data was gathered from sample consisting of 284 persons who is responsible for household food expenditure, with 18+ age and living in the central districts of Antalya province in the autumn of 2021. Structural Equation Model (SEM) is used to test whether the E-TPB is valid. The results confirmed that SR and agri-rural relations are important variables in consumers' decision to purchase collective brand. SR is also a mediating variable between intention and behaviour and between trust and behaviour. The results imply that consumers' purchasing decisions for the collective brand can be improved by increasing consumers' SR levels. As a result, public relation, communication, and advertising programs focused on raising consumers' awareness of SR towards agri-rural development can also support consumers' purchase of this brand.

Keywords*: Theory of Planned Behaviour, Collective food brand, Food consumer behaviour.*

1. Introduction

Food supply chain and consumer purchasing behavior have undergone considerable changes since the 1990s in Türkiye. Increase in consumer purchasing power, changes in demographic structure, urbanisation and accompanying lifestyles changes, and recognizing the relationship between diet and health outcomes can be counted as the main demand side drivers for food (Béné *et al*., 2020). Technological innovation and intensification (homogenisation) in production considered as major supply side drivers of food system (Béné *et al*., 2020). In worldwide, demographic factors, urbanisation, food industry marketing mix and trade liberalisation can be considered divers affecting shift in dietary patters and nutritional transition associated with increasing obesity and cardiovascular diseases (Kearney, 2010). Consumers have been increasingly interested about what they eat, how it is produced and the environmental impact of pro-

* Researcher (PhD), not affiliated.
** Selçuk University, Department of Agricultural Economics, Faculty of Agriculture, Konya, Türkiye.
Corresponding author: mckoc2004@hotmail.com

duction and consumption (Lappo *et al*., 2015). It was expected that food consumption in future will be effected by food quality-safety and health outcomes, social responsibility, methods of production and innovation, sustainability indicators, and origin (Lappo *et al*., 2015). As a matter of fact, sales of organic food, geographical indication-GIs food and fair-trade food products has exhibited an increasing trend in developed countries, particularly EU countries (Chilla *et al*., 2020; Kyrylov *et al*., 2018; Ruggeri & Corsi, 2021).

During last decade, popularity of agricultural cooperatives have increased in food sector (Qorri & Felföldi, 2024). In Europe, reported number of agricultural cooperatives are more than 400 thousand having 9 million farmer members and employing 600 thousand peoples (Ajates, 2020). In the EU, market share of agricultural cooperatives in food sales was 40% in 2010, of which dairy products (33%) and vegetable and fruits (27%) took the major shares (EC, 2012).

The share of agricultural cooperative in food sales exceed 50% in Austria, Denmark, Finland, France, Ireland, The Netherlands (70%) and Sweden. The market share is differ considerably with respect to sector (olive oil in Spain 75%, dairy in USA 75%) and countries such as 70% in the Netherlands (Candemir *et al*., 2021). As of 2021, cumulative revenue of the first-biggest 100 agricultural cooperatives in Europe exceed €255 billion which is up €17 billion from 2020 (Frey, 2023). In China, number of agri-cooperatives' members reported as 2.21 million in 2019 representing about 50% of whole farmers and higher than 50% of them are interested in fruits and vegetable growing industry (Wang *et al*., 2021).

In Türkiye, number of agricultural cooperatives are 11,982 (with about 3,931 thousand members) consisting of 53.73% agricultural development (multi-purpose cooperatives), 20.54% irrigation, 13.7% agricultural credit (ACC), 6.95% agricultural sales and marketing, 4.67% aquaculture, 0.26% sugar beet growers (31 cooperative) and 0.31% fresh fruits and vegetable cooperatives (Pakdemirli, 2019). While sugar beet grower's cooperative is the smallest one considering number of cooperatives, but it constitutes 35.86% of total cooperative members. Agriculture credit cooperatives (23.08%), sales and marketing cooperatives (14.05%) constituted second and third biggest one in terms of members.

Agricultural cooperatives have long been an important player in the food supply chain in Türkiye, both mid-stream with food processing plant and downstream level operation such as retailing and export marketing. Agricultural sales cooperatives' unions (ASCUs) such as Tariş Unions (olive and olive oil, olive paste, fig and fig based confectionery product, sultana raisin, vinegar, tahini and grape molasses, sauces and turnips, cosmetics and personal care products), Marmarabirlik (olive), Trakyabirlik and Karadenizbirlik (corn oil, refined sunflower oil and margarine), and Fiskobirlik (hazelnuts, hazelnut based confectionery products and honey) has been selling own-branded consumer food products. Besides distribution of their food product through dealer, some of the ASCUs has their own sales shops in their vicinity areas and use online sales channels to reach consumers.

During last decade, ACC entered food retail market operation in 2017, with medium-size hard discount format and more recently small grocery format and reached about 4,000 outlets at the end of 2023. Besides its conventional credit cooperative function, ACC can be regarded as an agribusiness group, with 1,805 service points, 17 regional unions, 800 thousand members and 15 subsidiary companies. ACC subsidiary companies are operating in area of input production facilities, mechanisation procurement service, food processing facilities, livestock production, contract farming, logistics service, insurance services, retirement service, licenced storage facilities, logistics, technology, and food retail market chain (TKK, 2023). In addition to the ACC, several agricultural development cooperatives have emerged in food processing and marketing operation during last two decade. Tire-Süt Agricultural Development Cooperative in Tire District in İzmir province (Tire-Dairy Coop) in Western Anatolia (https://tiresutkoop.org.tr) and Ovacık-94 Neighbourhood Agricultural Development Cooperative in Tunceli province in Eastern Anatolia (https://ovacikdogal.com) has recognised as successful cooperative examples in food supply chain. Tire-Dairy Coop has own

dairy and meat processing facilities, process raw milk and meat into cooperative branded dairy and meat products (under Tire-Süt brand), deliver these branded products to 3,378 dealers in the countrywide (including supermarket chains) and use online marketing channel. Ovacık Koop has 23 grocery outlets with its own brand called Ovacık Doğal (Naturel) in 13 different provinces and use also online sales channel. The products mix of this development cooperative includes honey, bee-milk and pollen, pulses (chickpea and beans) and roasted chickpea, walnut, jam, grape molasses, and salt. Another cooperative movement in food marketing during the last two decade is sugar beet grower's cooperative namely Pankobirlik which is a union of 16 regional beet grower cooperatives with about 900 thousand active members (Pankobirlik, 2024). While the cooperative has been an important player in sugar manufacturing since 1954 with Konya Sugar Factory, it has taken a step further in the food sector with chocolate products under Torku brand name in 2007. The Torku brand has extended its product ranges which varies from bakery products to chocolate, candy, frozen potatoes to processed product of dairy and meat, cereal product to fruit juice and vinegar to turnip juice (Pankobirlik, 2024). Currently, Torku branded products are available for consumers in many leading food retail chains as well as their own retail outlet located in 6 provinces with 45 outlets including Konya (27), Karaman (2), Ankara (16) and İstanbul (2). Besides first food grocery supermarket opened in 2013, Torku also entered in food catering sector with Torku Doğrudan (direct) Döner Restaurant in 2018.

In recent years, agricultural cooperatives owned branded foods in Türkiye have gained importance against to both retailer brands (private label) and manufacturer brands. In this context, the reminder "we are a cooperative brand" has been widely used in product advertising by agricultural cooperatives. A cooperative is currently marketing its product under the brand name "cooperative honey". In recent years, the visibility of agricultural cooperatives' products (branded) has increased in supermarket chains. Some chain stores use cooperative branded products especially as a product differentiation tool. Buying directly from producer cooperatives, enables retailers to buy the product at a more affordable price, solves the quality problem, and obtain other procurement advantages (i.e., volume, regularity). However, it was found that cooperatives in fresh produce (vegetable) supply chain are important organisation with respect to achieving coordination in supply chain activities including credit cooperation (Wang *et al*., 2021). Food retailer chains purchase from any producer regardless of their production scale if they can meet the specified requirements including quality and safety, volume and all year around supply or delivering capacity in pick season (Oparebea Boateng *et al*., 2023). The production and process requirements dictated by retailers implies investment and particular good agricultural production practices, which is a constraint for small-scale producer, particularly in developing countries. Inherent solidarity among farmers under cooperative organisation enables producers to collectively cop with perceived market risk and motivates common investment by sharing fixed cost. Moreover, provide incentives to change practices of its members and offer a stronger market position (Candemir *et al*., 2021). Members trust and support to cooperative management is important in sustainability of cooperatives success. In this context, cooperative management should proceed inclusive decision-making process (Kinikli & Yercan, 2023) and also marketing strategy based on evidence.

From the point of view of consumer-oriented marketing, it is important to understand consumer behavior in terms of marketing mix: product, price, distribution, and communication policies. In this context, understanding insight of consumers' cooperative branded food purchasing based on consumer behavior theories and obtaining results relying on reliable statistical analyses are important information for decision makers to understand the social-psychological factors and social responsibility affecting consumer purchase intentions and behaviours.

In the literature on food consumer behaviour study, it has not found a study that focuses on agricultural cooperatives' food brands purchasing decision of consumer and extending theory of planned behavior (TPB) with the dimension

of social responsibility perspective or general social issue component of socially responsible consumer. Excluding social responsibility perspective in consumer behaviour research with respect to agricultural cooperative branded food products context can be regarded as an important gap in the literature, as cooperatives are the most important producer organizations/social enterprises and their vertical integration is important to improve the position of farmers in the highly concentrated food supply chain in midstream and downstream level, especially at the retail level. In addition, socially responsible consumers have been emerging segment in society, particularly developed and upper-middle income countries during recent years. This study aimed to fill the gap in the literature on food consumer behavior research at the national level with using TPB. At the international level, it will enrich the relevant literature by using the method in a different context and extend the basic standard model with Social Responsibility perspective of consumer and the level of agri-rural relationship of consumers.

2. Literature Review

During last two decades, theory of planned behavior (Ajzen, 1985; 1991) relying on social psychology theory has been widely used in food consumer behavior research (Ajzen, 2016; Alam *et al*., 2020; Chang *et al*., 2019; Chen, 2017; Donahue, 2017; Essakkat *et al*., 2021; Giampietri *et al*., 2018; Kim & Kuo, 2022; Maichum *et al*., 2017; Paul *et al*., 2016; Qi & Ploeger, 2019; Raygor, 2016; Shah Alam & Mohamed Sayuti, 2011; Sogari *et al*., 2023; Tommasetti *et al*., 2018; Vermeir & Verbeke, 2008; Yadav & Pathak, 2016a, 2016b; Yousuf *et al*., 2019; Zhang *et al*., 2019). This model has been one of the widely used and the most influential empirical framework to predict reasoned consumer behaviour (Pandey *et al*., 2021).

When the literature is examined, it can be observed that TPB has been employed in different product context and countries to explain consumer behavior, particularly certified products such as organic and green products. In the literature, it is emphasized that TPB gives reliable results in understanding consumer behavior (Giampietri *et al*., 2018; Rozenkowska, 2023). Although TPB structure has used in different fields of social sciences (such as communication, sociology, agricultural economics, business administration) in Türkiye (Aktulay Çakır, 2014; Demirtaş, 2017; Mercan, 2015; Sığındı & Kavak, 2015; Sözüer *et al*., 2015; Taşçı-Duran, 2016), but, it has only recently used in very limited study focused on food consumer behavior. Furthermore, it was not found any study in international literature focused on consumer behaviour with respect to agricultural cooperatives' branded products and the standard model has not also been extended with Social Responsibility perspective of consumer in the reviewed literature. In the literature, general social issues component of socially responsible consumer has neglected (Han & Stoel, 2017). There are quite a few studies included social dimension only partially within TPB framework in their study (Fleșeriu *et al*., 2020; Kim, 2014). Majority of the studies based on TPB framework in the literature have added a new dimensions or constructs into basic model either directly or as a mediating variable. Articles published during last decade were accessed through google scholar search using key terms as "extended-TPB in food consumer behaviour". The variables used in the extension of TPB is presented below. The variables used in the extended model includes confidence and values (Vermeir & Verbeke, 2008), positive moral attitude, ethical self-identity and food choices (Dowd & Burke, 2013), personal background factors (Menozzi *et al*., 2015), past behaviour (Raygor, 2016), ethnocentrism and collectivist behavior (Vabø & Hansen, 2016), personal norm depicted by "sense of obligation to take action" based on value, belief and norm theory (Hoeksma *et al*., 2017), product labels such as organic, green, ethical, geographic-GIs/PDO (Setyawan *et al*., 2018; Giampietri *et al*., 2018; Zhang *et al*., 2019), perceived usefulness and curiosity (Tommasetti *et al*., 2018), perceived benefit and risk affected by trust (Zhang *et al*., 2018), moral obligation and self-identity moderated by gender (Beldad & Hegner, 2018), trust, habits and behavioural beliefs (Spence *et al*., 2018), self-idendity and ecological motives (Zhu, 2018), trust, corporate social responsibility perception,

concern about GMO foods, and consumer promotion and prevention focus (Akbari *et al*., 2019), personal characteristics and confidence (Qi & Ploeger, 2019), trust, past behaviour, and green self-identity (Carfora *et al*., 2019), ethnocentrism (Maksan *et al*., 2019; Miguel *et al*., 2022), perceived health and monetary value affecting attitude formation (Fiandari *et al*., 2019), health consciousness and gender as mediator variable with rest of construct variables (Shin *et al*., 2020), trust to organic product (Canova *et al*., 2020), health consciousness (Fleşeriu *et al*., 2020; Rahamat *et al*., 2022), perceived communication, satisfaction and trust as a moderating variables (Sultan *et al*., 2020), alcohol-identity as a moderating variable between norms (injunctive and descriptive), attitudes and perceived behavioural control variables and beer purchase intention (Wang, 2020), price sensitivity (Wang *et al*., 2020), perceived benefit: health, sustainability and price (Dorce *et al*., 2021), knowledge, perceived barriers and sensory attributes (Pandey *et al*., 2021), experience and behavioural intention (Bae & Choi, 2021), environmental concern (Auza & Mouloudj, 2021; Fleşeriu *et al*., 2020), perceived severity, vulnerability, fear, rewards, efficacy and response cost variables corresponding to the protection motivation theory (Pang *et al*., 2021), trust, environmental issues and habits (Dionysis *et al*., 2022), well-being perception variables (D'Souza, 2022), trust (organisational, product, interpersonal, chain), knowledge (subjective and objective), uncertainty and past behaviour (Loera *et al*., 2022), trust, face consciousness and policy support (Ding *et al*., 2022), product availability, product quality, health concern and environmental concern (Teixeira *et al*., 2021), self-determination theory (intrinsic and extrinsic motivation), trust and price consciousness (Khan *et al*., 2023), marketing mix as a direct determinant of intention and also as a moderating variable between attitudes, norms and perceived behavioural control construct and intention (Farid *et al*., 2023). The results of studies indicated that additional constructs included in the TPB for extending to the model is generally contributed the robustness and predictive capability of the theoretical model (Dionysis *et al*., 2022).

Among the cited literature in this study, it has not found a study that focuses on agricultural cooperatives' food brands purchasing decision of consumer and extending TPB with the dimension of social responsibility perspective or general social issue component of socially responsible consumer. This can be considered an important gap in the literature, as cooperatives are the most important producer organizations/ social enterprises and their vertical integration is important to improve the position of farmers in the highly concentrated food supply chain in midstream and downstream level, especially at the retail level.

Today, there are consumers called socially responsible who act with a social responsibility motive, such as those who have environmental, health, religious-cultural and ethical concerns. Consumers who prefer cooperatives' branded products can also act with motives such as supporting cooperatives as social enterprises and contributing to rural development. Socially responsible consumer (SRC) is as an individual who takes into account of the externalities of his/her private consumption with respect to environmental and general social concerns (Han & Stoel, 2017). It was criticized that treating SRC as equivalent to ecologically or environmentally friendly behavior of consumers and portraying consumer profile and strategies only based on environmental issues. In the literature, importance of taking into account of ethical issues including both environment and general social issues was emphasized (Han & Stoel, 2017).

3. Methodology

3.1. *Theoretical Model*

The TPB has been widely used theoretical model in consumer research for last decade. The extended version the TPB provided an uncountable amount of valuable contribution to the literature that will be informative and beneficial for advancing the theory and consumer behaviour researches (Rozenkowska, 2023). The variables collectively lead to construct of individual behavioural intention in the TPB are: (1) attitude, (2) subjective/social norms, and (3) perceived behavioural control (Ajzen, 1985; 1991; 2016).

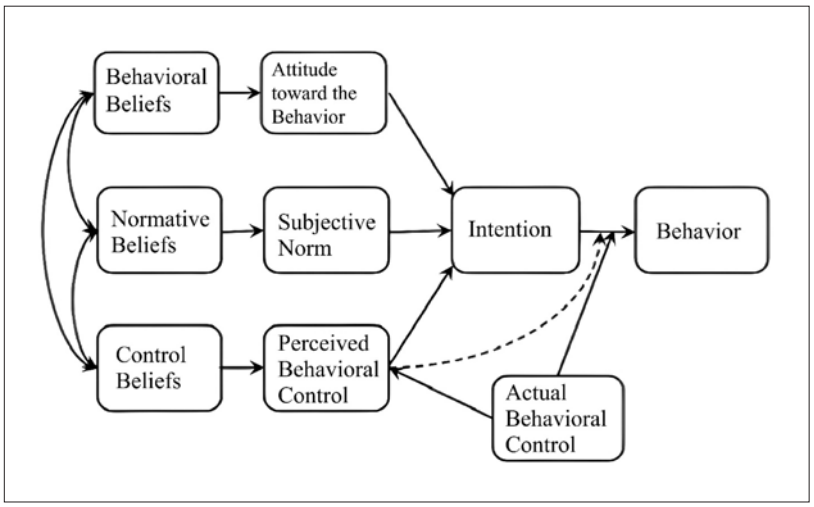


Figure 1 - Theory of Planned Behaviour.

Source: Ajzen, 1985, 1991, 2016.

In literature, attitude is defined as a degree to which an individual evaluation (positive or negative) or evaluation of an intention and behavior in a given context. Subjective norms are related to the perceived social pressure that plays an important role in performing the behavior. Perceived behavioural control represent whether behavioural performance is easy, difficult and under individual control (Qi & Ploeger, 2019). The TPB has been applied in a wide range of consumer food choice/ purchasing decision to reveal consumer intention and behavior.

According to the theory shown in Figure 1, consumer behaviour in question can be derived from the intention to perform the relevant behavior in a given context. The intention is a function of latent variables consist of attitude towards behavior, social norms, and perceived behavioural control. Each of these latent variables are formed by beliefs with respect to behavioural, normative and control, respectively. Behavioral beliefs and evaluation of behavioural outcomes affect the attitude towards behavior, normative beliefs, and the motivation to adapt to these norms, subjective norms and control beliefs, and perceived power affects perceived behavioural control (Raygor, 2016). TPB model encompass individual, social, and behavioural aspects all in one which is different from variety of the other model. It has conceptually three independent determinants consisting of attitudes, subjective norms, and perceived behavioural control of persons' intention to do a behaviour (Vabø & Hansen, 2016).

The basic assumption of the theory is that intention is a predecessor of behavior and jointly with perceived behavioural control determines to behaviour in question. The stronger these two determinants are, the more accurate the behavioural performance will be. As seen in the Figure 1, consumers' intention is determined by a combination of attitude toward behavior, subjective norms, and perceived behavioural control in relation to the behavior in question, and these are influenced by behavioural, normative, and control beliefs, respectively. The more favourable attitude and subjective norm, the greater the perceived behavioural control, the greater the probability of the consumer's intention to engage in the relevant behavior. In the literature, past behavior and self-identity has considered as additional determinants of intention within the original TPB (Giampietri *et al*., 2018). The TPB states that the intention to perform a behavior is formed by the latent variables (Szejda *et al*., 2020):

i) Attitudes (positive or negative feelings of the person towards the behavior)

ii) Subjective norms (important values and perceived social pressure from people)

iii) Perceived behavioural control (self-efficacy or whether the behavior is believed to be under one's own control).

As aforementioned, actual behavior of indi-

vidual is projected by combination of actual intention and actual behavioural control. Behavioral control is conceptualized as a component of self-sufficiency (internal factors) and controllability (external factors). This also explains why intentions do not always lead to behavior change.

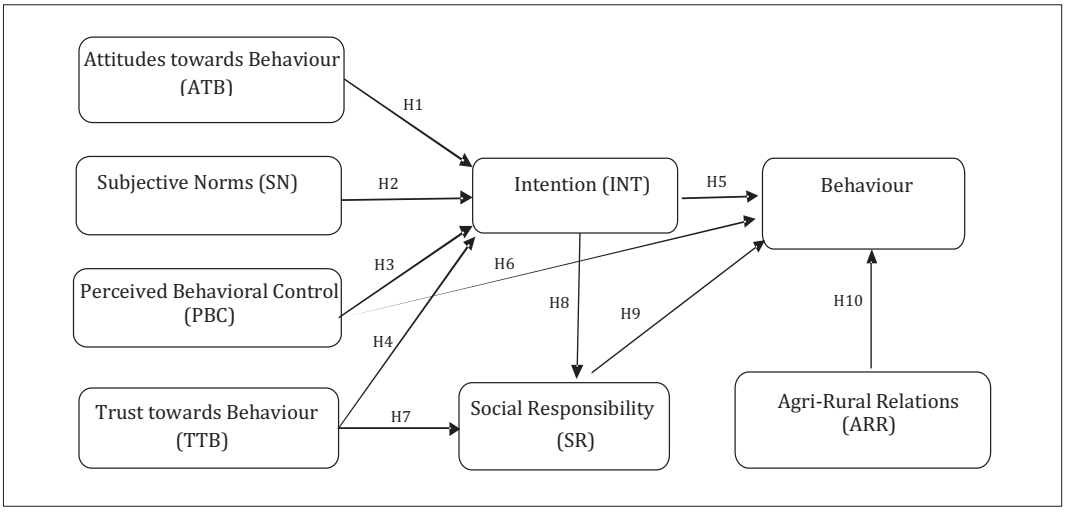
The extended-TPB structure used in this research is shown in Figure 2. The latent variables in the basic TPB are the variables consisted of attitude towards behavior (ATB), subjective norms (SN), and perceived behavioural control (PBC). In the basic model, it is accepted that these variables affect intention (INT) and intention affects to behavior. In the basic model, perceived behavioural control (PBC), one of the latent variables, is the variable that affects both the intention and the behavior directly or indirectly as a mediating factor between intension and behaviour.

As can be seen in Figure 2, it is assumed that the variable of trust towards behavior (TTB) in the extended-TPB can affect behavior both directly through intention and through social responsibility (SR) variable. Lack of consumer trust affect to consumer behavior both directly and indirectly. The indirect impact occurs through less favourable beliefs that causes less favourable attitudes and intentions. The direct affects occur by reducing consumers' propensity to purchase the product even if consumers indent to do so (Loera *et al*., 2022).

In this study, TPB is extended with SR, trust and the ARR constructs, which represents social responsibility, trust, and the agricultural-rural relationship level of consumers respectively, were considered as variables that directly affect the behavior. In a study conducted in Italy (Giampietri *et al*., 2018), the variable of trust towards behavior (with single item construct), fair-trade certified product consumption habit and the residential areas of consumers (urban/rural) were included in the extended-TPB. Trust was also used in the TPB as a moderating variable between intention and buying, and also between PBC and buying (Sultan *et al*., 2020). The ARR variable, reflects rural ties and consumer involvement in agriculture because of food security issues and supporting traditional cultural values, can be considered as a proxy for ethnocentric behaviour.

Social dimension of consumer buying behaviour was only included in the study conducted in Romania (Fleșeriu *et al*., 2020). The author extended the TPB with a variable called as "social consciousness" in their consumer organic food behavior study. They reasoned that besides tangible attributes, consumers purchasing decision is also affected by intangible attributes such as contributing local community development and keeping traditions alive. Because of this reason, consumers exhibit ethnocentric behaviour and to prefer domestic product against to imported

Figure 2 - Extend-Theory of Planned Behaviour.



products. The inclusion of the variable called SR in the TPB framework in this study is based on assumption that there are significant numbers of consumers acting with the motivation of supporting rural areas and agricultural cooperatives due to various reasons (i.e., acting as anti-capitalist or supporting farmer's collective actions, contribution to rural development, supporting short food supply chain etc.).

In recent years, success of various types of agricultural cooperatives in food supply chain, particularly, mid-and-down-stream part of the chain in Türkiye such as agricultural development, agricultural sales and agricultural credit may be influenced by the motivation of consumers to realize their social responsibilities towards agriculture/rural areas.

Structural equation model (SEM) was used to determine whether the socio-psychological variables in the theoretical model, as well as the social responsibility behaviours of consumers and the level of agro-rural relationship, influence the consumer purchasing decision. The model was used to test the validity of the H0 hypothesis that there is no relationship between consumer purchasing decision (behavior) and attitude towards behavior, subjective norm, perceived behavioural control, social responsibility perspective and consumer agricultural-rural relationship level. If the H0 hypotheses are not valid, the results will answer the questions of how much each variable explains the behavior. The hypothesis of the search is given in the Figure 2. The direction of the arrows indicates causal relationship, for instance social norm effects intention and intension effect behaviour.

3.2. *Data*

The data employed in this study was collected through a face-to-face interview with 284 adult persons (18+ years old) who are responsible for their household food purchase in three central districts of Antalya province (namely Kepez, Muratpaşa and Konyaaltı) in Türkiye. These three districts constituted almost 50 percent of province population as of 2021 (TurkStat, 2022). Currently, in terms of population size, Antalya is ranked first-fifth across 81 provinces in Türkiye. The province is keeping high population growth rate and becoming multicultural metropolitan area with new citizens from different countries (TurkStat, 2020). The sample size was calculated by using equation (1) below (Cochran, 1977; Oğuz & Karakayacı, 2017; Yamane, 2001).

$$n = \frac{\mathrm{N}*p*(1-p)}{(N-1)*\mathrm{D}^2+p*(1-p)} = \quad (1)$$

$$= \frac{405304*0.5*(1-0.5)}{(405304-1)*(0.0304)\text{^}2+0.5*(1-0.5)} = 270$$

In equation (1), lower case n stands for sample size, capital N stands for total household numbers in central districts (N = 405304 household), p represents ratio of consumer regularly and frequently purchasing cooperative branded food products in population (assumed 0.5), $D^2 = (d/t)$ represent varyans, d represents permitted error (0.5), and t stands for assumed confidence level (90%, t=1,65). In the field study, five percent more respondent were interviewed than the calculated sample size, so total 284 fully completed questionnaires were achieved. The survey study was implemented during October-December 2021 period[1]. The instrument used in data gathering includes questions on socioeconomic and demographic characteristic of respondents, frequency of purchasing cooperatives' branded products (behaviour construct variable: regularly, seldomly, only at discounted price, non-purchased, and not available in the store where I regularly shopping) and scaled questions involved in the E-TPB (See Table 2). A Semantic scale with 7 point was used to obtain rating score of consumers on dimension of variables are given in the model (Figure 2). Following question was asked to respondent to obtain their ratings on dimensions of the variables. Please mark your opinion on the following statements about purchasing

[1] It was the period that COVID-19 pandemic lockdown was substantially removed and partially allowed normal life including school attendance, business opening and work life, but aligning with social distance and protective measures and hygiene conditionalities. The questionnaire was implemented at a convenience place just outside of entrance of different supermarket chain outlet located at different part of the districts considered.

cooperatives' branded food products as [1] indicates "I strongly disagree" while [7] indicates "I strongly agree" opinion of respondent.

4. Research results and discussion

4.1. *Descriptive Statistics*

Descriptive statistics with respect to respondent characteristics of survey data is reported in Table1. As seen in the Table 1, average age of respondent is 36.5 year, 59.5% of respondents are female, 54.5% of respondents are married, and 35.2% of respondents are single. Respondents are highly educated persons (almost 50% is university graduated, 33.5% high school graduated), 71.8% is working age population and almost one-third of respondent is professional occupation and one-quarter is unqualified service workers. Respondents purchasing frequency of cooperatives' food brand is found as 13%, 51.1%, 13%, 16.5%, and 3.5% respectively for purchase of regularly, seldomly, only at discounted price, non-purchased, and not available in the store where I regularly shopping. These results indicate that majority of consumers are familiar with cooperatives' branded food products.

Mean score corresponding to the dimension of latent variables of extended-TPB is reported in Table 2. The scale used in measuring respondent response was ranged from 1 to 7 with higher score corresponding more positive statements about the constructs and vice versa. The results are closer to positive-end since mean scores are greater than mid-point (3.5).

In deciding whether the data are suitable for SEM, in addition to mean scores, it is customary to check correlations between the data pairs and the Cronbach's Alpha value (Sogari *et al*., 2023). The correlation matrix of the latent construct is given in Table 3 shows that there is a statistically significant and positive relationship between the variables used in SEM. Cronbach's Alpha value was computed to confirm whether the latent variables are formed from observed variables (variables given in Table 2) are suitable for the analysis. As seen in Table 3, the Cronbach's Alpha values of the latent variables are relatively high (lowest 0.81 and highest 0.97) which confirms

Table 1. Demographic and Socioeconomic Characteristics of Survey Respondents.

Gender	*Frequency*	*%*
Female	169	59,5
Male	115	40,5
Marital status of respondents		
Married	155	54,6
Single	100	35,2
Divorced	16	5,6
Others (married lives apart, wife/husband died)	13	4,6
Education attained of respondents		
Primary School (or <)	39	13,7
Secondary	7	2,5
High school	95	33,5
Vocational School (or associate degree)	14	4,9
Faculty Graduate	112	39,4
Postgraduate (4.2% in master's degree)	17	6,2
Distribution of age among household members		
0-6 (infancy)	54	8,6
7-12 (childhood)	45	7,1
13-18 (adolescent/young)	66	10,5
19-64 (working population)	453	71,8
65+ (aging population)	13	2,1
Average age (year) and std. deviation	*36,5*	*11,8*
Employment status of respondent		
Public Servant	67	23,6
Working in the private sector (2.5% in part-time workers)	107	37,7
Works at own business or works	38	13,4
Agriculture producer/ employee	2	0,7
Retired	18	6,3
Unemployed looking for work	9	3,2
Unemployed not looking for a job	13	4,6
Students	18	6,3
Other (i.e., housewife)	12	4,2
Profession of respondent		
Executive/Manager	19	6,7
Office workers	35	12,3
Artist etc. service worker	8	2,8
Professional occupation	96	33,8
Service and salesperson	44	15,5
Technician, machine operator and assembler	6	2,2
Agriculture, Forestry, Fishing etc. service worker	4	1,4
Unqualified service workers	72	25,4

Table 2 - Average Score Measured by Semantic Scale on Dimension of E-TPB Model Variable.

7-Point Semantic Scale: [1] "I strongly disagree"... and [7] "I strongly agree"	*Av.*	*S.dev.*
1 - If I buy cooperative branded food products, it gives me happiness (ATB)	4.43	1.65
2 - If I buy cooperative branded food products, it gives me a feeling of solidarity with the producer (ATB)	4.98	1.65
3 - If I buy cooperative branded food products, it gives me a sense of quality and healthy product consumption (ATB)	4.90	1.58
4 - Most people who are important to me want me to buy cooperative branded food products (SN)	4.36	1.76
5 - Most people who are important to me approve of purchasing cooperative branded food products (SN)	4.59	1.72
6 - Most people who are important to me think that I should buy cooperative branded food products (SN)	4.34	1.82
7 - It is easy for me to buy cooperative branded food products (ease of purchase) (PBC)	4.65	1.68
8 - I can buy cooperative branded food products if I want (very easily accessible) (PBC)	4.95	1.7
9 - It is entirely up to me (I can buy) whether to buy cooperative branded products (PBC)	5.39	1.71
10 - I am thinking of purchasing cooperative branded products in the next shopping. (INT)	4.85	1.56
11 - I plan to buy cooperative branded products in the next shopping (INT)	4.85	1.55
12 - I am ready to buy cooperative branded products in the next shopping (INT)	4.80	1.66
13 - I think cooperative branded products are safe (TTB)	5.28	1.56
14 - Cooperative branded products seem reliable to me (TTB)	5.31	1.58
15 - I trust cooperative branded products (TTB)	5.31	1.59
16 - If I buy cooperative branded products, farmers become stronger against intermediaries (SR)	5.40	1.58
17 - If I buy cooperative branded products, I will contribute to agriculture and rural development (SR)	5.36	1.63
18 - If I buy cooperative branded products, I will provide social support to the farmers (SR)	5.35	1.6
19 - If I buy cooperative branded products, I will support the strengthening of cooperatives (SR)	5.47	1.57
20 - If I buy cooperative branded products, I protect the exploited (farmers and consumers) against the capitalist system (SR)	5.32	1.73
21 - If I buy cooperative branded products, I will fulfil my social responsibility towards agriculture and rural development (SR)	5.37	1.65

the scale validity of the variables. These results confirm that the scores given to the observed variables used to obtain each latent variable are internally consistent.

4.2. *Confirmatory Factor Analysis*

The overall fit of the measurement model was assessed by confirmatory factor analysis (CFA). The composite reliability (CR), average variance extracted (AVE) and discriminant validity are commonly used indexes for the construct validity. In this study, as seen in Table 4, the composite reliability of all six constructs is varied between 0.812 to 0.966, which exceeds the minimum reference threshold of 0.70 suggesting internal consistency of multiple items for each construct is adequate. The measurement model was also assessed using the AVE index to determine whether each construct satisfies convergent validity. The AVEs of all constructs are higher than the 0.5 threshold indicating convergent validity of the all the constructs. Comparing the AVE values with the squared correlations among constructs is used to verify discriminant validity of the construct. All the squared correla-

Table 3 - Correlation relationship between implicit variables and scale validity.

	ATB	*SN*	*PBC*	*TTB*	*INT*	*SR*	*ARR*
ATB	1						
SN	0,706**	1					
PBC	0,499**	0,492**	1				
TTB	0,717**	0,679**	0,540**	1			
INT	0,699**	0,666**	0,574**	0,747**	1		
SR	0,664**	0,541**	0,446**	0,798**	0,622**	1	
ARR	0,199**	0,118*	0,101	0,189**	0,191**	0,170**	1
Cronbach's Alfa	0,92	0,93	0,81	0,93	0,97	0,96	

*Asterisk ** and * indicates that coefficients are statistically significant at the 5% and 10% significance level.*

tions of the involved constructs were found less than the value of AVEs, confirming discriminant validity of the constructs.

In SEM analysis, confirmatory factor analysis (CFA) is used to overall fit of the measurement model. The result of CFA in this study is given below in the Table 5. The factor loads obtained by CFA have the lowest value of 0.61 and the highest value of 0.97. It has been generally reported that the factor loads of the variables observed in the CFA should be at least 0.70 (Dorce *et al*., 2021; Rahamat *et al*., 2022). The factor load of the PBC3 variable (one of the dimensions of perceived behavioural control) was 0.61 that falls below commonly agreed threshold value of 0.70. In the literature, it was reported that a factor loading value higher than 0.50 for an item is significant (Beldad & Hegner, 2018). The values between 0.6 and 0.7 were also considered adequate (Dorce *et al*., 2021; Hair *et al*., 2010). Factor loadings of all observed variables except PBC3 are higher than the commonly accepted critical value and all but except PBC3 are in the range of 0.85-0.97 points. The commonly used statistics for model goodness of fit is the Chi-square statistic (CMIN/DF statistic), comparative fit index (CFI) and root mean square error (RMSE) statistics (Beldad & Hegner, 2018; D'Souza, 2022; Essakkat *et al*., 2021; Farid *et al*., 2023; Giampietri *et al*., 2018; Kim, 2014; Menozzi *et al*., 2015; Pandey *et al*., 2021; Rahamat *et al*., 2022; Zhu, 2018). The value of Chi-square (CMID/DF) statistic was 3.1 ($\chi 2$ = 532.7 / Degrees of Freedom = 173). According to this result, the goodness of fit of the model is interpreted as perfect. Another commonly used statistic for model goodness of fit is the index expressed as comparative fit index (CFI). In this study, the CFI value was found as 0.95. According to this result, the goodness of fit can be expressed as excellent.

The root mean square error (RMSE) is generally accepted statistic for model goodness of fit measure. In this study, the RMSE value was obtained as 0.086, according to this value the model is acceptable range at the margin. Ac-

Table 4 - Validity Analysis.

	CR	*AVE*	*SR*	*SN*	*PBC*	*ATT*	*ATB*	*INT*
SR	0,965	0,823	**0,907**					
SN	0,932	0,819	0,564***	**0,905**				
PBC	0,812	0,598	0,493***	0,635***	**0,773**			
ATT	0,966	0,905	0,827***	0,717***	0,592***	**0,951**		
ATB	0,922	0,798	0,714***	0,754***	0,572***	0,763***	**0,894**	
INT	0,931	0,819	0,664***	0,711***	0,658***	0,785***	0,745***	**0,905**

Table 5 - Confirmatory Factor Analysis.

			Standard Factor Loads	*CR*	*AVE*
SS3	←	SR	0.94	0,965	0,823
SR2	←	SR	0.92		
SS6	←	SR	0.89		
SS1	←	SR	0.90		
SS5	←	SR	0.87		
SS4	←	SR	0.89		
SN3	←	SN	0.92	0,932	0,819
SN2	←	SN	0.92		
SN1	←	SN	0.88		
PBC2	←	PBC	0.74	0,812	0,598
PBC3	←	PBC	0.61		
PBC1	←	PBC	0.94		
ATB1	←	ATB	0.85	0,922	0,798
ATB3	←	ATB	0.93		
ATB2	←	ATB	0.90		
INT2	←	INT	0.93	0,931	0,819
INT1	←	INT	0.89		
INT3	←	INT	0.90		
TTB3	←	TTB	0.94	0,966	0,905
TTB2	←	TTB	0.97		
TTB1	←	TTB	0.94		

cording to all three goodness of fit statistics, it can be said that the sub-dimensions (observed variables) of latent construct are statistically significant and well representing latent variables.

4.3. *Path analysis*

The second stage of SEM is the path model which is given below in Table 6. The path model shows the relationships between the observed variables and their latent counterpart variables, the bilateral co-variance relationships among the latent variables, and the relationship between the latent variables and the behavior variables.

In Table 6, three (***), two (**), and one (*) asterisk on the standardized beta coefficients (regression coefficients) are p values which confirming the coefficients are statistically significant at the level of 0.1%, 1%, and 5%, respectively (Beldad & Hegner, 2018; D'Souza, 2022; Giampietri *et al*., 2018; Kim, 2014; Menozzi *et al*., 2015; Pandey *et al*., 2021; Rahamat *et al*., 2022; Wang, 2020).

CMIN/DF, CFI and RMSE statistics are commonly used for model goodness of fit (Beldad & Hegner, 2018; D'Souza, 2022; Essakkat *et al*., 2021; Farid *et al*., 2023; Giampietri *et al*., 2018; Kim, 2014; Menozzi *et al*., 2015; Pandey *et al*., 2021; Rahamat *et al*., 2022; Zhu, 2018). Among these goodness-of-fit statistics, CMIN/DF ($\chi 2$ = 405.7/ Degrees of Freedom = 149) value is 2.72 (excellent), CFI value is 0.954 (>95 excellent), and "root of mean square error approximation" (RMSE) is found 0.078 (< 0.08 is at an acceptable level).

In the model, except for SN, other variables (ATB, TTB and PBC) affect intention (INT) in a statistically significant and positive way. Intention does not affect SR and TTB affects SR. Intention has not direct statistically significant

Table 6 - Path Analysis Results.

Estimator	*Result*	*Standardized Beta*	*Estimator*	*Result*	*Standardized Beta*
ATB	INT	0,224 **	INT	INT2	0,931 ***
SN	INT	0,114	INT	INT1	0,888
TTB	INT	0,400 ***	INT	INT3	0,895 ***
PBC	INT	0,221 ***	TTB	TTB3	0,945 ***
INT	SR	0,019	TTB	TTB2	0,966 ***
TTB	SR	1,028 ***	TTB	TTB1	0,942
SR	SR2	0,688	ATB	ATB1	0,843
SR	SR1	0,749 ***	ATB	ATB3	0,936 ***
SR	SR5	0,740 ***	ATB	ATB2	0,898 ***
SN	SN3	0,92	*INT*	*Behaviour*	-0,007
SN	SN2	0,918 ***	*SR*	*Behaviour*	0,189 *
SN	SN1	0,877 ***	*PBC*	*Behaviour*	0,243 **
PBC	PBC2	0,744	*ARR*	*Behaviour*	0,171 **
PBC	PBC3	0,607 ***			
PBC	PBC1	0,934 ***			

effect on behavior. Social Responsibility (SR), Perceived Behavioral Control (PBC) and Agricultural-Rural Relationship Level (ARR) significantly and positively affect behavior at 5%, 1% and 1% levels, respectively.

According to the results obtained by SEM estimation, H1, H3, and H4 hypothesis given in Figure 2 are accepted and H2 is rejected. The results of hypothesis test suggest that attitude towards behavior affects the intention to purchase cooperatives' branded food. Similarly, the perceived behavioural control also affects the intention to purchase cooperatives' branded food. Behavioral trust has impact on cooperatives' food purchase intention. Previous studies found that either directly or as a mediating variable, trust was found statistically significant variable on buying behaviour (Ding *et al*., 2022; Giampietri *et al*., 2018; Khan *et al*., 2023; Sultan *et al*., 2020). The subjective (normative) norm variable does not have a statistically significant effect on the intention to purchase cooperatives' branded food. According to the results of the first four hypotheses, latent variables related to attitude, perception and trust have a statistically significant effect on purchase intention, while subjective norm (SN) has no significant effect on intention. Similarly, intention has no effect on cooperatives' branded food purchase, therefore H5 was rejected. Perceived behavioural control (PBC) variable has a direct effect on purchasing behavior, so hypothesis H6 is accepted.

Considering that the intention, which is accepted as the antecedent of the behavior, can affect the behavior through social responsibility, the H7 hypothesis, which is established as intention affects social responsibility, is rejected. On the other hand, the H8 hypothesis, which was established with the statement that trust toward behavior (TTB) affects social responsibility, is accepted under the assumption that TTB can have an impact on behavior through social responsibility. It can be assumed as SR a mediating variable between trust and buying behaviour. A similar result was found in a study realized in Romania on organic food buying behaviour. However, it was found that social consciousness is not significant variable affecting buying intention, but has effect on personal attitudes and indirectly buying intension (Fleşeriu *et al*., 2020). It was indicated that traditional eating (reflecting cultural identity) is an mediating variables between intention and basic TPB latent variables social norms, attitudes and perception (Sogari *et al*., 2023).

Path analysis results show that the H9 and H10 hypotheses defined in the E-TPB model are valid. In other words, the hypothesis that SR is decisive on cooperatives' branded food purchasing behavior is accepted. The H10 hypothesis, which is defined as the level of agricultural and rural relationship of respondents is decisive on the behavior of purchasing cooperatives' branded food, is also accepted. Previous studies indicated that consumer ethnocentrism is significantly effecting buying intention directly and also basic latent variables of TPB (Miguel *et al*., 2022) or predictor of attitudes towards purchase intention (Maksan *et al*., 2019), as aforementioned, agricultural and rural ties (relation level) in this study reflects consumer ethnocentrism and directly effecting the behaviour.

5. Conclusion

This study aims to determine whether social responsibility (SR) perspective and agri-rural interest of consumers play a significant role in purchasing the agricultural cooperatives' branded food (collective brand). The theory of planed behaviour-TPB is extended with consumers' trust, social responsibility (SR), and agri-rural interest of consumers to verify validation of the TPB in collective food brand context. A survey data gathered via a face-to-face questionnaire implemented with a sample consisting of 284 person who is responsible for household food expenditure, with 18+ age and living in the central districts of Antalya province in the autumn of 2021. Simple random probability sampling based on finite population ratios was used to determine the sample size. Structural Equation Model (SEM) is used to test whether the E-TPB is valid.

Descriptive statistics indicates that average age of respondents is 36.5 year, 59.5% of respondents are female, 54.5% of respondents are married, and 35.2% of respondents are single. According to the results, respondents are highly educated persons (almost 50% is university graduated, 33.5% high school graduated), majority (71.8%) is working age population, almost one-third of respondent is professional occupation and one-quarter is unqualified service workers. Respondents purchasing frequency of cooperatives' food brand is found as 13%, 51.1%, 13%, 16.5%, and 3.5% respectively for purchase of regularly, seldomly, only at discounted price, non-purchased, and unavailable in the store. These results indicate that majority of consumers are familiar with cooperatives' branded food products.

According to the SEM results of the E-TPB model, social responsibility and agricultural-rural relationship level of consumer are found as significant model construct variables in the cooperatives' branded food purchasing behavior of consumers. Empirical results show that highlighting "social responsibility and agro-rural ties" in the marketing and communication of cooperatives' branded food products will strengthen the preference of these brands by consumers. In addition, strengthening the perception towards cooperatives' food brands will also positively affect purchasing behavior.

It can be said that strengthening the social responsibility behavior of consumers in relation to agriculture and rural development, highlighting the agricultural and rural relationship level of individuals, and strengthening the positive perception towards cooperatives' branded products can be a part of the product promotion and public relations strategy of the cooperatives. Additionally, cooperatives should support programs and projects aimed to enhance socially responsible consumers or requesting fund to such program from the agricultural support budgets.

This research will further be extended with nationwide representative survey or actual purchase data enabling to determine regional differences. Although the present study provides important information about consumer purchasing behaviour with respect to cooperatives' food brand, it has also some limitation since self-reported measures of behaviour does not reflect actual behaviour as sated in many studies. In addition, survey is not regionally and nationally representative, therefore results can be used with caution for generalizing results for nationwide or regionwide.

References

Ajates R., 2020. Agricultural cooperatives remaining competitive in a globalised food system: At what cost to members, the cooperative movement and food sustainability? *Organization*, 27(2): 337-355.

Ajzen I., 1985. From intentions to actions: A theory of planned behavior. In: Kuhl J., Beckmann J. (eds), *Action control*. Berlin-New York: Springer, pp. 11-39.

Ajzen I., 1991. The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2): 179-211.

Ajzen I., 2016. Consumer attitudes and behavior: the theory of planned behavior applied to food consumption decisions. *Italian Review of Agricultural Economics*, 70(2): 121-138.

Akbari M., Ardekani Z.F., Pino G., Maleksaeidi H., 2019. An extended model of Theory of Planned Behavior to investigate highly-educated Iranian consumers' intentions towards consuming genetically modified foods. *Journal of Cleaner Production*, 227: 784-793.

Aktulay Çakır T.M., 2014. *Postmodern tüketim ve tüketicinin değişen özellikleri "Online satın almaya yönelik tutumların ayrıştırılmış planlı davranış teorisi çerçevesinde incelenmesi"*. Maltepe Üniversitesi, Sosyal Bilimler Enstitüsü.

Alam S.S., Ahmad M., Ho Y.-H., Omar N.A., Lin C.-Y., 2020. Applying an extended theory of planned behavior to sustainable food consumption. *Sustainability*, 12(20): 8394.

Auza K.A., Mouloudj K., 2021. Using the theory of planned behavior to explore green food purchase intentions. *University of South Florida (USF) M3 Publishing*, 5: 96.

Bae Y., Choi J., 2021. Consumer acceptance of edible insect foods: an application of the extended theory of planned behavior. *Nutrition Research and Practice*, 15(1): 122-135.

Beldad A., Hegner S., 2018. Determinants of fair trade product purchase intention of Dutch consumers according to the extended theory of planned behaviour: The moderating role of gender. *Journal of consumer policy*, 41: 191-210.

Béné C., Fanzo J., Prager S.D., Achicanoy H.A., Mapes B.R., Alvarez Toro P., Bonilla Cedrez C., 2020. Global drivers of food system (un) sustainability: a multi-country correlation analysis. *Plos one*, 15(4): e0231071.

Candemir A., Duvaleix S., Latruffe L., 2021. Agricultural cooperatives and farm sustainability–A literature review. *Journal of Economic Surveys*, 35(4): 1118-1144.

Canova L., Bobbio A., Manganelli A.M., 2020. Buying organic food products: the role of trust in the theory of planned behavior. *Frontiers in Psychology*, 11: 575820.

Carfora V., Cavallo C., Caso D., Del Giudice T., De Devitiis B., Viscecchia R., Nardone G., Cicia G., 2019. Explaining consumer purchase behavior for organic milk: Including trust and green self-identity within the theory of planned behavior. *Food Quality and Preference*, 76: 1-9.

Chang H.-P., Ma C.-C., Chen H.-S., 2019. Climate change and consumer's attitude toward insect food. *International Journal of Environmental Research and Public Health*, 16(9): 1606.

Chen M.-F., 2017. Modeling an extended theory of planned behavior model to predict intention to take precautions to avoid consuming food with additives. *Food Quality and Preference*, 58: 24-33.

Chilla T., Fink B., Balling R., Reitmeier S., Schober K., 2020. The EU food label 'Protected Geographical Indication': economic implications and their spatial dimension. *Sustainability*, 12(14): 5503.

Cochran W.G., 1977. *Sampling techniques*. New York: John Wiley & Sons.

Demirtaş B., 2017. Zeytin üreticilerinin tarımsal yayım programlarına katılımının değerlendirilmesi. *Harran Tarım ve Gıda Bilimleri Dergisi*, 21(3): 309-322.

Ding L., Liu M., Yang Y., Ma W., 2022. Understanding Chinese consumers' purchase intention towards traceable seafood using an extended Theory of Planned Behavior model. *Marine Policy*, 137: 104973.

Dionysis S., Chesney T., McAuley D., 2022. Examining the influential factors of consumer purchase intentions for blockchain traceable coffee using the theory of planned behaviour. *British Food Journal*, 124(12): 4304-4322.

Donahue M., 2017. *Theory of Planned Behavior analysis and organic food consumption of American consumers*. Minneapolis: Walden University.

Dorce L.C., da Silva M.C., Mauad J.R.C., de Faria Domingues C.H., Borges J.A.R., 2021. Extending the theory of planned behavior to understand consumer purchase behavior for organic vegetables in Brazil: The role of perceived health benefits, perceived sustainability benefits and perceived price. *Food Quality and Preference*, 91: 104191.

Dowd K., Burke K.J., 2013. The influence of ethical values and food choice motivations on intentions to purchase sustainably sourced foods. *Appetite*, 69: 137-144.

D'Souza C., 2022. Game meats: Consumption val-

ues, theory of planned behaviour, and the moderating role of food neophobia/neophiliac behaviour. *Journal of retailing and consumer services*, 66: 102953.

EC, 2012. *Support for Farmers' Cooperatives*. https://agriculture.ec.europa.eu/system/files/2020-02/ext-study-support-farmers-coop-leaflet_2012_en_0.pdf.

Essakkat K., Mattas K., Unay-Gailhard I., Baourakis G., 2021. Youth's potential of adopting the Mediterranean diet lifestyle in response to climate change: Empirical study in Crete, Greece. *New Medit*, 20(5): 85-95.

Farid M.S., Cavicchi A., Rahman M.M., Barua S., Ethen D.Z., Happy F.A., Rasheduzzaman M., Sharma D., Alam M.J., 2023. Assessment of marketing mix associated with consumer's purchase intention of dairy products in Bangladesh: Application of an extended theory of planned behavior. *Heliyon*, 9(6).

Fiandari Y.R., Surachman S., Rohman F., Hussein A.S., 2019. Perceived value dimension in repetitive fish consumption in Indonesia by using an extended theory of planned behavior. *British Food Journal*, 121(6): 1220-1235.

Fleşeriu C., Cosma S.A., Bocăneț V., 2020. Values and planned behaviour of the Romanian organic food consumer. *Sustainability*, 12(5): 1722.

Frey O., 2023. *Mapping and challenges of the largest European agricultural cooperatives*. https://olivier-frey.com/report-top-100-european-coops/

Giampietri E., Verneau F., Del Giudice T., Carfora V., Finco A., 2018. A Theory of Planned behaviour perspective for investigating the role of trust in consumer purchasing decision related to short food supply chains. *Food Quality and Preference*, 64: 160-166.

Hair J.F., Anderson R.E., Babin B.J., Black W.C., 2010. *Multivariate data analysis: A global perspective*, 7th ed. Upper Saddle River, NJ: Pearson.

Han T.-I., Stoel L., 2017. Explaining socially responsible consumer behavior: A meta-analytic review of theory of planned behavior. *Journal of International Consumer Marketing*, 29(2): 91-103.

Hoeksma D.L., Gerritzen M.A., Lokhorst A.M., Poortvliet P.M., 2017. An extended theory of planned behavior to predict consumers' willingness to buy mobile slaughter unit meat. *Meat science*, 128: 15-23.

Kearney J., 2010. Food consumption trends and drivers. *Philosophical transactions of the royal society B: biological sciences*, 365(1554): 2793-2807.

Khan Y., Hameed I., Akram U., 2023. What drives attitude, purchase intention and consumer buying behavior toward organic food? A self-determination theory and theory of planned behavior perspective. *British Food Journal*, 125(7): 2572-2587.

Kim S.-H., Kuo W.-Y., 2022. The Role of Beliefs, Pride, and Perceived Barriers in Decision-Making Regarding Purchasing Value-Added Pulse Products among US Consumers. *Foods*, 11(6): 824.

Kim Y.G., 2014. Ecological concerns about genetically modified (GM) food consumption using the theory of planned behavior (TPB). *Procedia-Social and Behavioral Sciences*, 159: 677-681.

Kinikli F., Yercan M., 2023. The relationship between members' participation and organizational trustin cooperative firms: a case of dairy cooperatives in Izmir province-Turkey. *New Medit*, 22(1): 85-98.

Kyrylov Y., Thompson S.R., Hranovska V., Krykunova V., 2018. The world trends of organic production and consumption. *Management theory and studies for rural business and infrastructure development*, 40(4): 514-530.

Lappo A., Bjørndal T., Fernández-Polanco J., Lem A., 2015. Consumers 'concerns and External Drivers In Food Markets. *FAO Fisheries and Aquaculture Circular*(C1102), I.

Loera B., Murphy B., Fedi A., Martini M., Tecco N., Dean M., 2022. Understanding the purchase intentions for organic vegetables across EU: a proposal to extend the TPB model. *British Food Journal*, 124(12): 4736-4754.

Maichum K., Parichatnon S., Peng K.-C., 2017. Developing an extended theory of planned behavior model to investigate consumers' consumption behavior toward organic food: A case study in Thailand. *International journal of scientific & technology research*, 6(1): 72-80.

Maksan M.T., Kovačić D., Cerjak M., 2019. The influence of consumer ethnocentrism on purchase of domestic wine: Application of the extended theory of planned behaviour. *Appetite*, 142: 104393.

Menozzi D., Sogari G., Mora C., 2015. Explaining vegetable consumption among young adults: An application of the theory of planned behaviour. *Nutrients*, 7(9): 7633-7650.

Mercan N., 2015. Ajzen'in Planlanmış Davranış Teorisi Bağlamında Whistleblowing (Bilgi İfşası). *Sosyal ve Beşerî Bilimler Dergisi*, 7(2): 1309-8039.

Miguel L., Marques S., Duarte A.P., 2022. The influence of consumer ethnocentrism on purchase of domestic fruits and vegetables: application of the extended theory of planned behaviour. *British Food Journal*, 124(13): 599-618.

Oğuz C., Karakayacı Z., 2017. *Tarım Ekonomisinde Araştırma ve Örnekleme Metodolojisi*. Konya: Atlas Akademi, 183.

Oparebea Boateng A., Kwasi Bannor R., Bold E., Helena O.-K., 2023. A systematic review of the supply of agriproducts to supermarkets in emerging markets of Africa and Asia. *Cogent Food & Agriculture*, 9(1): 2247697.

Pakdemirli B., 2019. Tarımsal kooperatiflerin dünya ve Türkiye'de mevcut durumunun karşılaştırılması. *ANADOLU Ege Tarımsal Araştırma Enstitüsü Dergisi*, 29(2): 177-187.

Pandey S., Ritz C., Perez-Cueto F.J.A., 2021. An application of the theory of planned behaviour to predict intention to consume plant-based yogurt alternatives. *Foods*, 10(1): 148.

Pang S.M., Tan B.C., Lau T.C., 2021. Antecedents of consumers' purchase intention towards organic food: integration of theory of planned behavior and protection motivation theory. *Sustainability*, 13(9): 5218.

Pankobirlik, 2024. https://www.pankobirlik.com.tr/ (retrieved March 2024).

Paul J., Modi A., Patel J., 2016. Predicting green product consumption using theory of planned behavior and reasoned action. *Journal of retailing and consumer services*, 29: 123-134.

Qi X., Ploeger A., 2019. Explaining consumers' intentions towards purchasing green food in Qingdao, China: The amendment and extension of the theory of planned behavior. *Appetite*, 133: 414-422.

Qorri D., Felföldi J., 2024. Research Trends in Agricultural Marketing Cooperatives: A Bibliometric Review. *Agriculture*, 14(2): 199.

Rahamat S., Jeong E., Arendt S.W., Xu Y., 2022. Menu labeling influence on purchase behaviors: Applying the theory of planned behavior and health consciousness. *Appetite*, 172: 105967.

Raygor A.D., 2016. *The theory of planned behavior: Understanding consumer intentions to purchase local food in Iowa*, Iowa State University.

Rozenkowska K., 2023. Theory of planned behavior in consumer behavior research: A systematic literature review. *International Journal of Consumer Studies*, 47(6): 2670-2700.

Ruggeri G., Corsi S., 2021. An Exploratory Analysis of the FAIRTRADE Certified Producer Organisations. *World*, 2(4): 442-455.

Setyawan A., Noermijati N., Sunaryo S., Aisjah S., 2018. Green product buying intentions among young consumers: extending the application of theory of planned behavior. *Problems and Perspectives in Management*, 16(2): 145-154.

Shah Alam S., Mohamed Sayuti N., 2011. Applying the Theory of Planned Behavior (TPB) in halal food purchasing. *International journal of Commerce and Management*, 21(1): 8-20.

Shin Y.H., Jung S.E., Im J., Severt K., 2020. Applying an extended theory of planned behavior to examine state-branded food product purchase behavior: The moderating effect of gender. *Journal of Foodservice Business Research*, 23(4): 358-375.

Sığındı T., Kavak B., 2015. Satın alma niyetinin öngörüsünde planlı davranış modeli'nin farklı ürün sınıfları için denenmesi. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 15(2): 111-127.

Sogari G., Pucci T., Caputo V., Van Loo E.J., 2023. The theory of planned behaviour and healthy diet: Examining the mediating effect of traditional food. *Food Quality and Preference*, 104: 104709.

Sözüer A., Altuntaş G., Semercioöz F., 2015. Planlı Davranış Kuramını Yenilikçiliğe Uyarlamak: Kavramsal Bir Model. *Organizasyon ve Yönetim Bilimleri Dergisi*, 7(1): 30-43.

Spence M., Stancu V., Elliott C.T., Dean M., 2018. Exploring consumer purchase intentions towards traceable minced beef and beef steak using the theory of planned behavior. *Food Control*, 91: 138-147.

Sultan P., Tarafder T., Pearson D., Henryks J., 2020. Intention-behaviour gap and perceived behavioural control-behaviour gap in theory of planned behaviour: Moderating roles of communication, satisfaction and trust in organic food consumption. *Food Quality and Preference*, 81: 103838.

Taşçı-Duran E., 2016. Göç Eden Gebe Kadınların Planlı Davranış Kuramına Göre Doğum Öncesi Bakım Almaya Yönelik Niyet ve Tutumlarını Etkileyen Etmenler. *Journal of Hacettepe University Faculty of Nursing*, 3(1).

Teixeira S.F., Barbosa B., Cunha H., Oliveira Z., 2021. Exploring the antecedents of organic food purchase intention: An extension of the theory of planned behavior. *Sustainability*, 14(1): 242.

TKK (Agricultural Credit Cooperatives of Turkey), 2023. *2022 Yıllık Faaliyet Raporu*, retrieved March 2024 from https://www.tarimkredi.org.tr/faal%C4%B1yetler/kooperatif-market/.

Tommasetti A., Singer P., Troisi O., Maione G., 2018. Extended theory of planned behavior (ETPB): investigating customers' perception of restaurants' sustainability by testing a structural equation model. *Sustainability*, 10(7): 2580.

TurkStat (Turkish Statistic Institute), 2020. *International Migration Statistics*, retrieved 15 December from https://data.tuik.gov.tr/Kategori/GetKategori?p=nufus-ve-demografi-109&dil=1.

TurkStat (Turkish Statistic Institute), 2022. *Address Based Population Registration System*, retrieved 15 January from https://data.tuik.gov.tr/Bulten/Index-

?p=The-Results-of-Address-Based-Population-Registration-System-2022-49685&dil=2.

Vabø M., Hansen H., 2016. Purchase intentions for domestic food: a moderated TPB-explanation. *British Food Journal*, 118(10): 2372-2387.

Vermeir I., Verbeke W., 2008. Sustainable food consumption among young adults in Belgium: Theory of planned behaviour and the role of confidence and values. *Ecological economics*, 64(3): 542-553.

Wang E.S.-T., 2020. Hypotheses for the reasons behind beer consumer's willingness to purchase beer: an expanded theory from a planned behavior perspective. *Foods*, 9(12): 1842.

Wang J., Zhang Y., Liu S., 2020. *Analysis of consumers' perception and purchasing intention toward organic food based upon the Theory of Planned Behavior*. 2020 International Conference on Big Data Economy and Information Management (BDEIM),

Wang L., Luo J., Liu Y., 2021. Agricultural cooperatives participating in vegetable supply chain integration: A case study of a trinity cooperative in China. *Plos one*, 16(6): e0253668.

Yadav R., Pathak G.S., 2016a. Young consumers' intention towards buying green products in a developing nation: Extending the theory of planned behavior. *Journal of Cleaner Production*, 135: 732-739.

Yadav R., Pathak G.S., 2016b. Intention to purchase organic food among young consumers: Evidences from a developing nation. *Appetite*, 96: 122-128.

Yamane T., 2001. *Temel Örnekleme Yöntemleri*, 1. Baskı, *Çev. A. Esin, MA Bakır, C. Aydın ve E. Gürbüzsel*. Istanbul: Literatür Yayıncılık.

Yousuf J.B., Bose S., Kotagama H., Boughanmi H., 2019. Preferences and intentions of seafood consumers in Oman: An empirical analysis. *Journal of International Food & Agribusiness Marketing*, 31(2): 175-203.

Zhang L., Fan Y., Zhang W., Zhang S., 2019. Extending the theory of planned behavior to explain the effects of cognitive factors across different kinds of green products. *Sustainability*, 11(15): 4222.

Zhang Y., Jing L., Bai Q., Shao W., Feng Y., Yin S., Zhang M., 2018. Application of an integrated framework to examine Chinese consumers' purchase intention toward genetically modified food. *Food Quality and Preference*, 65: 118-128.

Zhu Y., 2018. Using the theory of planned behavior to investigate what influences Chinese intention to purchase organic food. *China-USA Business Review*, 17(6): 324-333.

NEW MEDIT

Scope

New Medit is an international open access journal published quarterly on behalf of the Italian Institute of CIHEAM of Bari.

The Journal seeks to act as a bridge between the relevant regional studies to enhance regional cooperation and sustainable development, by providing a platform for theoretical and empirical debate in the fields of social and economic sciences, regional development and environmental studies in the Mediterranean Region.

The main subjects of interest include the economic and social transformations in Mediterranean Countries in particular agro-food economics, rural development, environmental economics and sustainability issues around all Mediterranean region.

New Medit welcomes studies tackling the various problems characterising the economy and agribusiness of Mediterranean countries with a multi-disciplinary approach and from diverse theoretical perspectives.

Types of article

New Medit publishes *Research/Themes* papers and *Notes*/*Reports*. *Research/Themes papers* include original essays and cutting-edge research closely related to the aims and scope of the Journal. *Notes/Reports* include comments or studies on singles experiences which aim to facilitate the debate and dissemination of the real questions among the different territorial areas of the Mediterranean.

The journal operates with a double-blind peer review policy.

Papers are published in English or in French, with an abstract in English.

Guidelines

http://www.newmedit.iamb.it.

Printed in Italy
December 2024