Forecasting the dynamics of farm income: The case of the olive sector in Spain

José A. Gómez-Limón*, Sandra M. Sánchez-Cañizares**

DOI: 10.30682/nm2303a
JEL codes: Q18, Q12, C54, C33

Abstract

The objectives of this paper are to determine the factors that influence interannual movements of farms between income categories, and to forecast future income categories of farms under several different market, climate, and policy scenarios. To achieve these goals, a methodology combining the Markov chain model with a partial proportional odds model is proposed. Spanish olive farms are taken as an illustrative case study. The results show that the income dynamics of these farms are mainly influenced by off-farm uncontrollable factors such as the output prices, the weather conditions, and the policy support. Moreover, farm-, farmer-, and management-specific factors also play a relevant role.

Keywords: *Farm viability, Farm accountancy data network, Markov chain model, Partial proportional odds model, Scenario analysis.*

1. Introduction

Low income is the main factor driving farm abandonment (van der Zanden *et al.*, 2017). For this reason, ensuring farmers receive a 'fair' income has been an objective of the European Common Agricultural Policy (CAP) since its origins in 1957, as a way to maintain productive activity and guarantee food supply for the population, as well as to support the vitality of rural areas and encourage the provision of multiple ecosystem services (Finger and El Benni, 2021). In fact, "support viable farm income and resilience of the agricultural sector across the Union to enhance long-term food security and agricultural diversity" is the first of the nine specific objectives set out to guide the design and implementation of the CAP during the next programming period 2022-2027 (EC, 2018).

Despite this stated objective of the CAP, the European Union (EU) has never established any norms on what should be understood by a 'fair' or 'viable' income (Hill and Bradley, 2015). This lack of specificity means there is no normative reference level with which to compare the income actually obtained by European farms. To fill this gap, scholars have studied farm income in an attempt to establish different reference levels based on objective criteria, applying different analytical methodologies (e.g., Vrolijk *et al.*, 2010; Barnes *et al.*, 2020).

The study of farm income is a recurrent research topic within the agricultural economics

* Water, Environmental and Agricultural Resources Economics (WEARE) research group, Universidad de Córdoba, Spain.
** Department of Business Organization, Universidad de Córdoba, Spain.
Corresponding author: jglimon@uco.es

literature, especially when periods of difficulty are detected in certain farming subsectors. This is the case with the olive sector in Spain, which recently experienced an acute market crisis caused by a cycle of low olive oil prices (2018-2020), negatively impacting farm income (MAPA, 2021). This situation sparked large protests by olive growers throughout 2019 and 2020, leading to an intense social and political debate in production regions about the current role and prospects for olive farming. The high volatility of farm income experienced by the olive sector justifies both its use as an illustrative case study and the interest in analysing farm income from a dynamic point of view, assessing the factors explaining interannual changes in farm income.

Within this framework, the objective of this paper is threefold. First, to analyse how olive farm income has evolved during the period 2009-2018, using accounting data from a representative sample of Spanish olive farms to assess the share of these farms that achieve an adequate income level, thus ensuring their viability in the medium-to-long term. Second, to determine the structural and socio-economic factors explaining the heterogeneity in income dynamics of these farms. And third, to estimate the effects of several feasible market (changes in olive oil prices), production (reduction in olive yields because of climate change), and policy (reductions in the CAP support) scenarios on their income.

To achieve the abovementioned objectives, a farm typology is proposed based on different farm income levels. Considering the farm-level accounting information provided by the Spanish Farm Accountancy Data Network, every farm sampled has been classified into an income category for each year in the analysed period, allowing the modelisation of the dynamics of farm income, observing how individual farms move between categories across the years. For this purpose, the Markov chain model is used. This methodological approach has already been used in the agricultural economics literature, especially in studies focused on farm structural change (e.g., Rahelizatovo and Gillespie, 1999; Zimmermann and Heckelei, 2012). However, it has seldom been used to analyse the dynamics of farm income (Phimister *et al.*, 2004; Barnes *et al.*, 2015). In fact, this paper adds to the existing literature by combining the Markov chain model with an ordinal regression (partial proportional odds) model for ex-ante policy assessment of future market, production, and policy scenarios. Moreover, this methodological contribution is of interest because the method can be easily replicated using the same data source in any other farming sector and member state within the EU, allowing useful comparative studies to be carried out (e.g., comparison of the dynamics of farm income across olive farms in Spain, Italy, and Greece).

2. Measuring farm income: a typology

2.1. *Data*

The analysis of farm income necessarily relies on microeconomic data at the farm level, adequately reflecting the heterogeneity of these production units in terms of their capacity to generate revenue and remunerate the inputs employed. In this sense, the information provided by the Farm Accountancy Data Network (FADN) is the best available option in EU countries. For the case of Spain, these data are provided by the Spanish Farm Accountancy Data Network (*Red Contable Agraria Nacional*, RECAN), the Spanish branch of the FADN.

The RECAN annually collects structural, productive, economic, and financial information on a representative sample of Spanish commercial farms. Among the main advantages of using this data source are:

1. The sampling of farms is carried out by quotas according to the EU's farm typology (Regulation (EC) 1242/2008), considering the strata established by: a) economic dimension, quantified in terms of total standard gross margin (SGM) expressed in Euros; b) type of farming (TF); and c) Spanish Autonomous Communities.
2. The sample size of the RECAN annually exceeds 8,700 farms. This large size and the quota sampling procedure guarantee that the sample collected by the accounting

network is representative of the population of commercial farms in Spain.

3. The RECAN data gathering is carried out using the methodology applied throughout the EU (Regulation (EC) 1217/2009), thus contributing to a harmonised source of microeconomic data on farms at the European level. Therefore, the income indicators proposed for this work for Spanish olive farms could be replicated for other EU countries and other types of farming, enabling comparative analyses.

For all these reasons, the RECAN is a suitable database for the proposed analysis of the income dynamics of Spanish olive farms, allowing a reliable approximation of the heterogeneity in this agricultural sector.

The only limitation of the RECAN worth mentioning is that the population of farms analysed is not the whole population of Spanish farms (945,024 according to the latest official figures). As with all the other national FADN branches, the population considered by the RECAN consists only of "commercial" farms; that is, those with an annual SGM greater than 8,000 Euros (about 430,000 farms in Spain). Nevertheless, it should be pointed out that the sample collected annually by the RECAN represents a population of farms that manages 89% of the farmland in Spain (20.6 million hectares) and produces 96% of agricultural output at the national level. For this reason, the data and results obtained using this source are useful for policy analysis.

The analysis carried out was based on the microdata of the farms classified as the type TF 37 (specialised olive farms) included in the RECAN samples from 2009 to 2018. The size of the annual subsamples of farms belonging to the TF 37 has ranged throughout the period analysed between 224 (in 2009) and 363 (in 2018), with information available for a total of 3,156 observations (i.e., total number of farms for the full ten years).

2.2. *Farm income typology*

Many authors (e.g., Vrolijk *et al.*, 2010; Barnes *et al.*, 2020) propose assessing farm viability by taking several different income levels as references or benchmarks. In this paper, we follow this approach considering two references to measure the viability of olive farms in Spain. These two benchmarks are presented below, along with the viability indicators derived from them, which then allow us to classify the analysed farms according to their level of income.

The first income reference to be considered is the *total opportunity costs* incurred by the farmer because of the use of all internal resources (i.e., factors of production owned by the farmer) in his/her farming activities (O'Donoghue *et al.*, 2016; Coppola *et al.*, 2020). In the case where the farm income is enough to remunerate (i.e., higher than) all the opportunity costs for the use of the labour, capital, and land factors provided by the farmer, it can be said that factor allocation is economically efficient, making the farming activity viable in the long term. This income level would allow the generation of an economic surplus that can be reinvested in the farm, not only ensuring its economic sustainability but even enabling its growth.

To operationalise this first reference, a first viability indicator ($VI1$) is proposed as the ratio between the Farm Net Income (FNI) and the sum of the estimated values of the farmer's opportunity costs (labour, land, and capital):

$$VI1 = \frac{FNI}{OC_{labor} + OC_{land} + OC_{capital}} \quad (1)$$

The opportunity costs are estimated by calculating the potential remuneration that could be obtained if the factors of production provided by the farmer were used in the best possible alternative:

a) *Opportunity costs of labour* (OC_{labor}). For this case study, we valued this cost considering the average wage in the Spanish economy as the reference (EC, 2018). Thus, OC_{labor} was obtained at the farm level by multiplying this average wage by the agricultural work units provided by the farmer and his/her family.

b) *Opportunity costs of owned land* (OC_{land}). It is assumed that the best alternative use for owned land factor is renting it out (Coppola *et al.*, 2020). Thus, this opportunity cost was calculated by multiplying the number of hectares of owned farmland by the average rental price paid by farmers in the TF 37 (olive farming), with the latter data also being obtained from the RECAN microdata set.

c) *Opportunity costs of owned capital* (other than owned farmland) ($OC_{capital}$). Similar to previous studies (e.g., Vrolijk *et al.*, 2010; Coppola *et al.*, 2020), this opportunity cost was calculated on the basis of the interest paid for long-term public debt. Thus, $OC_{capital}$ was obtained by calculating the value of the farm equity less the value of owned land multiplied by the tax-free yield of 10-year government bonds.

The second income reference considered to assess the degree of viability of farms is the opportunity cost of the labour provided by farmers (Argilés, 2001; EC, 2018). Considering OC_{labor} as a reference for the farm income, the second of the viability indicators ($VI2$) is defined as follows:

$$VI2 = \frac{FNI}{OC_{labor}} \tag{2}$$

If the income of a farm is higher than this benchmark, it can be affirmed that this farm is viable in the short term, insofar as this income is a suitable remuneration for the labour provided by the farmer, allowing him/her to have a fair livelihood, similar to other people working in other economic sectors. Conversely, those farms with an income below the opportunity cost of labour can be considered as non-viable since farming is achieved at the cost of undervaluing the labour provided by the farmer. Indeed, under these circumstances, the farm is economically unsustainable in the long run, and the continuity of production is only explained by the farmer's lack of labour opportunities.

Taking into account the two abovementioned viability indicators, olive farms can be classified into three categories. Those farms with $VI2$ values lower than or equal to one are considered "*non-viable*" (category 1 −*C1*− of the viability scale). Those farms with $VI2$ values higher than one, but with $VI1$ values lower than one are considered "*viable in the short term*" (category 2 −*C2*−of the viability scale). Finally, those farms with a value of $VI1$ greater than or equal to one can be qualified as "*viable in the long term*" (category 3 −*C3*− of the viability scale).

3. Methodological approach for analysing the dynamics of farm income

3.1. *Markov chain approach*

Zimmermann *et al.* (2009) provide a literature review of relevant methods for forecasting change in the distribution of farm characteristics (i.e., number of farms in classes or categories defined by a typology). These authors conclude that the Markov chain model (MCM) is the most suitable approach to analyse the dynamics of farm changes (i.e., movements of farms between categories).

An MCM focused on the dynamics of farms is based on three basic elements (Rahelizatovo and Gillespie, 1999): a) a farm typology considering a finite set of C farm categories; b) the initial distribution of farms according to this typology, described by the matrix X^0 ($1 \times C$), where x_i^0 represents the number (or share) of farms in the category i in the first period analysed (t=0); and c) the stochastic transition probability matrixes (TPM) P^t ($C \times C$) showing the probabilities of moving between farm categories during the T periods considered (t=1,...T).

When the TPM does not change over time, it is said the MCM is stationary. However, this is not generally the case for economic phenomena such as farm income, which is affected by multiple exogenous variables (e.g., product prices, input costs, production technology, public support, legal requirements, etc.). Since changes in these exogenous variables impact farm viability, transition probabilities are time-varying, leading to a non-stationary MCM (i.e., different TPMs for each period t). Accounting for non-stationarity, any change process considering an initial farm distribution X^0 and the TPMs P^t can be represented as follows:

$$X^0_{(1\times C)} \times P^1_{(C\times C)} \times P^2_{(C\times C)} \times \ldots \times P^T_{(C\times C)} = X^T_{(1\times C)} \tag{3}$$

where the matrix X^T ($1 \times C$) presents the farm distribution in period T. Thus, this general expression can be used to forecast the future distribution of farms among the C categories considered when the matrixes P^t structure is known.

Each element p_{ij}^t of the TPM P^t represents the probability of a single farm classified in category i in period t−1 being classified in category j in period t. These transition probabilities also have the following two characteristics: a) $0 \leq p_{ij}^t \leq 1$ for every category i and j, and every period t, and b) $\sum_{j=1}^{C} p_{ij}^t = 1$ for every category i and period t.

It is usually assumed that the movement of farms from one farm category to another follows a first-order Markov chain; this is, that the probability of the movement of a farm in period t-1 to another farm type in period t is independent of movement in earlier periods. In these cases, the number of farms in category j in period t (n_j^t) depends on the number of farms in all farm categories i in the preceding period (t-1) multiplied by their respective transition probabilities p_{ij}^t:

$$n_j^t = \sum_{i=1}^{C} n_{ij}^{t-1} \times p_{ij}^t \qquad (4)$$

If microdata are available to account for single farm movements between categories for each period, transition probabilities p_{ij}^t can be estimated as follows:

$$\hat{p}_{ij}^t = \frac{m_{ij}^t}{\sum_{j=1}^{C} m_{ij}^t} \qquad (5)$$

where m_{ij}^t is the number of farms in category i in period t-1 that moved to category j in period t.

Note that transition probabilities obtained using Equation (5) based on observed data are just estimated values of real unknown parameters (p_{ij}^t). However, the values recovered by the use of the microdata are proven to be the maximum likelihood estimators of real transition probabilities (Gourieroux, 2012), allowing their use as unbiased values of these parameters for empirical applications.

Moreover, transition probabilities explaining the dynamics of farms' characteristics are functions of a full array of exogenous factors. Most of these factors are time-varying, justifying the non-stationary MCM. For this reason, an econometric model estimating the effect of these independent variables on transition probabilities is also required:

$$\hat{p}_{ij}^t = f_{ij}(Z^t, \beta_{ij}) \qquad (6)$$

where f_{ij} is the function of the vector of explanatory variables Z^t and the matrix of parameters β_{ij} which relates to the independent variables considered.

Based on this theoretical framework, we implement a two-step approach: the first step involves calculating the non-stationary transition probabilities using Equation (5), and the second step estimating the influence of the exogenous variables on these probabilities using Equation (6).

3.2. *Factors determining the dynamics of farm income*

Coppola *et al.* (2020) and Barnes *et al.* (2020) have recently reviewed the factors affecting the income levels and viability of EU farms. They highlight the influence of the farmer's socio-demographic characteristics, the farm's structural characteristics, and the farmer's productive choices. Moreover, it is also worth pointing out the role of off-farm uncontrollable factors in farms' income dynamics, notably those related to the volatility of agricultural markets and changing weather conditions (e.g., Poon and Weersink, 2011) and those linked to shifting agricultural policy instruments (e.g., Biagini *et al.*, 2020; Cardone *et al.*, 2021). Taking into account this evidence and the information available for the empirical analysis, four kinds of factors were considered as explanatory variables that may potentially shape the functions f_{ij} (i.e., influence the dynamics of olive farm income):

1. *Farmer's socio-demographic characteristics*: age (*AGE*), agricultural training (*AGTRAIN*), family labour (*FAMLAB*), and land ownership (*LANDOWN*).
2. *Farm's structural characteristics* (economies of scale, agronomic suitability): farm size (*FSIZE*), agronomic suitability for olive production (*AGSUIT*), and olive area under irrigation (*IRRIG*).
3. *Farmer's productive choices* (production technology and financial situation): specialisation in olive production (*SPEC*), intermediate consumption intensity (*ICINT*), capital intensity (*CAPINT*), outsourcing (*OUTSOUR*), and debt-equity ratio (*DEBEQRAT*).
4. *Off-farm uncontrollable factors* (market, climatic, and policy conditions): a) bulk olive oil price (*OPRICE*), b) annual weather con-

ditions accounting for precipitation and temperature differences impacting olive yields (*WEATHER*), c) CAP decoupled payments (*CAPDP*), and d) interest rate (*IRATE*).

Table 1 shows the details about how these variables are operationally defined, the units of measurement, and the sources of the data. Table 2 shows the descriptive statistics of the explanatory variables included in the analysis.

The regression models estimated (see next section) do not include the variables *IRATE* (interest rate) and *IRRIG* (olive area under irrigation) because of multicollinearity problems. Specifically, *IRATE* presented a high correlation with *OPRICE*, while *IRRIG* was collinear with *AGSUIT* and *ICINT*. Although several alternative regression models were proposed to overcome the multicollinearity problems, we opted for those that elim-

Table 1 - Potential explanatory variables of the dynamics of olive farm income.

Theoretical concept	*Variable*	*Acronym*	*Type of variable*	*Measurement*	*Units*	*Source*
Farmer's socio-demographic characteristics	Age	AGE	Farm	Age	Years	RECAN
	Agricultural training	AGTRAIN	Farm	Dummy variable: only practical experience (0); agricultural degree (1)	---	RECAN
	Family labour	FAMLAB	Farm	Family labour as a percentage of total farm labour	Percentage	RECAN
	Land ownership	LANDOWN	Farm	Owned land as a percentage of total farmland	Percentage	RECAN
Farm's structural characteristics	Farm size	FSIZE	Farm	Farm size	Hectares	RECAN
	Agronomic suitability for olive production	AGSUIT	Farm	Average olive yield 2009-2018 as a time-invariant factor measuring land productivity	kg olive oil/ hectare	RECAN
	Olive area under irrigation	IRRIG	Farm	Irrigated olive area as a percentage of the total olive area	Percentage	RECAN
Farmer's productive choices	Specialization in olive production	SPEC	Farm	Olive area as a percentage of total farmland	Percentage	RECAN
	Intermediate consumption intensity	ICINT	Farm	Value of intermediate consumption (fertilizers, phytosanitary products, fuel, etc.) per hectare	€/hectare	RECAN
	Capital intensity	CAPINT	Farm	Non-land assets per hectare	€/hectare	RECAN
	Outsourcing	OUTSOUR	Farm	Agricultural practices subcontracted over total costs	Percentage	RECAN
	Debt-equity ratio	DEBEQRAT	Farm	Total debt over equity	Percentage	RECAN
Off-farm uncontrollable factors (climatic, market, and policy conditions)	Bulk olive oil price	OPRICE	National	Olive oil price index based on the average bulk olive oil price in Spain 2009-2018=100%	Percentage	Spanish Ministry of Agriculture (MAPA)
	Annual weather conditions	WEATHER	Province	Province yield index based on the average yield for rain-fed olive 2009-2018=100%	Percentage	Spanish Ministry of Agriculture (MAPA)
	CAP decoupled payments	CAPDP	Farm	CAP decoupled payments per hectare	€/hectare	RECAN
	Interest rate	IRATE	National	Spanish Government 10Y Bond yield	Percentage	Spanish Ministry of Finance

Note: Monetary variables were deflated using the Spanish Consumer Price Index (CPI) (Instituto Nacional de Estadística, INE, www.ine.es).

Table 2 - Descriptive statistics of explanatory variables of the dynamics of olive farm income.

Variable	*Acronym*	*Average*	*St. Dev.*	*Skewness*	*Kurtosis*
Age	AGE	58.38	10.91	0.103	0.430
Agricultural training	AGTRAIN	0.11	0.32	2.462	4.064
Family labour	FAMLAB	62.40%	25.56%	-0.305	-0.604
Land ownership	LANDOWN	87.08%	28.24%	-2.220	3.652
Farm size	FSIZE	39.50	59.44	6.723	64.007
Agronomic suitability for olive production	AGSUIT	668.09	384.66	1.894	12.043
Olive area under irrigation	IRRIG	28.75%	41.61%	0.939	-0.941
Specialization in olive production	SPEC	93.25%	14.56%	-2.386	5.317
Intermediate consumption intensity	ICINT	584.21	410.13	1.452	2.649
Capital intensity	CAPINT	3,998.92	5,322.31	9.713	134.594
Outsourcing	OUTSOUR	4.65%	9.53%	2.660	7.755
Debt-equity ratio	DEBEQRAT	1.47%	24.43%	34.638	1,410.547
Bulk olive oil price	OPRICE	105.93%	24.25%	0.169	-1.189
Annual weather conditions	WEATHER	102.20%	36.66%	0.362	1.695
CAP decoupled payments	CAPDP	544.93	433.34	1.630	5.075
Interest rate	IRATE	4.14%	0.54%	-0.369	-1.394

Note: The descriptive statistics reported have been calculated using the 3,156 observations gathered by the RECAN subsamples for the TF 37 (specialized olive farms) from 2009 to 2018. However, it is worth noting that the subsample size for this type of farming has not remained constant throughout the period analysed; it has gradually increased from 224 farms in the year 2009 to 363 farms in the year 2018. This explains why the averages of the variables WEATHER or OPRICE are not equal to 100% as might be expected. In these cases, mean values slightly higher than 100% actually indicate that olive yields and prices for olive oil were higher during the last years considered (those with larger sample size) than over the first years analysed (those with smaller sample size).

inate *IRATE* and *IRRIG* since they had the best goodness-of-fit statistics and predictive power.

3.3. *Partial proportional odds model*

Different regression techniques have been used to estimate functions f_{ij} included in Equation (6): least squared procedures, multinominal logit, and ordinal regression models (e.g., Rahelizatovo and Gillespie, 1999; Zimmermann and Heckelei, 2012). Since the income level of olive farms in our case study is ranked from the least viable category (i.e., "non-viable", C1) to the most viable category (i.e., "viable in the long term", C3), the dynamics of farm income can be modelled using an ordinal regression model where the farm income category is considered as the dependent variable (y).

Among the different regression models for ordinal responses, the ordinal regression model is the most traditional. However, this model requires the proportional odds or the 'parallel lines' assumption, i.e., the effects of independent variables or beta coefficients are equal at different thresholds (categories) of the dependent variable. This assumption is often violated in the sense that one or more coefficients can differ across values of y. To solve this problem, Peterson and Harrell (1990) proposed the partial proportional odds model (PPOM), where the parallel lines assumption can be relaxed for a subset of explanatory variables in the model. This means the PPOM contains the proportional odds for independent variables that do not violate this assumption, but estimates additional coefficients for those predictors which do not fulfil it. This model provides a more accurate estimation than other available modelling techniques (e.g., multinomial logistic model) given that not all the independent variables have to violate the parallel lines assumption.

The PPOM can be written as:

$$\ln\left(\frac{Pr\{y \leq m|x\}}{Pr\{y > m|x\}}\right) = \tau_m - x\beta - \omega\eta_m \quad (1 \leq m \leq M) \tag{7}$$

where m is an ordered response category (3 levels in our empirical study), x and ω are vectors of explanatory variables which meet/do not meet the proportional odds assumption, respectively, β is a vector of unknown regression coefficients corresponding to x, η_m a vector of coefficients corresponding to ω that vary across cutpoint equations and, finally, τ_m the vector of thresholds or cut points for each category of y.

Therefore, in the empirical study carried out, three ordinal regression models were fitted taking each one of the income categories in turn as the dependent variable, and Brant tests (Brant, 1990) were conducted to test the proportional odds assumption for the models as a whole and for each of the explanatory variables. This test compares the β coefficients from m-1 binary logits; the null hypothesis is that these coefficients are equal for all logit models. Therefore, when the null hypothesis is rejected, it means that the β coefficient is violating the parallel assumption and different coefficients should be estimated for each category. The results of these tests showed that a subset of explanatory variables did not fulfil the parallel assumption in every model run, indicating the suitability of the PPOM proposed as the regression technique.

Subsequently, a PPOM was estimated for each dependent variable using the *gologit2* command by Williams (2006) in Stata 14.0 software specifying the *autofit* option. This prompts the *gologit2* command to go through an iterative process, running a series of Wald tests on each independent variable to check if their coefficients are different across equations. The final model imposes constraints on variables that do not violate the proportional odds to keep the same coefficient estimate, while the rest are unconstrained and show different values for each equation.

By estimating these three PPOMs, we seek to determine the explanatory factors of the income dynamics of the olive farms included in categories C1 ("non-viable"), C2 ("viable in the short term"), and C3 ("viable in the long term"). Each model presents two panels representing the probability of staying in the same category or changing to a different (lower or higher) one. Stata selects C3 as the reference category, which means that current and lower categories (C1, and C1 and C2 in the first and the second panels, respectively) are taken as the base group and then compared to the more viable groups (C2 and C3, and C3 in the first and the second panels, respectively).

3.4. *Scenario analysis*

The MCM approach can also be used for ex-ante policy assessment, allowing comprehensive and valid forecasts of future shares of farms included in each income category under different relevant scenarios (Zimmermann *et al.*, 2009). In this regard, the scenario analysis performed here is focused on the off-farm uncontrollable factors to evaluate how market, climatic, and policy conditions could impact the viability of olive farms.

The BASELINE scenario for the analysis proposed is defined by the current olive farms' structure, as described in the last available RECAN subsample for the TF 37 (data gathered in the year 2018), considering average market and climatic conditions for 2009-2018 (i.e., *OPRICE*=100%, *WEATHER*=100%, and *IRATE*=4.24%) and the latest data on policy support (i.e., *CAPDP*=farms' specific CAP payments in 2018). Assuming that farms' structure remains constant, the following three off-farm uncontrollable variables have been considered key for the definition of the scenarios to be analysed: *OPRICE, WEATHER*, and *CAPDP.*

Olive oil prices in the international markets are determined by the laws of supply and demand, but they are also affected by factors such as speculative activity in the market, information asymmetry, currency fluctuations, and government policies (Mili and Bouhaddane, 2021). Interannual imbalances between global olive oil production (affected by events related to weather, pests, and diseases) and demand (influenced by changes in the prices of other substitute vegetable fats), along with other factors shaping the market, lead to a high interannual (between crop years) price volatility (Abid and Kaffel, 2018). This volatility has meant the olive oil sector in Spain (and elsewhere) has experienced recurrent boom and bust cycles in producer prices, which has jeopardised olive farms' income stability and viability (Gontijo *et al.*, 2020). Moreover, there is no consensus about expected future trends in the price of olive oil since both production and demand are rising worldwide, and it is not clear which will be the dominant driver (Mili and Bouhaddane, 2019). This explains why price volatility has been a major concern for policymakers and justifies the inclusion of *OPRICE* as a key variable to define policy-relevant scenarios.

Over the decade analysed (2009-2018), olive oil prices in Spain ranged from €1.84 to €3.71/kg (average €2.58/kg). For this reason, two feasible price scenarios are considered. First, supposing that an increase in world production (modern high-density and super-high-density groves) would lead to a downward trend in prices, an average price of €2.00/kg is proposed (scenario OP_2EUR), with the variable *OPRICE* taking the value of 77.67%. Second, if increasing demand were to be the dominant driver, a scenario of rising prices is also suggested, considering an average price of €3.00/kg (scenario OP_3EUR), where *OPRICE* would be equal to 116.51%.

The temperatures that regulate olive tree phenology (dormancy period, flowering, and fruit maturation) and the precipitation that determines water availability for olive trees grown under rainfed conditions (69.9% in Spain) are considered the most important climatic factors conditioning olive yields (Fraga *et al.*, 2021). Thus, interannual variations in local weather conditions affect olive yields, both directly (depending on extreme events such as frosts or heatwaves) and indirectly (by influencing the incidence of pests and diseases), thereby determining olive farms' revenue and income. Wide interannual fluctuations in temperature and rainfall, and thus large variations in olive oil production, are distinctive features of the Mediterranean climate. However, according to the Intergovernmental Panel on Climate Change, future climate projections point to the Mediterranean Basin as a climate change "hotspot", where temperatures will continue to rise and precipitation patterns will shift (IPCC, 2015). These warming (meaning higher evapotranspiration and water demand) and drying (i.e., less water availability) trends are expected to strongly affect olive yields in Spain and all other Mediterranean countries (Arenas-Castro *et al.*, 2020; Cabezas *et al.*, 2021). For the Spanish case, a substantial decrease is projected in rainfed olive yields (down by 45%) (Fraga *et al.*, 2020). This evidence leads us to consider *WEATHER* as another key variable to define future scenarios. Thus, we analyse scenarios in which rainfed olive yields will be reduced by 20% (scenario YIELD-20%, where *WEATHER*=80%) and 40% (scenario YIELD-40%, where *WEATHER*=60%). The former scenario assumes technology innovation and adaptation measures will be able to minimise the negative effects of climate change, while the latter assumes such strategies will not be implemented.

Farm incomes in the EU have traditionally benefited from strong public support through the CAP. In the case of Spanish olive farming, the estimated Producer Subsidy Equivalent (PSE, an indicator measuring total monetary transfers to agricultural producers) reaches, on average, 42% of the gross olive producer revenues (Júdez *et al.*, 2017). Most of this public support for olive growers is received through decoupled payments per hectare, set based on past references (historical model). As a result, the average Spanish olive grower currently receives far more in CAP payments (€475.44/ha, RECAN, 2020) than the average Spanish farmer (€266.84/ha, RECAN, 2020). However, the new CAP reform has introduced updated regulations aimed at ensuring more equitable support for all European farmers. The required convergence in decoupled payments will lead to a reduction in the value of payment entitlements that exceed the national average, as is the case of olive growers (Chousou *et al.*, 2020). This likely reduction in the level of support will also negatively impact olive farms' income, which justifies the selection of *CAPDP* as another key variable worth considering when defining policy-relevant scenarios. Thus, a scenario involving a 30% reduction in these payments for all olive farms is proposed (scenario CAP-30%, where *CAPDP* is calculated for each farm as 70% of the CAP payments received in 2018). Additionally, taking into account that the reductions in decoupled payments will be targeted on the basis of farm size (i.e., the introduction of the new redistributive payment for the first hectares), another CAP scenario is suggested in which there is a 50% reduction in CAP support, but with this cut only being implemented after the first 10 of the hectares for which the farmer is entitled to receive payments (scenario CAP-50%+10HA, where *CAPDP* is calculated for each farm depending on the CAP payments received in 2018 and its size).

The scenarios proposed above are illustrative of the versatility of MCM for ex-ante policy analysis. In fact, as the reader might suppose, any other scenario affecting the variables explaining olive farm income dynamics could be defined for future predictions.

Considering the values of the explanatory variables for the base year ($t=k$, the year 2018 in our case study) in each scenario, the PPOM (Equation (7)) to be estimated can be used to predict the probability of the movement of each farm between income categories ($\hat{p}_{ij}^{t}$). These predictions give us the TPMs P^t for each scenario, and thus farm distribution among categories for the next year ($t=k+1$, the year 2019 in our case study):

$$X^k \times P^k_{scenario\_X} = X^{k+1 \text{ under } scenario\_X} \quad (8)$$

However, this kind of prediction for the next year is not very useful since the results obtained would be out of date by the time they are calculated, and they do not reflect the actual impact of the scenarios considered, since these results for the $k+1$ period are highly dependent on the initial farm distribution (X^k). For this reason, it is worth assuming that the variables defining the scenarios will remain constant for the next m years, until the farm distribution became stationary; that is, when the distribution X^t remains constant for any $t \geq k+m$:

$$X^k \times P^k_{scenario\_X} \times P^{k+1}_{scenario\_X} \times \ldots \times \times P^{k+m-1}_{scenario\_X} = X^{k+m \text{ under } scenario\_X} \quad (9)$$

As the stationary distribution X^{k+m} does not depend on the initial distribution X^k, it reflects the actual impact on farm income of the scenario considered. The scenario analysis is thus aimed at calculating and analysing the stationary distributions for each scenario, allowing us to assess the impact of the proposed changes in the different off-farm uncontrollable factors on the near future viability of Spanish olive farms.

Finally, note that caution should be taken when comparing the stationary distribution under the BASELINE scenario with those resulting from the different policy scenarios proposed. It is worth recalling that the latter results do not take into account possible responses of olive growers to the scenario changes in terms of their income and cost structure (farms' structure is assumed to remain constant). Despite this shortcoming, these comparative analyses are useful for exploring the primary effects of the three factors studied on the viability of the farms analysed (Vrolijk *et al.*, 2010).

4. Results and discussion

4.1. *The dynamics of olive farm income: transition probabilities*

The RECAN annually collects data from a rotating panel of farms. As shown in Table 3, for the case of the TF 37, olive farms remain in the RECAN panel for varying lengths of time. In fact, only 111 out of the 576 olive farms sampled from 2009 to 2018 have remained in the TF 37 annual subsamples for the whole period. In any case, the 3,156 observations collected in the TF 37 annual subsamples throughout this decade yield 2,555 interannual observations (i.e., single farms sampled in two consecutive years) to analyse the dynamics of farm income. These interannual observations make it possible to account for single farm movements between farm income categories from year t-1 to year t (m_{ij}^{t}) and regress the corresponding transition probabilities (p_{ij}^{t}) as ordinal dependent variables with the set of independent variables proposed (Z^t) using the PPOM approach, as explained in Equation (7).

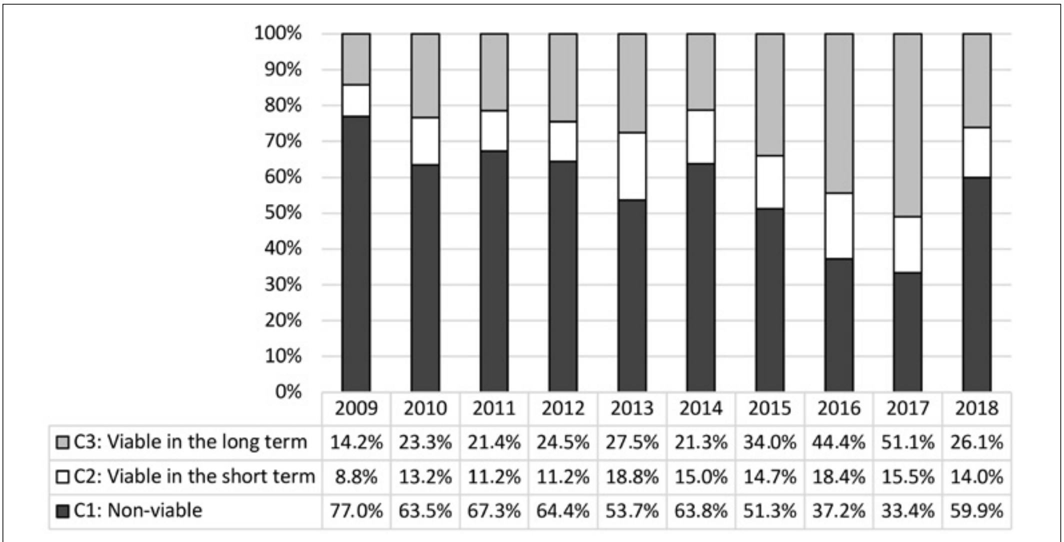
Table 3 - Farms included in the TF 37 (specialized olive farms) subsamples from 2009 to 2018.

Consecutive years in the subsample	*Num. of farms*	*% farms*	*Num. annual observations*	*Num. interannual observations*
1	84	14.6%	109	0
2	35	6.1%	70	35
3	58	10.1%	174	116
4	81	14.1%	324	243
5	33	5.7%	165	132
6	31	5.4%	186	155
7	131	22.7%	917	786
8	7	1.2%	56	49
9	5	0.9%	45	40
10	111	19.3%	1,110	999
Total	*576*	*100.0%*	*3,156*	*2,555*

Using the data from this unbalanced panel of farms could lead to some attrition bias as olive growers voluntarily leaving the RECAN sample cannot be controlled for, and their replacement could generate some sampling noise. However, as pointed out by Barnes *et al.* (2015), this bias is found to be low since the average farm remains in the sample for a reasonable length of time (70% of olive farms remain in the RECAN sample for 4 years or more). Thus, interannual observations obtained as explained above can be judged suitable enough for implementing the microdata MCM approach proposed.

Figure 1 shows the proportions of the olive farms included in the empirical analysis (i.e., those with interannual observations available; n=2,555) classified as "non-viable" (C1), "viable in the short term" (C2), and "viable in the long term" (C3) from 2009 to 2018. As expected, these shares fluctuate over the course of the decade under analysis. For instance, the proportion of "viable in the long term" farms ranges from

Figure 1 - Distribution of farms among farm income categories from 2009 to 2018.



51.1% in 2017 to only 14.2% in 2009, while the share of "non-viable" farms varies from 77.0% in 2009 to 33.4% in 2017. As will be analysed in the next section, these fluctuations can be explained by off-farm uncontrollable factors affecting olive oil price (market conditions) or olive yields (weather conditions), and also by farm factors such as farm size, production technology, or the farmer's management skills.

Taking the three income categories proposed, the MCM approach was implemented by shaping the TPMs P^t as follows:

$$P^t(\hat{p}_{ij}^t) = \begin{pmatrix} \hat{p}_{C1,C1}^t & \hat{p}_{C1,C2}^t & \hat{p}_{C1,C3}^t \\ \hat{p}_{C2,C1}^t & \hat{p}_{C2,C2}^t & \hat{p}_{C2,C3}^t \\ \hat{p}_{C3,C1}^t & \hat{p}_{C3,C2}^t & \hat{p}_{C3,C3}^t \end{pmatrix} \quad (10)$$

Available RECAN microdata allow the calculation of the maximum likelihood estimators of non-stationary transition probabilities $\hat{p}_{ij}^t$ on an annual basis following Equation (5). Thus, considering the initial farm distribution X^0 (related to the year 2009), and the TPMs P^t from $t=1$ (year 2010) to $t=9$ (year 2018), the final farm distribution X^9 (the year 2018) can be expressed using Equation (3) as follows: $X^0 \times P^1 \times P^2 \times \ldots \times P^9 = X^9$.

Table 4 shows the average transition probabilities over time for the period analysed (the diagonal is shaded), along with the corresponding standard deviations. Transition probabilities show both high variability across categories and high variability over time.

The highest values in Table 4 are found on the diagonal (except for C2) and represent the probabilities of remaining in the same income category as in the year before. In other words, viable (non-viable) farms tend to remain viable (non-viable) in the following year. This pattern is widely seen in TPMs representing economic phenomena, and in our case study indicates that single farms 'resist' transitioning to other categories. This could be explained by the fact that their income level is strongly influenced by the structural features of olive farms (olive growing is based on a perennial crop for which it is hard to make changes in production technology in the short term). These results also suggest transitions of olive farms between income categories are mainly caused by off-farm uncontrollable factors, such as the price of olive oil, weather conditions determining olive yields, and CAP payments. In any case, these hypotheses will be tested in the next section.

The income category C2 is the exception (aggregated $\hat{p}_{C2,\,C2}^{t}$=22.4%, while aggregated $\hat{p}_{C2,\,C1}^{t}$ and $\hat{p}_{C2,\,C3}^{t}$ are higher than 30%), which can be explained by the relative 'narrowness' of this category (i.e., the short-range of farm income defining the requirements for inclusion in this category: $PI1<1$ and $PI2>1$). In fact, only 14.5% of the interannual observations taken into account (371 out of 2,555) are considered "viable in the short term" in the year t-1. This means that even small changes in the variables determining the dynamics of farm income lead "viable in the short term" farms to transition to another income category. Furthermore, this situation also explains why probabilities adjacent

Table 4 - Transition probabilities (p_{ij}^t) and standard deviations over time.

Profitability category in year t-1	*Profitability category in year t*	*C1*	*C2*	*C3*
C1: *Non-viable*	Average	71.6%	14.0%	14.5%
	St. Dev.	0.084	0.053	0.056
	Num. observations	1,010	197	201
C2: *Viable in the short term*	Average	42.8%	22.4%	34.7%
	St. Dev.	0.216	0.091	0.148
	Num. observations	160	85	126
C3: *Viable in the long term*	Average	22.2%	16.8%	61.0%
	St. Dev.	0.216	0.091	0.148
	Num. observations	195	117	464

to the diagonal (i.e., indicating the relative frequency of transition to neighbouring categories) are not higher than those farther away from the diagonal (i.e., indicating the relative frequency of transition from C1 to C3 or C3 to C1), as is usually found in TPMs.

4.2. *Partial proportional odds models*

Table 5 displays the coefficients (β) and odds ratio (OR) estimates for the independent variables considered for each of the three estimated PPOMs (section 3.3).

Goodness-of-fit statistics are also reported for the three models, showing good values for all these measures. The LR χ^2 test allows us to reject the null hypothesis that the performance of the estimated model is similar to a null model with only the intercept, thus indicating the overall estimated model is statistically significant. McFadden's pseudo R^2 (or LR index) ranges between 19.7% and 22.9%. The count statistic, which reports the proportion of correct predictions, fluctuates between 60.3% (model C2 has the worst predictive potential) and 73.6-73.5% for models C1 and C3, respectively. These values both for McFadden's R^2 and Count can be considered fairly high when compared with other papers that also use PPOM regressions (O'Connell and Liu, 2011). Finally, Log-likelihood at zero and at convergence, as well as Akaike (AIC) and Bayesian information criterion (BIC), are statistics used to compare models. All these statistics (except for LL at zero, which logically stays the same) present better results here than in the initial ordinal logit regression models, confirming that the PPOM provides a more robust estimation.

When interpreting the results of each panel in Table 5, the coefficients are equivalent to those of a binary logit model where categories 1 to m are coded as zero (as the base group) and categories $m+1$ to M are coded as one. Positive coefficients or odds ratios greater than one mean that higher values of an explanatory variable increase the probability of a farm moving to a higher category than the current one. Negative coefficients or odds ratios lower than one imply that the higher the value of the independent variable, the higher the probability of the farm staying in the current category or moving to a lower one (Williams, 2006).

According to the PPOM estimates, the variables *OPRICE*, *WEATHER*, *CAPDP*, and *OUTSOUR*, all of which have positive and statistically significant coefficients (i.e., odds ratios significantly above unity) in the three models, are factors that have a positive impact on olive farm income. That is, higher values of these variables (i.e., good weather conditions, high prices for olive oil, high CAP decoupled payments, and a large share of agricultural practices subcontracted) imply an increase in the probability of moving from categories C1 or C2 to the most viable category, C3, or simply staying in the latter. On the other hand, the variables *ICINT* and *FAMLAB* are also significant in all models, but their coefficients are negative (i.e., odds ratio significantly below one). This means a higher probability of the farm moving to a worse income category for higher values in these variables (i.e., intensive use of intermediate consumption inputs and family labour representing a high percentage of total farm labour).

Some independent variables do not have fixed coefficients in the two panels for each model: *SPEC*, *AGSUIT*, *CAPINT*, and *LANDOWN* in model C1; *SPEC, AGSUIT*, and *CAPINT* in model C2; and *FSIZE*, *AGSUIT*, *DEBEQRAT*, and *FAMLAB* in model C3. This means they do not meet the parallel lines assumption, and thus they can show a significant coefficient and OR estimate in one panel and non-significant ones in the other panel.

In the first model explaining the income dynamics of farms included in category C1 (non-viable), *SPEC* and *AGSUIT* show positive and statistically significant coefficients in both panels, with higher values in the second one, which reports coefficients related to categories C1 and C2 *vs*. C3. Consequently, the higher these two variables, the higher the probability of moving from the non-viable category to viable in the short term or viable in the long term, and the probability of changing from non-viable or viable in the short term to the viable in the long term category is even higher. Moreover, in this first model, the variable *LANDOWN* has a significant negative coefficient

Table 5 - Coefficients and OR estimates for PPOMs.

	C1: Non-viable		*C2: Viable in the short term*		*C3: Viable in the long term*	
C1 vs. C2 and C3	*Coef.*	*Odd Ratio*	*Coef.*	*Odd Ratio*	*Coef.*	*Odd Ratio*
AGE	0.006	1.006	0.039 **	1.039 **	0.028 ***	1.028 ***
AGTRAIN	0.380 *	1.462 *	-0.273	0.761	*-0.415*	*0.661*
FAMLAB	-2.541 ***	0.079 ***	-2.457 ***	0.086 ***	*-3.801 ****	*0.022 ****
LANDOWN	*0.119*	*1.126*	*0.371*	*1.449*	-0.407	0.665
FSIZE	0.004	1.004	0.002	1.002	*0.004*	*1.004*
AGSUIT	*0.002 ****	*1.002 ****	*0.000*	*1.000*	0.000	1.000
SPEC	*1.804 ****	*6.071 ****	*-0.864*	*0.422*	*2.891 ***	*18.007 ***
ICINT	-0.002 ***	0.998 ***	-0.002 ***	0.998 ***	-0.002 ***	0.998 ***
CAPINT	0.000	1.000	*1.52E-05*	*1.000*	0.000	1.000
OUTSOUR	5.021 ***	151.604 ***	3.199 *	24.516 *	5.619 ***	275.484 ***
DEBEQRAT	-0.176	0.838	46.253	1.2E+20	*-35.613 **	*0.000 **
OPRICE	1.659 ***	5.256 ***	3.003 ***	20.137 ***	3.530 ***	34.127 ***
WEATHER	0.622 **	1.862 **	1.492 ***	4.448 ***	0.774 ***	2.168 ***
CAPDP	0.002 ***	1.002 ***	0.002 ***	1.002 ***	0.002 ***	1.002 ***
Constant	*-5.258 ****	*0.005 ****	*-4.540 ***	*0.011 ***	*-5.578 ****	*0.004 ****
C1 and C2 vs. C3	*Coef.*	*Odd Ratio*	*Coef.*	*Odd Ratio*	*Coef.*	*Odd Ratio*
AGE	0.006	1.006	0.039 **	1.039 **	0.028 ***	1.028 ***
AGTRAIN	0.380 *	1.462 *	-0.273	0.761	*0.283*	*1.327*
FAMLAB	-2.541 ***	0.079 ***	-2.457 ***	0.086 ***	*-2.294 ****	*0.101 ****
LANDOWN	*-0.582 **	*0.559 **	*0.371*	*1.449*	-0.407	0.665
FSIZE	0.004	1.004	0.002	1.002	*0.008 ***	*1.008 ***
AGSUIT	*0.003 ****	*1.004 ****	*0.002 ***	*1.002 ***	*0.001 ****	*1.001 ****
SPEC	*3.606 ****	*36.815 ****	*2.351 **	*10.495 **	*2.891 ***	*18.007 ***
ICINT	-0.002 ***	0.998 ***	-0.002 ***	0.998 ***	-0.002 ***	0.998 ***
CAPINT	0.000 ***	1.000 ***	*-6.16E-05 **	*1.000*	0.000	1.000
OUTSOUR	5.021 ***	151.604 ***	3.199 *	24.516 *	5.619 ***	275.484 ***
DEBEQRAT	-0.176	0.838	46.253	1.2E+20	*7.446*	*1,713.6*
OPRICE	1.659 ***	5.256 ***	3.003 ***	20.137 ***	3.530 ***	34.127 ***
WEATHER	0.622 **	1.862 **	1.492 ***	4.448 ***	0.774 ***	2.168 ***
CAPDP	0.002 ***	1.002 ***	0.002 ***	1.002 ***	0.002 ***	1.002 ***
Constant	*-7.894 ****	*0.000 ****	*-9.942 ****	*0.000 ****	*-8.735 ****	*0.000 ****
N. of observations	1,219		343		732	
LR χ^2	326.92***		148.01***		269.10***	
Pseudo R^2	0.197		0.199		0.229	
Count	0.736		0.603		0.735	
LL at zero	-998.54		-371.56		-587.01	
LL at convergence	-802.96		-297.56		-452.46	
AIC	1,644.39		633.12		946.92	
BIC	1,746.51		706.03		1,043.43	

Note: Coefficients and Odds Ratios of explanatory variables that do not meet the parallel assumption are in italics.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

but only in the second panel, suggesting that the probability of moving directly to the viable in the long term category increases when the percentage of owned land increases. Finally, the variable *AGTRAIN* also shows significant coefficients, with the same positive value in both panels since this variable meets the proportional odds assumption in this model. This result means that a farmer with an agricultural degree managing a farm initially included in the non-viable category (C1) has a higher probability of the farm moving to a more viable one (C2, C3), while farms initially included in C2 also have the same higher probability of moving to C3.

Regarding the second model explaining the income dynamics of farms included in category C2 (viable in the short term), there are three variables which violate the proportional odds assumption showing significant coefficients (*SPEC, AGSUIT*, and *CAPINT*), although these coefficients are only significant in the second panel (related to categories C1 and C2 *vs*. C3). *SPEC* and *AGSUIT* have significant positive coefficients implying that higher values of these two variables (i.e., higher share of olive farming area and better local pedoclimatic conditions for olive production, respectively) increase the probability of the farm moving to the viable in the long term category (C3). The variable *CAPINT* shows a negative and significant parameter, indicating an increasing probability of moving directly to the viable in the long term category when the capital intensity increases. Additionally, in this second model, there is also a variable (*AGE*) meeting the proportional odds assumption showing positive and significant coefficients in both panels. This means that older farmers have a higher probability of staying in the C2 category or moving to the viable in the long term category (C3).

Lastly, the model explaining the income dynamics of farms included in category C3 (viable in the long term) also shows three variables that do not meet the parallel assumption and that show significant coefficients: *FSIZE*, *AGSUIT*, and *DEBEQRAT*. The variables *FSIZE* and *AGSUIT* have positive coefficients but they are only significant for the second panel, meaning a higher probability of staying in the viable in the long term category (C3) when these variables have higher values. The variable *DEBEQRAT* presents a significant parameter only in the first panel (related to categories C1 *vs*. C2 and C3) showing a negative value. Thus, a rise in the value of this factor (i.e., higher debt) increases the probability of moving from category C3 directly to category C1. In addition, two variables (*SPEC* and *AGE*) meeting the proportional odds assumption exhibit significant coefficients in this model. These two variables have positive coefficients, meaning that those olive farms with higher values for both variables have a higher probability of remaining in category C3.

Most of the results reported above are aligned with those found in the literature focused on other agricultural systems elsewhere. Thus, our PPOM estimates corroborate the crucial role in farm income dynamics played by off-farm uncontrollable factors such as the agricultural commodity prices (e.g., Baek and Koo, 2009; Zimmermann and Heckelei, 2012) and the subsidies granted by the CAP (e.g., Biagini *et al.*, 2020; Piet and Desjeux, 2021).

Moreover, the empirical results obtained also confirm that much of the interannual variations in farm income can be explained by farm-specific structural factors as the suitability of the farmland for agricultural production (e.g., Zimmermann and Heckelei, 2012; Allanson *et al.*, 2017), the farm's productive specialisation (e.g., Barnes *et al.*, 2020; Biagini *et al.*, 2020), the age of the farmer (e.g., Gloy and LaDue, 2003; Piet and Desjeux, 2021), or the farmer's managerial ability related to his/her agricultural training (e.g., Allanson *et al.*, 2017; Barnes *et al.*, 2020).

However, our results differ from other common findings in the literature. Probably the most notable discrepancy is that farm size did not yield significant coefficients (except in the second panel of model C3), contradicting evidence from many previous studies (e.g., Allanson *et al.*, 2017; Coppola *et al.*, 2020) showing increasing return to scale in farming production. Two circumstances could explain this divergence. First, it is worth noting that very small olive farms (those with an SGM of less than 8,000 Euros per year or "non-commercial" farms) are not included in the RECAN samples. Thus, our results just suggest that, for an economic dimension above the threshold to be considered a commercial farm, the differences in return to scale are rather small given that available olive production technologies can be adapted to a wide range of farm sizes. The second explanation is related to the insensitivity of olive production to labour and intermediate consumption (i.e., changes in labour and intermediate consumption cause little difference in the total output obtained). Our results suggest that smaller olive farms are more likely to opt for more intensive production (i.e., higher labour −usually family labour− and intermediate consumption input use) as a strategy to obtain higher output per hectare and thus compensate for any possible handicap regarding returns to scale. However, it has been proven that this is not an effective strategy since higher values in the variables *ICINT* and *FAMLAB* increase the probability of worsening the farm viability, which can only be explained by the farmers' undervaluation of the labour, land, and capital inputs they contribute to farming activities. Conversely, our results show a positive impact on farm viability of the outsourcing strategy (i.e., subcontracting more complex agricultural practices). To the best of the authors' knowledge, no such evidence on outsourcing has been reported before, although this finding could probably only be translated to other agricultural systems with a similar level of managerial complexity to modern olive production.

Finally, other factors commonly reported in the literature as affecting farm income, such as land ownership (e.g., Barnes *et al.*, 2015; Biagini *et al.*, 2020) or the farm business leverage (e.g., Gloy and LaDue, 2003; Allanson *et al.*, 2017), were not found to be significant in our case study. This divergence can probably be explained by specific characteristics of the Spanish olive sector, which largely relies on owned farmland (average *LANDOWN*=87.1%) and owned capital resources (average *DEBEQRAT*=1.5%).

4.3. *Scenario analysis*

The three PPOM obtained in the previous section allowed us to estimate the transition probabilities between farm categories ($\hat{p}^t_{ij}$) and the TPMs P^t for any year $t \geq k+1$. Thus, the BASELINE scenario and each of the six alternative scenarios proposed have been predicted following Equation (9), according to the different assumptions for the specific explanatory factors involved in the scenario analysis (*OPRICE*, *WEATHER*, and *CAPD*). Table 6 shows the farm distributions obtained once they became stationary. In all cases, the stationary distributions were reached after just 2 years (i.e., for $t = k + 2$).

The BASELINE scenario shows a farm distribution that is fairly well balanced between categories C1 and C3, although the share of farms that are viable in the long term is 3% higher. Thus, under the business-as-usual scenario, a clear duality is observed, with the commercial olive farms being evenly split into viable in the long term and non-viable categories. Bearing in mind the results detailed above, the former farms are those with more suitable farmland, more specialised in olive production and managed by older and better trained olive growers who avoid implementing excessive intensive production techniques while subcontracting more complex agricultural practices. The latter farms are those

Table 6 - Results for scenario simulations: farms distribution among profitability categories.

Scenario	*C1*	*C2*	*C3*
BASELINE	46.2%	4.3%	49.5%
OPRICE_2EUR	61.7%	5.2%	33.0%
OPRICE_3EUR	37.9%	4.9%	57.1%
WEATHER=80%	50.3%	3.7%	46.0%
WEATHER=60%	53.5%	3.0%	43.5%
CAP-30%	53.5%	4.9%	41.6%
CAP-50%+10HA	54.1%	4.0%	41.9%

that do not display said features.

Category C2 is practically non-existent, accounting for just 4.3% of the farms, a low portion also seen in all other scenarios (in all cases, C2 lies between 4.0% and 5.2%). This can be explained by the relative 'narrowness' of this category as explained above, meaning that in stationary states the farms tend to be classified into one of the two extremes categories (non-viable −C1− or viable in the long term −C3) according to their specific characteristics and the assumptions made in each scenario.

As expected, the variable capturing olive oil prices (*OPRICE*) causes the largest changes in farm distributions. Thus, in the OPRICE_2EUR scenario, the percentage of non-viable farms (C1) increases to 61.7%, while only one-third of the farms would remain viable in the long term. Conversely, in the favourable price scenario considered (OPRICE_3EUR), the percentage of non-viable farms (37.9%) would be the lowest among all the scenarios considered and, simultaneously, the highest percentage of farms that are viable in the long term would be reached (57.1%).

Regarding the variable *WEATHER*, a 20% worsening (i.e., WEATHER=80% scenario) leads to a 4 percentage point increase in non-viable farms compared to the BASELINE scenario, which corresponds to a 0.6 percentage point reduction in farms that are viable in the short term and 3.5 percentage point drop in ones that are viable in the long term. An additional 20% reduction in the *WEATHER* variable (i.e., WEATHER=60% scenario) would lead to a further increase of 3.2 percentage point in non-viable farms (7.2 percentage point increase over the BASELINE scenario), while category C3 would drop another 2.5 percentage point (6 percentage point decrease over the BASELINE scenario). All these estimations describing the potential impact of climate change suggest that Spanish olive farms are rather resilient, especially the third of the Spanish olive area under irrigation, where climate change impacts are expected to be minimised.

Finally, CAP payments also exert a considerable influence on the farms' distribution among income categories. A 30% decrease in the variable *CAPDP* (i.e., CAP-30% scenario) yields a similar influence to that caused by the WEATHER=60% scenario, with the same percentage of farms in category C1; however, the situation is worse in terms of farms that are viable in the long term, which decrease to 41.6% in this scenario (8.1 percentage point less than in the BASELINE scenario). These results can be taken as evidence that olive farm income is highly dependent on CAP subsidies. However, it is worth noting that in the scenario with a 50% reduction in the CAP decoupled payments variable, but with this cut not affecting payments granted for the first 10 hectares, the impact on farm distribution would be very similar to the CAPD-30% scenario, with variations below 1 percentage point in all three categories. This provides evidence that new CAP payments could be designed to minimise the impact of any support cut on the income distribution of olive farms.

5. Concluding remarks

This paper presents a relevant theoretical contribution relating to the analysis of farm income dynamics. By combining the Markov chain and ordinal regression models, the proposed approach allows us to determine which factors explain interannual changes in farm income (i.e., individual farm movements between income categories) and predict the impact of future scenarios on individual farms income (i.e., stationary income category). The empirical application of this approach to Spanish olive farms has shown it is sound and easily replicable for any other farming sector in the EU using the data provided by the FADN (or similar accountancy data networks in individual countries). Moreover, the empirical case study performed has also provided evidence that the results obtained using this approach are useful for ex-ante policy analysis supporting policy decision-making.

The empirical results obtained have shown, on the one hand, that interannual income variations in Spanish olive farms are determined by a combination of off-farm uncontrollable factors such as the price of olive oil, the annual weather conditions, and the CAP subsidies. Other factors also influencing the dynamics of these farms' income are: a) farm-specific structural features such as the agronomic suitability of the farmland and the farm's productive specialisation; b) farmer-specific characteristics such as age and agricultural training; and c) management factors such as the production intensity and the outsourcing strategy. However, it is worth remarking that farm size has not yielded significant results in the case of the Spanish olive sector, unlike what has been found in many other agricultural sectors.

On the other hand, the application of the method proposed to forecast the distribution of farms among income categories under several alternative scenarios has provided useful results that can support policy analysis. Future scenarios proposed show that the viability of Spanish olive farms is very sensitive to market conditions (i.e., olive oil prices), such that an increase (decrease) in olive oil prices contributes to a rise (reduction) in the proportion of viable olive farms. Likewise, the worsening of climatic conditions (i.e., decrease in olive yields) and policy support (i.e., reduction of CAP decoupled payments) leads to a decrease in the share of viable farms, although these two explanatory variables are less relevant than olive oil prices.

Regarding the limitations of the empirical analysis performed, it is worth mentioning the 2-3 years lag in the FADN data (i.e., the latest data currently available are for the year 2018). In this regard, a faster release of annual accountancy data could be very valuable.

Two avenues for future research are suggested. First, a comparative analysis with other agricultural sectors and/or other countries could be carried out. This could yield useful information for policy-makers to support the design and implementation of policy instruments, and also for farmers themselves, who would be able to distinguish between the uncontrollable and controllable factors that really influence income generation and competitive position. Second, from a methodological perspective, some refinements in the methodological approach could be tested (e.g., a farm typology with more viable categories), and the proposed approach could be extrapolated to analyse the dynamics of other farm outcomes (e.g., environmental performance measured using agri-environmental indicators).

Acknowledgement

This work was supported by the Andalusian Department of Economy and Knowledge and the European Regional Development Fund by the research project TRANSECOag [Grant PROYEXCEL_00459]. The authors thank the Spanish Ministry of Agriculture, Fisheries, and Food (MAPA) for providing the microdata files from the Spanish Farm Accountancy Data Network (RECAN) used for the research.

References

Abid F., Kaffel B., 2018. The extent of virgin olive-oil prices' distribution revealing the behavior of market speculators. *Review of Quantitative Finance and Accounting*, 50(2): 561-590. https://doi.org/10.1007/s11156-017-0638-9.

Allanson P., Kasprzyk K., Barnes A.P., 2017. Income mobility and income inequality in Scottish agriculture. *Journal of Agricultural Economics*, 68(2): 471-493. https://doi.org/10.1111/1477-9552.12192.

Arenas-Castro S., Gonçalves J.F., Moreno M., Villar R., 2020. Projected climate changes are expected to decrease the suitability and production of olive varieties in southern Spain. *Science of the Total Environment*, 709: 136161. https://doi.org/10.1016/j.scitotenv.2019.136161.

Argilés J.M., 2001. Accounting information and the prediction of farm non-viability. *European Accounting Review*, 10(1): 73-105. https://doi.org/10.1080=09638180120052221.

Baek J., Koo W.W., 2009. On the dynamic relationship between U.S. farm income and macroeconomic variables. *Journal of Agricultural and Applied Economics*, 41(2): 521-528. https://doi.org/10.1017/S1074070800002972.

Barnes A.P., Hansson H., Manevska-Tasevska G., Shrestha S.S., Thomson S.G., 2015. The influence of diversification on long-term viability of the agricultural sector. *Land Use Policy*, 49: 404-412. https://doi.org/10.1016/j.landusepol.2015.08.023.

Barnes A.P., Thomson S.G., Ferreira J., 2020. Disadvantage and economic viability: characterising vulnerabilities and resilience in upland farming systems. *Land Use Policy*, 96: 104698. https://doi.org/10.1016/j.landusepol.2020.104698.

Biagini L., Antonioli F., Severini S., 2020. The role of the Common Agricultural Policy in enhancing farm income: A dynamic panel analysis accounting for farm size in Italy. *Journal of Agricultural Economics*, 71(3): 652-675. https://doi.org/10.1111/1477-9552.12383.

Brant R., 1990. Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, 46(4): 1171-1178. https://doi.org/10.2307/2532457.

Cabezas J.M., Ruiz-Ramos M., Soriano M.A., Santos C., Gabaldón-Leal C., Lorite I.J., 2021. Impact of climate change on economic components of Mediterranean olive orchards. *Agricultural Water Management*, 248: 106760. https://doi.org/10.1016/j.agwat.2021.106760.

Cardone G., Bottalico F., Prebibaj M., 2021. Assessment of the economic sustainability of an organic olive oil farm in Puglia region (Italy) under the voluntary regional quality scheme. *New Medit*, 20(1): 113-129. https://doi.org/10.30682/nm2101h.

Chousou C., Tsakiridou E., Mattas K., 2020. A retrospective view of the EU policy reforms in the olive oil sector and future perspectives. In: Mattas K., Kievit H., van Dijk G., Baourakis G., Zopounidis C. (eds.), *Sustainable Food Chains and Ecosystems. Cooperative Management*. Cham: Springer, pp. 79-91. https://doi.org/10.1007/978-3-030-39609-1_6.

Coppola A., Scardera A., Amato M., Verneau F., 2020. Income levels and farm economic viability in Italian farms: An analysis of FADN data. *Sustainability*, 12(12): 4898. https://doi.org/10.3390/su12124898.

EC (European Commission), 2018. *CAP Specific Objectives Explained. Ensuring Viable Farm Income*. Bruxelles: DG AGRI, European Commission.

Finger R., El Benni N., 2021. Farm income in European agriculture: new perspectives on measurement and implications for policy evaluation. *European Review of Agricultural Economics*, 48(2): 253-265. https://doi.org/10.1093/erae/jbab011.

Fraga H., Moriondo M., Leolini L., Santos J.A., 2021. Mediterranean olive orchards under climate change: a review of future impacts and adaptation strategies. *Agronomy*, 11(1): 56. https://doi.org/10.3390/agronomy11010056.

Fraga H., Pinto J.G., Viola F., Santos J.A., 2020. Climate change projections for olive yields in the Mediterranean Basin. *International Journal of Climatology*, 40(2): 769-781. https://doi.org/10.1002/joc.6237.

Gloy B.A., LaDue E.L., 2003. Financial management practices and farm profitability. *Agricultural Finance Review*, 63(2): 157-174. https://doi.org/10.1108/00215060380001147.

Gontijo T.S., Rodrigues A.d.C., Muylder C.F.D., Falce J.L.l., Pereira T.H.M., 2020. Analysis of olive oil market volatility using the ARCH and GARCH techniques. *International Journal of Energy Economics and Policy*, 10(3): 423-428. https://doi.org/10.32479/ijeep.9138.

Gourieroux C., 2012. *Econometrics of Qualitative Dependent Variables*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511805608.

Hill B., Bradley B.D., 2015. *Comparison of Farmers' Incomes in the EU Member States*. Brussels: Policy Department B: Structural and Cohesion Policies, European Parliament. https://doi.org/10.2861/892960.

IPCC (Intergovernmental Panel on Climate Change), 2015. *Climate Change 2014: Synthesis Report*. Geneva: IPCC.

Júdez L., Rosario D.A., Mili S., 2017. Investigating the impacts of EU CAP reform 2014-20 and developments in sustainable olive farming systems. *New Medit*, 16(3): 2-10.

MAPA (Ministerio de Agricultura Pesca y Alimentación), 2021. *Cadena de valor del aceite de oliva virgen extra en el mercado español. Campaña 2019-2020*. Madrid: Ministerio de Agricultura, Alimentación y Medio Ambiente.

Mili S., Bouhaddane M., 2019. Delphi-based foresight of global olive oil market trends. Proceedings in System Dynamics and Innovation in Food Networks 2019. *International Journal of Food System Dynamics*, 113-132. https://doi.org/10.18461/pfsd.2019.1912.

Mili S., Bouhaddane M., 2021. Forecasting global developments and challenges in olive oil supply and demand: a Delphi survey from Spain. *Agriculture*, 11(3): 191. https://doi.org/10.3390/agriculture11030191.

O'Connell A.A., Liu X., 2011. Model diagnostics for proportional and partial proportional odds models. *Journal of Modern Applied Statistical Methods*, 10(1): 139-175. https://doi.org/10.22237/jmasm/1304223240.

O'Donoghue C., Devisme S., Ryan M., Conneely R., Gillespie P., Vrolijk H., 2016. Farm economic sustainability in the European Union: A pilot study. *Studies in Agricultural Economics*, 118(3): 163-171. https://doi.org/10.7896/j.1631.

Peterson B., Harrell F.E.J., 1990. Partial proportional odds models for ordinal response variables. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 39(2): 205-217. https://doi.org/10.2307/2347760.

Phimister E., Roberts D., Gilbert A., 2004. The dynamics of farm incomes: Panel data analysis using the Farm Accounts Survey. *Journal of Agricultural Economics*, 55(2): 197-220. https://doi.org/10.1111/j.1477-9552.2004.tb00093.x.

Piet L., Desjeux Y., 2021. New perspectives on the distribution of farm incomes and the redistributive impact of CAP payments. *European Review of Agricultural Economics*, 48(2): 385-414. https://doi.org/10.1093/erae/jbab005.

Poon K., Weersink A., 2011. Factors affecting variability in farm and off-farm income. *Agricultural Finance Review*, 71(3): 379-397. https://doi.org/10.1108/00021461111177639.

Rahelizatovo N.C., Gillespie J.M., 1999. Dairy farm size, entry, and exit in a declining production region. *Journal of Agricultural and Applied Economics*, 31(2): 333-347. https://doi.org/10.22004/ag.econ.15372.

RECAN (Red Contable Agraria Nacional), 2020. *Red Contable Agraria Nacional 2018*. Madrid: Ministerio de Agricultura, Pesca y Alimentación.

van der Zanden E.H., Verburg P.H., Schulp C.J.E., Verkerk P.J., 2017. Trade-offs of European agricultural abandonment. *Land Use Policy*, 62: 290-301. https://doi.org/10.1016/j.landusepol.2017.01.003.

Vrolijk H.C.J., De Bont C.J.A.M., Blokland P.W., Soboh R.A.M.E., 2010. *Farm Viability in the European Union: Assessment of the Impact of Changes in Farm Payment*. The Hague: LEI, Wageningen UR

Williams R., 2006. Generalised ordered logit/partial proportional odds models for ordinal dependent variables. *The Stata Journal*, 6(1): 58-82. https://doi.org/10.1177/1536867X0600600104.

Zimmermann A., Heckelei T., 2012. Structural change of European dairy farms – A cross-regional analysis. *Journal of Agricultural Economics*, 63(3): 576-603. https://doi.org/10.1111/j.1477-9552.2012.00355.x.

Zimmermann A., Heckelei T., Domínguez I.P., 2009. Modelling farm structural change for integrated ex-ante assessment: review of methods and determinants. *Environmental Science & Policy*, 12(5): 601-618. https://doi.org/10.1016/j.envsci.2009.01.014.