# Does a farm household income problem still exist in the European Union?

Rocchi B.[1*], Marino M.[1], Severini S.[2],

[1]Università degli Studi di Firenze Department of Economics and Management, Italy
[2]Università degli Studi della Tuscia, Department of Agriculture and Forest Science, Italy
*benedetto.rocchi@unifi.it

Paper prepared for presentation at the 7th AIEAA Conference
"Evidence-based policies to face new challenges for agri-food systems"

14-15 June 2018
Conegliano (TV), Italy

## Summary

*This paper compares income condition of farm and non-farm households using a single, harmonized database, the European Union Statistics on Income and Living Conditions (EU-SILC), a survey collecting information on income and well-being of a representative sample of all European households. We adopt approaches that are currently used in the literature on income analysis outside the specific field of agriculture and rooted within the family of the quasi-experimental methods. Starting from a simple comparison of income levels, we move to Regression Adjusted (RA) methods and to Covariate Matching (CM) techniques that are based on a "counterfactual" logic.*
*The analysis produces two main results. The first shows that on average, considering the whole European Union, a negative differential between farming and non-farming households exists. The gap is estimated after controlling for differences between farm and nonfarm households in family size, age, marital status, education, health, rural residence and country belonging. The result is robust to different empirical approaches and, in our opinion, represents a policy relevant outcome of the study. The second main achievement of the study refers to the heterogeneityof income differentials across countries and household groups. Within the group of self-employed households, the farming ones earn on average a lower income, but they show an income level above the average of the total populationin the Western Continental country.*

Keywords: Agricultural Households, Income, Covariate Matching, EU-SILC, European Union
JEL Classification codes: D31,I31, Q12, Q18

# Does a farm household income problem still exist in the European Union?

## 1. INTRODUCTION

One of the main and traditional justifications for the policy support provided to the farm sector is the assumption that the income of productive factors in agriculture is systematically lower than in other production activities.

Gardner (1992) refers to this as one of the main dimension of the "farm income problem". He reviews the theoretical justifications of such an assumption and the empirical evidence provided to support it. Finally, he concludes that the evidence of the low level of farm income is weak and shows that a tendency toward convergence exists between farm and non-farm income levels in the USA. A similar trend has been detected also in the European Union (EU), even though results are sparse and not fully comparable (e.g., Stefani et al., 2012; de Frahan et al., 2017). These results, if confirmed, should reorient the government 's role in the sector.

In the EU, while the Common Agricultural Policy (CAP) nowadays pursues a larger set of objectives other than farm income enhancement, the income disparity remains one of the main justifications of important policy tools. One of the CAP's objectives reported in the Consolidated Version of the Treaty on the Functioning of the EU[1] is "…to ensure a fair standard of living for the agricultural community, in particular by the increasing of the individual earnings of persons engaged in agriculture". The first pillar of the 2014-20 CAP budget dedicates indeed most of CAP funds to support and stabilize farm incomes. This is clearly the case of decoupled direct payments such as the Basic Payment Scheme or the Single Area Payment Scheme, which account for a large share of the CAP Pillar 1's budget. Some measures within the EU Rural Development Policy (RDP) support rural development by, for example, enhancing the take-over of young farmers. Finally, some other RDP measures are specifically aimed at supporting farm investments and innovation based on the idea that these strategies improve farm competitiveness.

In recent years the debate has shifted from *farm* to *farmer* income. This is because the agricultural households often have additional sources of income other than farming[2].This calls for addressing the analysis more on the total income of farm household rather than only on income from farming. Unfortunately, recent empirical investigations on farm household income disparity are limited and mostly focused on American data (Mishra et al., 2002; Hopkins and Morehart, 2004; El-Osta et al., 2007; Katchova, 2008; Peake and Marshall, 2009). The last comprehensive analysis for Europe refers to year 2001 (Eurostat, 2003)[3]. More recent (but partial) research has been developed by De Frahen et al. (2017) who analyse OECD countries and Stefani et al. (2012), Rocchi et al. (2012)and Severini and Tantari (2014) who workon Italian data.

[1]Article 39 of the Official Journal of the European Union, C 326 of the 26.10.2012.

[2]Mishra and Sandretto (2002) reported that 94% of farm households collect some type of nonfarm income. Mishra et al. (2002)showthat more than half of U.S. farm operators were employed off-farm, and 80% of them held full-time jobs off-farm. Indeed, according to many authors, the level of inequality between farm and nonfarm incomes has declined also because the decision of farmers to supplement their income with off-farm work (e.g.,Mishra and Goodwin, 1997; Mishra and Sandretto, 2002).

[3] In Europe there are not harmonized statistical sources on farm households income as in the case of US. This is the reason why the last comparative analysis among EU countries relies on a diversity of data and estimation methodologies (Eurostat, 2003), showing an overall convergence of incomes among agricultural and non-agricultural households. According to Hill and Bradley (2015) "...the evidence points to farmers notbeing a particularly low-income sector of society in most Member States judged on the basis of their household disposable incomes".

Katchova (2008) compares the economic well-being of US farm and nonfarm households focusing on two potential dimensions affecting the income differential among farm and nonfarm households: the degree of involvement in business activities and the family life-cycle stage. She finds that on average farm and nonfarm households show similar life-cycle patterns for income and wealth, even though nonfarm households exhibit more variation across different age groups. Peake and Marshall (2009) conclude that farm households report significantly higher income levels than other non-farm self-employment households. A key contribution of their work is using certain household characteristics (e.g., gender, education level, race, location, farm size) to explain the difference between farm and nonfarm self-employed income. Finally, De Frahen et al. (2017) suggest that on average low income is not a chronic problem among farm households in none of the ten surveyed OECD countries. However, they show that disparity exists among countries.

The above studies show that consensus on the presence of a farm household income problem cannot easily achieved. This can also be due to the different approaches and sources used. Katchova (2008) also states: “When income comparisons do exist, …. they are sensitive to the sources of information, the methods of estimation, and the definitions of incomes and farm households versus non-farm households that are used”. Moreover, some of the cited studies are somehow limited. For example, De Frahen et al. (2017) utilise the Luxembourg Income Study (LIS)database and rely on simple comparisons of income levels of farm and non-farm households, without accounting for the systematic differences that exist among these two groups as well as other methodological problems that our work aims at facing. Other studies explicitly account for household characteristics but consider only few of them. For example, Katchova (2008) disaggregates the farm household population according to the type of business involvement (i.e., residential, intermediate and corporate farms) and life stage of the agricultural household.

Considering as a starting point the research discussed above, we aim at estimating the income differential among farm and non-farm households within the EU. Our analysis adds to the previous literature in three ways.

The first is that we compare income condition of EU farm and non-farm households using a single, harmonized database. This is the European Union Statistics on Income and Living Conditions (EU-SILC), a survey collecting information on income and well-being of a representative sample of all European households. This dataset, as far as our knowledge, has not yet been used to investigate income conditions of farm households neither at the whole EU level nor by considering single Member States or groups of countries[4]. Because of differences in terms of income level, in our study we consider 3 groups of EU countries apart the whole EU: Western Continental, Central Eastern and Mediterranean EU counties.

The second contribution of our paper is to consider not only differences in observable characteristics but also the possibility that unobservable factors might affect both the decision of becoming a farmer and the income determination. A concern about self-selection and a potential endogeneity issue is relatively standard in the literature analysing income differential among sectors (e.g., Lucifora and Meurs, 2006; Glinskaya and Lokshin, 2007; Gimpelson and Lukiyanova, 2009). For such reason, we also test for the presence of these unobservable factors[5]using two alternative but slightly different empirical strategies.

The last contribution of this study is relying on approaches that are currently used in the literature on income analysis outside the specific field of agriculture, rooted within the family of the quasi-experimental methods. Starting from a simple comparison of income level, we then move to Regression Adjusted (RA) methods and to Covariate Matching (CM) techniques that are based on a "counterfactual" logic.

The results of the analysis support that the income of farm households is lower than the whole population, at the EU level. However, this is not the case for Western Continental EU countries where

[4]The only exception is the analysis for Italy proposed by Rocchi (2014).

[5]This is important to have unbiased estimates of the income differentials between the two groups under scrutiny.

income of farm-families is in line with the overall population. When comparing farm households only to nonfarm,self-employed ones, the results of the RA and CM analyses suggest that some of these differences can be explained by differences in the characteristics of the two groups of households. When accounting for observable family characteristics, the income differential strongly declines. This is particularly the case of Eastern Central countries and less of Mediterranean countries. Finally, results of the RA and CM analyses suggest that an income differential remain between the income of farm and non-farm self-employed households across the whole EU. This suggests that, in these cases, there is room for supporting the farm sector to improve their relative income conditions.

All these results provide recommendations for improving future analyses to be developed at the EU level. However, the empirical findings allow also some preliminary policy considerations that could fuel the debate on how to redesign CAP and RDP policies. This seems a very appropriate timing to do so giventhat a not negligible cut of CAP spending and a further reform of the CAP is now discussed in the EU. This may require a reorientation of the scarce resources between measures aimed at providing income support and measures aimed at other policy relevant objectives. Furthermore, the comparison among groups of EU countries allows verifying whether a redistribution of the CAP support among Member States is needed.

Section 2 presents the challenges associated with a proper estimation of the income differential among farm and nonfarm households and the methodology that we propose to face them. Section 3 shows the data, how we define the farm households and the relevant covariates used in the analysis. Section 4 comments the most important results and finally the paper concludes in Section 5 providing also tentative policy recommendations.

## 2. METHODOLOGY

### *2.1 Assessing income level differences between groups as a treatment effect*

A simple comparison of average incomes across the two sectors can show several limits. First, individuals working in different sectors might be systematically different in terms of their observable characteristics (age, education, etc.). Second, some unobservable characteristics might affect both the farm decision (being or not being a farmer) and the determination of the income, creating self-selection and a potential endogeneity issue related to the choice of the production sector.

The ideal setting to estimate such income differential would be an experiment that randomizes households dividing them in two groups: a first “treated” group whose main income source comes from the farm sector and a second “control” group whose income comes from a different sector. Phrasing it differently, working in the farm sector can be viewed as a specific treatment applied to a group of households, which in turn influences their incomes with respect to the other group.

Such experiment is clearly not feasible and the analysis has to rely only on observational data. A possible strategy is to mimic the experiment creating a counterfactual control group that has similar characteristics to the treated group according to the potential outcome framework. Since the main source of income of farm households comes from self-employed work, we compare farm households with other households whose main source of income comes also from self-employed work but it is generated within other sectors. Such approach can also be found in the empirical literature estimating inter-sectorial differences (e.g., Lucifora and Meurs, 2006; Glinskaya and Lokshin, 2007; Gimpelson and Lukiyanova, 2009).

We start by reporting the ‘raw’ or unadjusted average income differential in Europe between the two institutional sectors (farm and nonfarm households). Then to obtain robust estimates of the differential, as it is customary, we complement the analysis using several econometric approaches to properly take into account both observable and unobservable characteristics affecting the decision to farming and the income determination.

We start with a regression adjusted(RA) estimator that controls for observables characteristics. Such technique allows adjusting for the difference in well-known covariates that affect income.

Second, we take into consideration the presence of unobservable factors. To address it, we employ two alternative strategies. According to the first we perform a bivariate recursive mixed model with normal distributed error terms that jointly estimate the equation of the sector decision and the income determination (Roodman, 2011). The second strategy, known as a control-function approach, includes residuals from the estimation of the sector equation in the equation of the income (Wooldrigde, 2010). Section 2.3 gives more details on how such strategies are employed to control for the presence of unobservable factors

Finally, we use statistical matching techniques to identify a "counterfactual" control group of nonfarm households to be compared with the farm ones. To give strength to our empirical analysis, we employ covariate matching (nearest-neighbor matching) that imputes missing potential outcome for each subject by using an average of the outcome of similar individuals belonging to the opposite group. In other words, matching methods allow constructing comparison groups when the assignment to the treatment is done on the basis of observable variables.

Before going into details in these strategies, we show the basic assumptions of the potential outcome framework. To estimate the income differential between farm and nonfarm household, we need to address a missing-data problem, since we do not observe what a group of households would have earned if it would have belonged to the other group. Creating a counterfactual according to this framework provides a solution to the missing data problem, since potential outcomes can be generated (Rubin, 1974; Heckman and Navarro-Lozano, 2004; Imbens, 2004; Wooldridge, 2010).

The advantage of using the potential-outcome framework is that we could derive animportant parameter, the average treatment effect on the treated (ATET),whose interpretation is important for the aim of our work.

Consider two potential outcomes $y_{0i}$e $y_{1i}$, with the first indicating the outcome that would be obtained if the unit *i* does not get the treatment, and the second is the outcome that would be obtained if *i* gets the treatment. Generally, we only observe one of these variables for each *i* (missing data problem). Suppose f denotes our “treatment”, with $f_i = 0$ indicating households that work in the non farm sector, and $f_i = 1$ indicating householdsworking in the farm sector, then the ATET can be described as:

$$\mathrm{ATET}_i = \mathrm{E}\,(y_{1i} - y_{0i} | f_i = 1)$$

To estimate the ATET, we need to estimate the missing potential outcome for each observation in the sample.

Now, suppose that $x_i$ is a vector of covariates that affect the income level and $z_i$ is a vector of covariates that affect the probability of belonging or not to the farm sector, with the possibility that the two vectors have elements in common. The functional forms for $y_1$ and $y_0$ are:

$$y_0 = x'_0\beta_0 + \varepsilon_0$$

$$y_1 = x'_1\beta_1 + \varepsilon_1$$

where $\beta_0$and $\beta_1$are coefficients to be estimated, and $\varepsilon_0$and $\varepsilon_1$ are i.i.d. error terms that are supposed to be uncorrelated with the two above vectors of covariates.

Since $f_i$ is binary, its process can be described by the following:

$$\mathrm{f_i} = \begin{cases} 1 \text{ if } \mathrm{z}_i{}'\delta + \xi_i > 0 \\ 0 \text{ otherwise} \end{cases}$$

with $\delta$ being a coefficient vector of covariates and $\xi$ the error term that is supposed to be uncorrelated with $z_i$ and distributed by the standard logistic distribution (logistic regression) or by a standard normal distribution (probit regression).

Three main assumptions are behind the potential outcome model (as well as for the matching estimator). First, the conditional independence assumption (CIA): once conditioning on the covariates, $y_0$e $y_1$ are conditionally independent of f (selection on observables or unconfoundedness). In other words, observable covariates are supposed to sufficiently identify the effect of f with unobservable factors that affect f having no effects on the incomes and vice versa. Since such hypothesis is, generally speaking, a strong one, we will test it in Section 2.3.

Second, the overlap assumption that requires that for each possible covariate in the population, the following holds:

$$0 < p(\mathrm{f}) < 1$$

Finally, the last assumption claims that we must have an i.i.d. sample from the population[6].

To estimate the potential outcomes for each observation in the sample, the strategies that follow in the next subsections are implemented.

*2.2 Regression adjustment (RA)*

The regression adjustment estimator allows controlling for observable factors and generally is suitable when a weaker assumption than the conditional independence assumption (CIA) holds, namely the conditional mean assumption (CMA).

$$\mathrm{E}\ (\mathrm{y_{1i}}|\mathrm{f_i}, \mathrm{x}_i) = \mathrm{E}\ (\mathrm{y_{1i}}|\mathrm{x}_i)$$

$$\mathrm{E}\ (\mathrm{y_{0i}}|\mathrm{f_i}, \mathrm{x}_i) = \mathrm{E}\ (\mathrm{y_{0i}}|\mathrm{x}_i)$$

Since in our case the dependent variable is a continuous one, RA simply fits two OLS regressions, one for each group and then uses means of predicted outcomes for both farm and nonfarm households and computes the ATET.

The advantage of such approach is that we can easily assume unconfoundedness; conversely, we cannot easily see non-overlap problems. Moreover, if the difference between the average values of the covariates in the two groups is large, the results are sensitive to the linearity assumption.

Another advantage is that the regression adjustment model does not need to make any specific assumption on the model for the conditional probability of being farm. However, choice of covariates to be included in the model strongly affects the results (model dependence problem).

Despite these limits, the utility of the RA approach is well recognized, especially when we have good reasons to assume that the selection is made only on observable characteristics.

[6]Wooldridge (2010) shows that the estimation of the ATET requires a weaker form of conditional independence and overlap assumption (only for the “treated”) than the estimation of the "average treatment effect" (ATE). Furthermore, the ATET is closer to our research question about the income the farming households would earn if they were "not treated".

*2. 3 Testing for the presence of unobservable factors*

Unadjusted or regression-adjusted comparisons may provide misleading estimates of causal effects when we suspect that the assignment into the two groups is endogenous. This might happen because there are some unmeasured and/or unmeasurable factors affecting both the income and the decision of being farmer. The endogeneity is a violation of the conditional mean assumption so if we suspect its presence, we must deal with it.

We adopt two alternative strategies to check for the presence of endogeneity. The first relies on the control-function approach that allows unobservable variables to affect both the equation of being farm and the equation of income by allowing the residuals of the first equation to be included in the second equation as covariates (Wooldridge, 2010). Naturally, the presence of endogeneity must be tested before computing the ATET through the control function approach. For such reason, we perform a Wald test to determine whether the estimated correlations between the two models (income determination and farming decision) are different from zero. The rejection of the null hypothesis (that the correlation is zero) suggests endogeneity.

The second strategy is to estimate a bivariate recursive mixed model that, contrary to the first strategy, employ a multi-equation setting (Roodman, 2010). The model is bivariate since two equations are jointly estimated with correlated errors: the farm and the income equation. It is also mixed, since the farm model is estimated through a probit regression while the income equation through a simple OLS. Moreover, it is recursive since the equations can be arranged so that the matrix of coefficients of the endogenous variable, f, in one another's equations is triangular. In other words, there is a sequential (one-way from f to y) rather than a simultaneous (i.e. two-way) relationship. Once the bivariate recursive model is estimated, we similarly estimate separately the two models, namely assuming that the correlation among errors is zero. Then we run a likelihood ratio test (LR test) under the null hypothesis that the two models are nested. Following this approach, we determine which model is better suited to describe the data.

In such settings, exclusion restrictions are usually needed to improve the identificability of the model and properly testing the presence of unobservable factors. If the two models, the control function and the bivariate recursive mixed model confirm the presence of endogeneity, then estimates of the RA model are likely to be biased and we should estimate an endogenous treatment effect model.

*2.4 Covariate Matching*

The last method employs non-parametric covariate matching methods. The conditional mean and overlap assumptions must hold. But the advantage of CM is that they do not require specifying the functional form of the income equation and are therefore not susceptible to bias due to misspecification of the model (King and Nielsen, 2016). However, they can be biased if the treatment is endogenous.

The idea is to match a 'control' group of nonfarm households that are similar to farm households in all relevant pre-treatment characteristics. In such a way, observed income differences between the two groups can be attributed to the sector effect. Despite being propensity score matching more popular than CM to estimate the income differential among production sectors (e.g. Gimpelson and Lukiyanova, 2009; Glinskaya and Lokshin, 2007), we opt here for a nearest-neighbor matching estimators. The reason can be found in King and Nielsen (2016) who claim that covariate matching methods approximate a fully blocked randomization while a propensity score matching only approximate a complete randomization withthe former beingmore efficient than the latter.

We implement the method proposed by Abadie et al. (2004) and Abadie and Imbens (2011) who also consider the bias deriving from multidimensional covariates and adjust for it.

What matching estimators simply do is, for each farm household, imputing the missing outcome by using the average outcome of similar nonfarm households. Similarity between subjects is based on a weighted function of the covariates for each observation. In particular, the Mahalanobis distance is used here, in which the weights are based on the inverse of the covariates' variance–covariance matrix. Also, the inverse variance is used as a robustness check. We match a treated observation with one control observation and, for robustness, we repeat the analysis also matching one treated with 4 control observations.

## 3. DATA

The European Union Statistics on Income and Living Conditions (EU-Silc) is a harmonized and EU-wide survey that collects and compares multidimensional microdata on income, poverty, social exclusion and living conditions in several countries of Europe (Eurostat, 2007). The availability of a detailed set of data on income level and composition for a representative sample of European households allows overcoming several methodological difficulties.

First, we can adopt several definitions of the agricultural households' institutional sector (Hill, 1995; United Nations, 2012). This offers a more complete picture of the sector, since many farm families today rely in a different extent on off-farm incomes. We consider three definitions of farm households ranging from a broad one to a narrow one. While they are all useful, we then focus onthe narrow definition, since it is the most suitable to test the existence of a farm income problem.

Second, the available data provides several variables that can be used to test for the presence of endogeneity and to create a counterfactual group to be compared with the farm households group.

### *3.1 Farm households definitions*

EU-SILC data gives the possibility to create several definitions of farm households. In particular, we focus on the following three definitions (OECD, 2002; United Nations, 2012). The first is a broad definition and considers as farm households all those that have at least one self-employed member working in the agricultural sector. The second definition is a narrower one and considers only those families whose total farm income is at least greater of the total labor income coming from other activities carried out by the members of the household. Finally, the third is the usual narrow definition of the "agricultural household sector": it considers as farm households only those that have a farm income that is at least half of the total household income. The three definitions allow improving our understanding of the farm problem and give policy makers a more informative picture on the sector.

Before going into the detail of how we implement the three definition of farm household, it is important to stress that EU-SILC includes all European countries, but our analysis is limited to countries belonging to the European Union. To build the two groups, we first identified self-employed individuals since farm income is generated by self-employment. We used the EU-SILC variables PL030 (before 2008) and PL031 (after 2008), which report the self-defined current economic status. Both part-time and full time working individuals are considered. Once self-employed subjects have been identified, we need to screen among them those who work in agriculture. Farmers are identified using mainly two variables: the ISCO-88 classification to which the EU-SILC Regulation refers and the NACE classification used at EU-level for both economic and social statistics. While the first is useful to identify the individual main occupation, the second can be used to identify the sector in which the individual is employed. The rule used in our work is the following: among the self-employed, individuals are classified as farmer if they respond to EU-SILC ISCO variables PL050 (before 2011) and PL051 (after 2011) that they are Skilled Agricultural, Forestry and Fishery Workers.[7] Additionally, among the self-employed, we also consider as farmers those individuals who

[7] The ISCO code for this category ranges from 61 to 63. Some corrections are necessary for some countries since from 2008 the ISCO variable is coded only with one digit, but the same criteria above are used EXCEPT that isco=1, 5 and 9 is not included.

respond to EU-SILC NACE variables PL110 (before 2008) and PL051 (after 2008) that they work in the agricultural sector and whose main occupation is one among the following: Production and Specialized Services Managers, Personal Services Workers, Agricultural, Forestry and Fishery Labourers.[8]

Once the farmers are identified, the first definition of farm households is straightforward since they are identified if including at least one farm individual. The second definition needs to compute income figures at the household level. In particular, we compute first the level of farm income earned at the household level by summing the self-employed income of farmers[9], then we compute the total household income coming from the employee adults[10] who are member of the same household. A farm household is now identified if the household farm income is greater than the rest of the household income from labour. Finally, the last definition, which is the narrowest, considers as farm households only those earning a farm income that is at least the half of the total disposable household income.[11]

Farm households (whatever the definition adopted) have been identified within the sub-sample of self-employed households, i.e. all the households included in the EUSILC database earning at least some income from self-employed labour. The subsample of nonfarm households has been defined in the same dataset simply excluding all households that have already been identified as farm.

*3.2 Observable covariates and exclusion restrictions*

Once the farm and nonfarm households are defined according to the three definitions, one might compare incomes differential among the two groups. The income variable considered for the comparison is the level of equivalised disposable income (HX090)[12].

To reduce income differences depending on observable characteristics, we consider a standard set of variables recognized in the literature as determinants of household income(Becerril and Abdulai, 2010; De Janvry andSadoulet, 2001; Kassie and Shiferaw and Muricho, 2011; Mendola, 2007; Yúnez-Naude and Taylor, 2001). Many variables are computed with respect to the householder[13]. Once the householder is identified, we consider his/her age, education, if he/she is married, male or female and if he/she has a good or very good health status[14].

Other covariates are defined at the household level. In particular, we consider the number of components and if the household lives in a rural area.

A final remark is with respect to exclusion restrictions needed for the endogeneity tests. In fact, as explained in Section 2.3, we need some variables, which are likely to affect the decision to become farmer, but arenot likely to affect income determination. In the 2011 wave of the survey the special module on

[8]The Nace Sector considered are A (Agriculture, hunting and forestry) or A+B (where B is Fishing).

[9] The Eusilc dataset includes both gross and net individual self-employed income (PY050G/PY050N). We prefer the net self-employed income to define the household farm income, but several countries only have gross income data. However, this is not a real problem since the information are used not from countries comparison, but only for the identification of the farm families.

[10] To compute the income coming from employee income we consider Eusilc variable PY010G/PY010N which corresponds to employee cash or near cash income and PY020G/PY020N which correspond non-cash employee income.

[11]The last figure, total disposable household income (variable HY020), is computed summing not only incomes at the household level, but also pensions, benefits and allowances.

[12]The variable refers to the total disposable household income (HY020) divided by the equivalized household size. The equivalence scale used in the EUSILC survey is equal to 1 for the reference person, 0.5 for other adult members (14 years old or elder) and 0.3 for members up to 13 years old.

[13]The householder is defined using the following criteria: he/she must be responsible of the accommodation; if there are two responsible of the accommodation, the one earning the greatest income is considered; if still no householder is identified, we consider the eldest.

[14]The codes of the variables used in the analysis are available upon request.

"Intergenerational transmission of disadvantages" includes detailed information on parents' education, jobs and income. These latter pieces of information allow assessing if the parents of the farmer were also working in the farm sector. Thus, we create a dummy variable if at least one of the two parents were farmers. Farm businesses are often passed from one generation to the next within the framework of the family because the agricultural sector is typified by a strong heredity (de Haan, 1994; Haagsma and Koning, 2005). Some countries consider agriculture to be a 'closed profession' (Symes, 1990). The most common way of entry to farming is therefore succession in the family business (Zagata and Sutherland, 2015).

Table 1reports the sample size and the corresponding population of self-employed households represented in 2011when considering the third definition of farm households. Excluding the estimates of average income for the total population, all results presented in the following section have been produced using this sub-sample[15].

Table 1
Sample size and population represented
EU-SILC - year 2011

| | Self-employed | Farm | Non farm |
|---|---|---|---|
| | *Total EU* | | |
| Sample | 12 731 | 2 233 | 10 498 |
| Population | 12 038 353 | 1 567 906 | 10 470 447 |
| | *Western Continental* | | |
| Sample | 3 886 | 678 | 3 208 |
| Population | 4 682 703 | 442 623 | 4 240 080 |
| | *Central Eastern* | | |
| Sample | 3 782 | 998 | 2 784 |
| Population | 2 509 966 | 725 307 | 1 784 658 |
| | *Mediterranean* | | |
| Sample | 5 063 | 557 | 4 506 |
| Population | 4 845 683 | 399 975 | 4 445 708 |

## 4. RESULTS

### *4.1 Income differences over time and among EU countries*

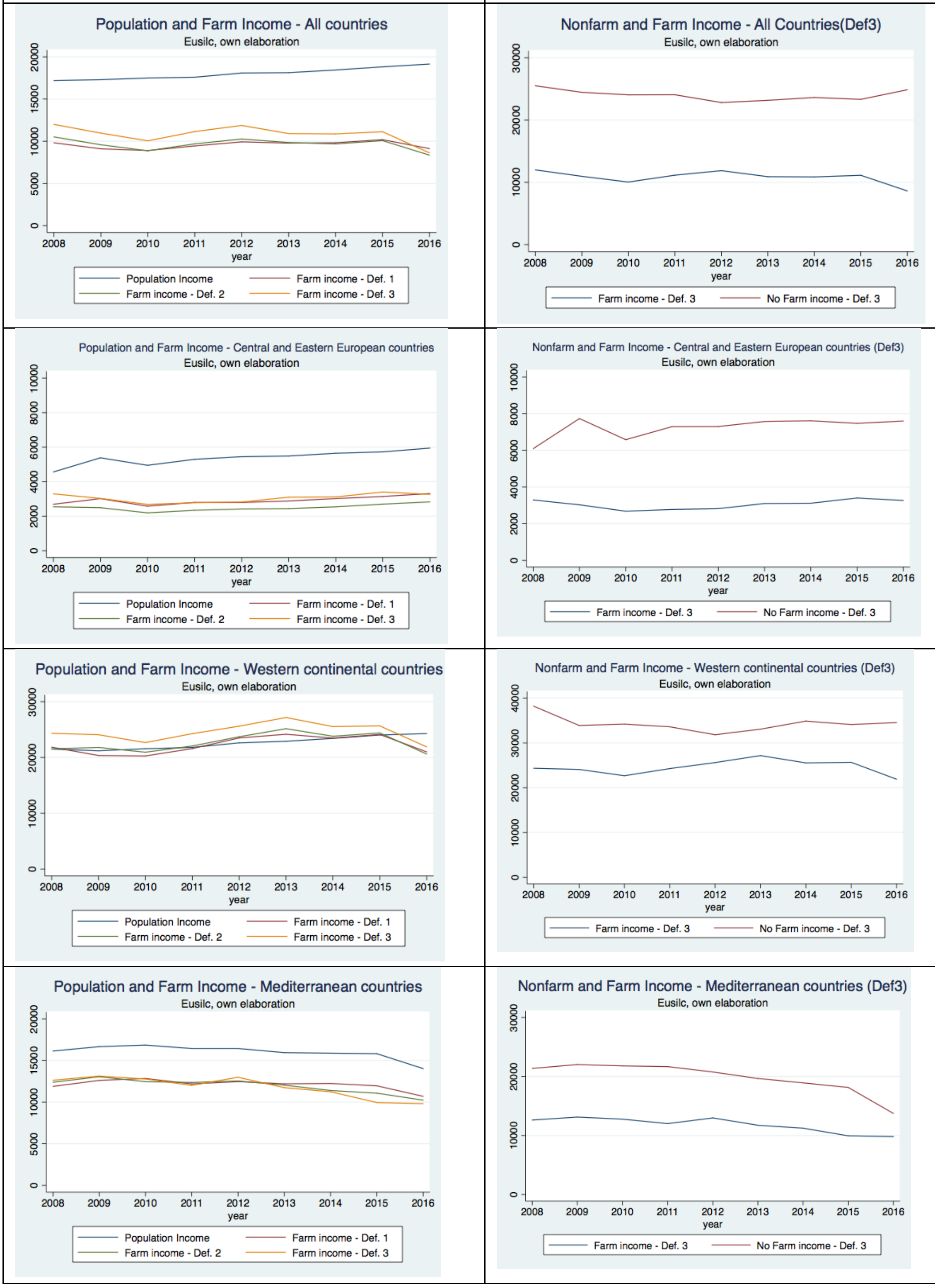
Figure 1 compares the evolution of farm households' average income compared to the overall EU population (left panels) and to non-farm self-employed households (right panels). We repeat the analysis also for three groups of EU countries, Western Continental, Mediterranean and Central Eastern. The data reported here refer to raw or unadjusted means.

A quick look at the graphs clearly suggests that a farm household's income problem still exists in the EU. However, the extent of such gap differs according to whether farm household income is compared with the whole population or only with the non-farm self-employed households. Differences also exist among EU countries.

[15]We used the R package *laeken* (Alfons and Temple, 2013) to recalibrate the weights and perform robust estimates of averages, considering the fact that we used only a sub sample of the EU SILC database.

Figure 1 Evolution of farm household income level total population (left) and only self-employed non-farm households (right), Current per capita equivalized disposable income (Euro), 2008 - 2016

| Overall population | Self-employed non-farm households |
|---|---|




Source: own elaboration on EUSILC data.

Farm household's income is generally lower when compared to the whole population at the EU level (left panels). This holds in Central Eastern and in Mediterranean EU countries but not when only Western

Continental countries are considered: in this case the income levels of farm and total population are very similar. Note that such differences are relatively stable over the whole considered period (i.e. 2008 – 2016)[16].

A larger income gap emerges when farm households (the narrower definition) are compared with the restricted, but more similar, group of non-farm self-employed households (right panels). The income differential between farm and non-farm self-employed is generally twice than that the one observed when farm household income is compared with the whole population. This is because the income of non-farm self-employed households is way larger than that of the overall population. At the EU level, income of farm households is around 60% that of the non-farm self-employed households over the whole considered period. However, the relative level of farm household income differs across the three groups of EU countries considered in the analysis: on average, it is around 70% in Central Eastern, 45% in Mediterranean and only 35% in in Western Continental EU countries.

These results suggest three main considerations. An income gap seems to exist but its extent largely depends on the group to which farm households are compared with. This suggests that results can be strongly affected bythe characteristics of the control group. Because of the nature of the farm business (i.e. an entrepreneurial activity), it seems more appropriate to compare it with the income of non-farm self-employed households. Second, the extent of the income gap differs among EU countries: this is true in terms of absolute as well as in relative levels. Hence, it seems important to account for the country location of households when comparing farm and non-farm households. Third, the differences between farmers and the control groups (population and self-employed) remain relatively stable over the considered period, suggesting that big differences should not be found by cross section analyses referred to different years.

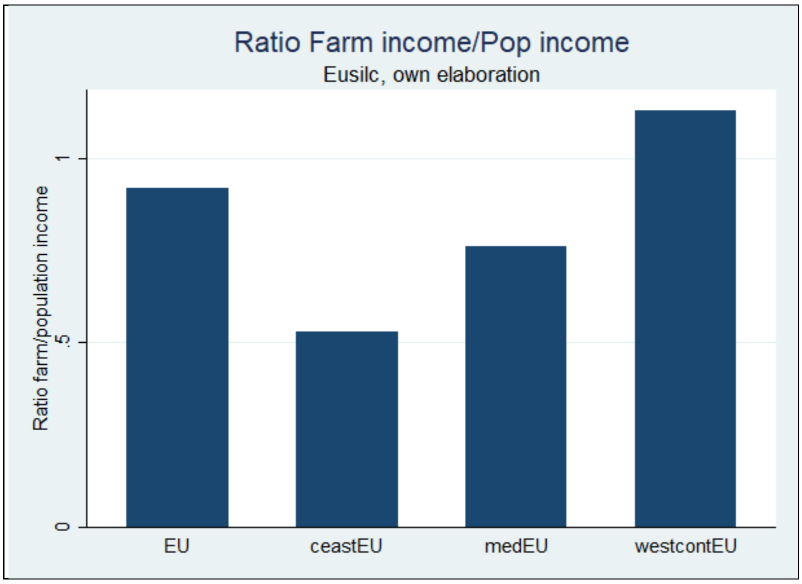
When focusing on 2011, which is our reference year for the reasons outlined in Section 3.2, we find that households whose income comes from self-employed work but not from the farm sector earn more than the average farm household.

The farm income problem can also be analyzed at the country level to shed light on the heterogeneity that exists among the different socio-economic realities of the European Union.

The two following graphs (Figures 3 and 4) consider only the "narrow" group of farm households (definition 3 above) and shows that country heterogeneity exists. For instance, when considering the ratio farm to total population income (Fig. 3), the group of Western Continental countries does not show any farm problem.

[16]Western Continental: Austria, Belgium, Germany, Denmark, Finland, France, Eire, Luxembourg, Netherland, Sweden and United Kingdom; Central Eastern: Bulgaria, Czech Republic, Estonia, Hungary, Lithuania, Latvia, Poland, Romania, Slovenia and Slovakia; Mediterranean: Cyprus, Malta, Greece, Spain, Italy and Portugal. According to the considered data, the Mediterranean countries show an average income of the total population that is just slightly lower than that of the overall EU, while the income in the Western continental countries is around 50% higher than the EU average and that of the Central Eastern countries is only 1/3 of the EU average. On the contrary, limited differences in income levels exist according to the definition of farm households both at the EU level and in the three groups of member States: in general, the income of farm households defined according to the narrow concept is higher than when the other two definitions are used.

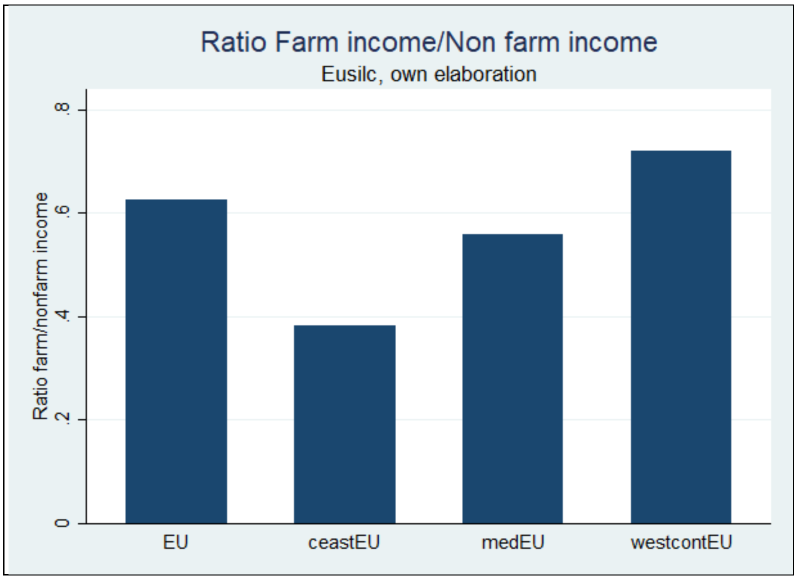
Figure 3
Farm households to total population income ratio by groups of countries
2011




Source: own elaboration on EUSILC data.

A different situation emerges when considering the ratio between the incomes of farm to other nonfarm self-employed households (Figure 4). All the three groups of EU countries show that households whose income comes from self-employment but not from agriculture earn on average more than farm households.

Figure 4
Farm households to other self-employed income ratio by groups of countries - 2011




Source: own elaboration on EUSILC data.

*4.2 Controlling for differences in households' characteristics*

So far, we showed the differences of incomes over time, among countries and across definitions and different types of households. We now focus on the differences on observable characteristics that might explain the gap between farm and (self-employed) nonfarm household income. Among them, we consider differences in mean across age, gender, marital status, education and health of the head holder. Moreover, we consider at the household level, the number of components and if they live in the rural area.

Table 2 reveals that significative differences in the average income of farm and non-farm households exist whatever observable characteristic we use to stratify the group of self-employed households.[17]

Table 2
Average income by group and household characteristics. Year 2011.

| | Non-farm | Farm | Difference |
|---|---|---|---|
| age16_40 | 20 051.79 | 11 307.82 | 8 743.97 |
| age 41_65 | 24 682.78 | 10 918.39 | 13 764.39 |
| age 66_80 | 45 920.90 | 14 789.15 | 31 131.75 |
| male | 24 052.81 | 11 103.87 | 12 948.94 |
| married | 22 341.55 | 9 258.27 | 13 083.28 |
| educ1 | 16 027.71 | 7 294.13 | 8 733.58 |
| educ2 | 17 303.39 | 8 806.74 | 8 496.65 |
| educ3 | 17 158.61 | 8 159.60 | 8 999.01 |
| educ4 | 19 502.35 | 11 688.69 | 7 813.65 |
| educ5 | 22 704.65 | 7 062.17 | 15 642.47 |
| educ6 | 34 313.77 | 20 427.88 | 13 885.89 |
| healhead | 23 865.94 | 10 158.63 | 13 707.31 |
| hhsize1 | 26 559.47 | 12 032.54 | 14 526.94 |
| hhsize2 | 29 996.36 | 13 527.15 | 16 469.21 |
| hhsize3 | 22 060.29 | 11 903.59 | 10 156.70 |
| hhsize4 | 19 304.33 | 9 396.16 | 9 908.17 |
| Rural residence | 17 269.45 | 9 012.41 | 8 257.04 |

Source: own elaboration on EUSILC data.

Table 3

| Differences in observable | Non-farm | Farm | Difference | |
|---|---|---|---|---|
| Age | 46.13 | 47.21 | -1.08 | *** |
| | | | (0.38) | |
| Male | 0.77 | 0.88 | -0.11 | *** |
| | | | (0.01) | |
| Married | 0.44 | 0.59 | -0.14 | *** |
| | | | (0.02) | |
| Education level | 3.42 | 2.64 | 0.79 | *** |
| | | | (0.04) | |
| Health status | 0.61 | 0.59 | 0.02 | |
| | | | (0.02) | |
| Household size | 2.63 | 2.96 | -0.33 | *** |
| | | | (0.04) | |
| Rural residence | 0.24 | 0.74 | -0.5 | *** |
| | | | (0.02) | |

Source: own elaboration of on EUSILC data.
Standard errors in parenthesis

[17] Parametric and non-parametric tests have been implemented and all the above differences are significative. Results are available upon request.

Table 2 suggest that the two groups, farm and nonfarm households, might have very different observable characteristics. In Table 3, we show the extent of such differences and we test if they are significative between the two groups. With the only exception of the level of head holder's health, the two groups appear to be quite different. For instance, among farmers, the head holders are generally older and less educated. The male and married proportion of farmers is greater than in the nonfarm household's counterpart. Finally, farm households are larger and they are (as expected) much more likely to live in rural areas.

Given the differences in observable characteristics between the two groups, the difference in income should be analysed adjusting for observables and, possibly, for unobservable characteristics. As explained in Section 3.1, we employ a specific empirical strategy in order to discriminate between the difference in income coming from such characteristics and the one coming from a "pure" sector effect. In particular, we start from a set of regressions that control from heterogeneity coming from covariates and from country belonging. The following Table 3 shows that controlling for differences in observable characteristics introducing covariates and countries dummies markedly reduces (82%) the income differential between the two groups. Still, an average difference in income of 2,286.41 Euro persists, highlighting the presence of a farm problem.

Table 4
Average income differentials between farm and non-farm households in the EU
Regression adjusting approach (ATET)

| | Estimated income differential | | [95% Conf. Interval] | |
|---|---|---|---|---|
| Regression, no covariates | -13 029.92 | *** | -14 349.92 | -11 709.93 |
| | (673.48) | | | |
| Regression with covariates | -5 341.49 | *** | -6 528.87 | -4 154.11 |
| | (605.82) | | | |
| Regression with covariates and country dummies | -2 286.41 | *** | -3 506.79 | -1 066.03 |
| | (622.66) | | | |

Source: own elaboration of on EUSILC data. Standard errors in parenthesis

The results of the above regression analysis couldbe affected by a potential endogeneity issue, since some unobservable characteristics could affect both the farming decision (being or not being a farmer) and the determination of the income (see Section 2.3 for details). We thus employ two estimation strategies to run the endogeneity tests and verify the presence of unobservables. The first relies on the control-function approach. The Wald test on the correlation between the two estimated models (income determination and farming decision) with a chi-square of 3.17 (p-value = 0.2047) does not allow rejecting the null hypothesis of noendogeneity.

The second test for endogeneity relies on the bivariate recursive mixed model**.** We run a likelihood ratio test (LR test) under the null hypothesis that the model with the correlated errors and the model with the restriction that the correlation between the two errors is equal to zero are nested. The null hypothesis could not be rejected with a LR chi square of 0.92 (p-value = 0.3369). Again, this confirms the absence of endogeneity. Since the two tests do not suggestthe presence of unobservable factors, then the results of the RA model are likely to not be biased.

The final step of our empirical strategy is the implementation of non-parametric covariate matching methods that allow imputing the income differential to the sectorial gap. Although we find that there are no unobservable factors in our data and the estimates of the RA are likely to not be biased, in the CM approach farm households are strictly matched with a 'control' group of nonfarm households with similar observable characteristics[18]. As a consequence, the later approach is likely to provide a more accurate estimate of the income differential.

We use two metrics to measure the distance among covariates (Inverse matrix and Mahalanobis distances) and two different numbers of matches (one and four) to be made for each treated observation.

The difference is larger when considering the inverse matrix distance, which uses only the diagonal matrix of the inverse sample standard errors (Table 5). Conversely, the income differentials are not very different from the ones found with the RA approach when considering the Mahalanobis distance. Given the metric, the number of matches doesn't affect the results in a significant extent[19].

Table 5
Average income differentials between farm and non-farm households in the EU
Covariate matching approach (ATET)

| Metric | Estimated income differential | | [95% Conf. Interval] | |
|---|---|---|---|---|
| Inverse Matrix (1:1) | -3 056.52 | *** | -5 022.54 | -1 090.50 |
| | (1 003.09) | | | |
| Inverse Matrix (1:4) | -3 055.72 | *** | -5 021.74 | -1 089.70 |
| | (1 003.09) | | | |
| Mahalanobis (1:1) | -2 319.20 | ** | -4 326.22 | -312.19 |
| | (1 024.01) | | | |
| Mahalanobis (1:4) | -2 319.50 | ** | -4 326.52 | -312.49 |
| | (1 024.01) | | | |

Source: own elaboration of on EUSILC data - Standard errors in parenthesis

*4.3 Going beyond averages*

The obtained results clearly suggest that in 2011 a relevant income gap still existsbetween farm and non-farm self-employed households. However, the extent of such gap differs according to the considered groups of countries and the approach used to compare income of these two groups of households.

The figures in table 6 refer to a set of results disaggregated by groups of countries and obtained adopting different methodologies(raw means, regression adjusted and covariate matching comparisons). Raw mean values suggest that at the EU level and in the three considered groups of countries, a relevant income gap exists. The income of farm-household is slightly less than 50% that non-farm self-employed households at the EU level. The income differential is higher than the EU average in Central Eastern countries while lower in the other two groups. Indeed, the Western Continental countries show the lowest gap, with the farm households' average income level around 70% of non-farm households.

Also results of the RA and CM confirm that an income gap exists. However, as seen before, accounting for household characteristics strongly reduces the extent of the gap. At the whole EU, RA and CM results suggest that the farm household income is between 79% and 82% of non-farm household's income.

[18] In the RA approach, conversely, the "treatment" effect is calculated imposing to the farmer households the average impact of observable characteristics on income estimated using the control group of non-farming self-employed households.

[19] We also produced diagnostic and test statistics to assess whether the matching balances covariates in the two groups. When considering the average treatment over the treated, the weighted standardized differences are almost all very close to zero and the variance ratios are almost all close to one.

Table 6. Income differentials between farm and non-farm self-employed households in 2011.Whole EU and country groups.

| | Farm Households Euro | Non-Farm Households Euro | Farm - Non Farm ratio |
|---|---|---|---|
| | Raw mean values | | |
| Whole EU | 11 121.03 | 24 150.95 | 0.460 |
| Central Eastern | 2 841.11 | 7 332.85 | 0.387 |
| Mediterranean | 11 974.41 | 21 605.08 | 0.554 |
| Western Continental | 23 917.82 | 33 899.05 | 0.706 |
| | Regression Adjusted (RA*) | | |
| Whole EU | 11 121.03 | 13 407.43 | 0.829 |
| Central Eastern | 2 841.11 | 3 541.20 | 0.802 |
| Mediterranean | 11 974.41 | 14 698.97 | 0.815 |
| Western Continental | 23 917.82 | 30 127.07 | 0.794 |
| | Covariate Matching (CM*) | | |
| Whole EU | 10 870.59 | 13 682.41 | 0.794 |
| Central Eastern | 3 072.69 | 4 079.86 | 0.753 |
| Mediterranean | 12 380.57 | 15 352.37 | 0.806 |
| Western Continental | 23 544.23 | 29 529.85 | 0.797 |

Source: own elaboration on EUSILC data.
*Data used in CMand RA estimates refer only to the farm households observations. RA reports the estimated coefficients for the full specification (covariates and country dummies), while the CM reports the estimated coefficients computed with the Mahalanobis (1:1).

Similar values are estimated also in the three groups of countries, with the Mediterranean one showing the highest value of the ratio. In the Western Continental group the RA and CM estimates yield results that are closer to those obtained calculating the raw mean values. This can be interpreted as if the income gap between farmers and the other self-employed households in Western Continental EU countries is due for the largest part to a "pure" sector-effect, i.e. the fact that incomes from self-employed labour of the farmers' group are earned with agriculture. An additional insight gained by the results of RA and CM is that differences in terms of relative income gap between farm and non-farm households among the three groups of EU countries decline.

Differences also exist in the results comparing RA andCM. In the case of CM, the average income gap in absolute terms results higher at the EU level and also in the country groups. Conversely the ratio between farm and non-farm income slightly decreases at the EU level and also in Central Eastern and Mediterranean countries; while remain almost the same in the Western Continental group.It should be taken into account that, differently from raw means, the estimation according to the RA and CV approach considers only the subsample of farm. Referringonly to farm households to estimate the "treatment effect" of being farmers (controlling for all other household characteristics) seems more appropriate than using the whole sample of self-employed households, given that the underlying research question refers to the potential changes in the income level a farm household could face if leaving the sector[20].

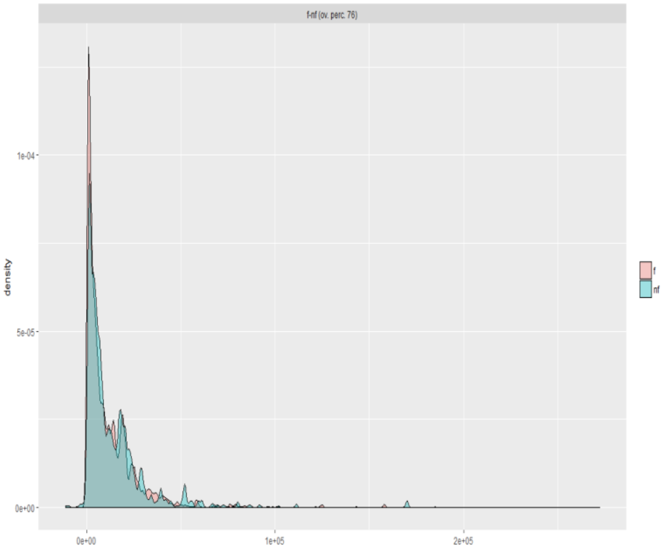
A further advantage of using the CM approach is that the production of a "counterfactual" subsample of observations of non-farming households allows a direct comparison of the distribution of income within the two groups of farm and non-farm households. Figure 5 display the kernel density functions representing

[20] According to paragraph 2.1, the RA and CV estimates can be considered a measurement of the ATET of "being farmer".

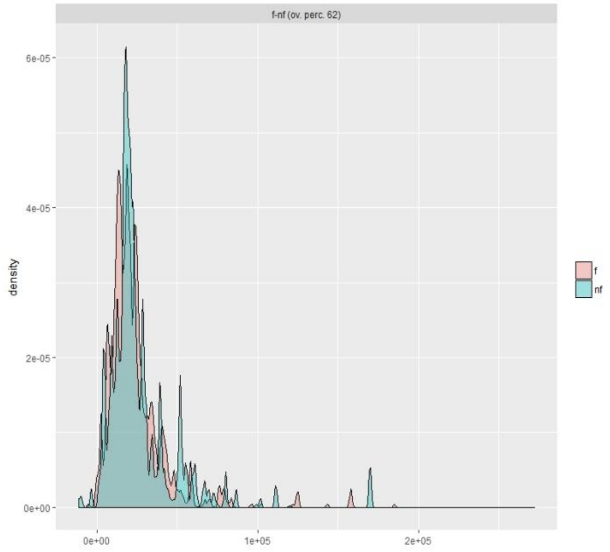
the distribution of households in the two groups according to the income level[21] while Table 7 provides some summary statistics on the distributions.

Figure 5
Overlapping of income density functions for farm and non-farm households
2011 – All EU and groups of countries

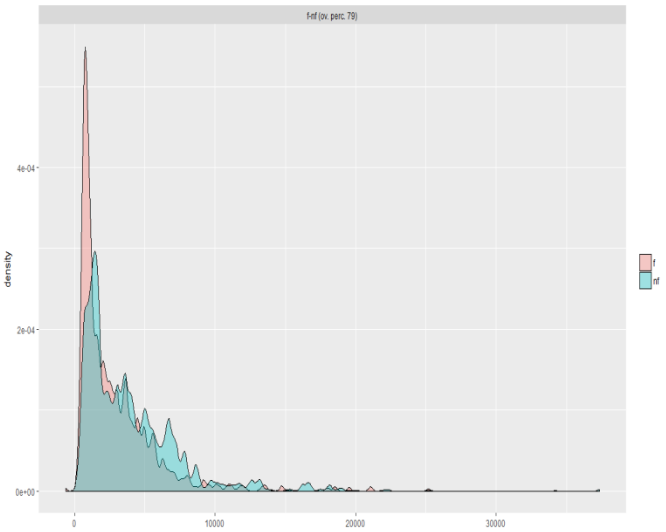
All EU



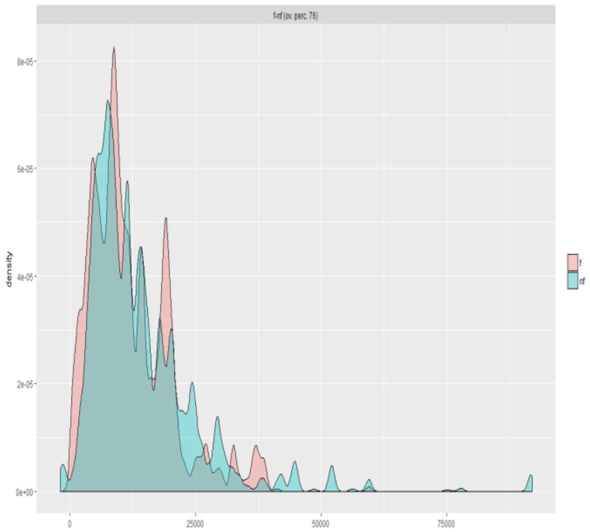

Western Continental EU




Central Eastern EU




Mediterranean EU




Source: own elaboration on EUSILC data.

[21] The graphs have been produced with R using the package "overlapping". The package estimates a kernel density function with Gaussian smoothing kernel and Silverman's "rule of thumb" determination of the bandwidth (Silverman, 1986).

Table 7
Income distribution within the farm and non-farm groups
Summary statistics by groups of countries

| | All EU | Western Continental | Central Eastern | Mediter-ranean |
|---|---|---|---|---|
| overlapping (%) | 75.9% | 62.4% | 79.2% | 75.5% |
| max_f | 272 276.81 | 272 276.81 | 26 627.55 | 64 728.11 |
| max_nf | 169 987.00 | 169 987.00 | 37 358.47 | 92 000.00 |
| mode_f | 851.58 | 17 711.09 | 763.92 | 7 553.30 |
| mode_nf | 1 343.92 | 18 668.90 | 1 460.72 | 8 769.95 |
| median_f | 6 026.00 | 19 630.00 | 1 917.10 | 10 004.65 |
| median_nf | 7 624.34 | 20 674.13 | 3 061.79 | 11 193.07 |

Source: own elaboration on EUSILC data.

The total overlapping of the two (farm vs. non –farm) distributions in the All EU case is equal to 76% and varies from a minimum of 62 % (Western Continental) to a maximum of 79% (Central Eastern) when different groups of countries are considered. The left-sided skewness of the distribution of households within the farm group is clearer in the Central Eastern group of countries while both in the Mediterranean and even more in the Western Continental group of countries, the presence of farm households along the whole field of variation of incomes is more evident. The values of the three summary statistics (mean, median and mode) for the farm household's case are closer, in relative terms, to those of the non-farm control group in the Western Continental Countries. Interestingly, the highest value of income is recorded in a farming household in the Western Continental countries. Even though this information does not allow any generalization from a statistical point of view (given the small number of observations included in the right tail of the distribution), it represents a piece of evidence about the complexity of the farm household problem. In other words, despite the asymmetry towards the left tail of the distribution within the group of farm households, in all countries farmers are also present in the highest quantiles of income distribution, even when compared with a similar group of households. This may suggest the presence of a heterogeneity within farm households that is likely to affect the distributive outcomes of farm income support measures.

## 5. CONCLUSIONS

In this study we addressed the farm household's income problem with the aim of providing an empirical test of its existence in the European Union. The use of a harmonized pan-European survey on incomes and living condition of families (EU SILC) allowed a systematic comparison of income levels of farm and non-farm households across all the 27 EU Member States. To our knowledge this is the first time that such analysis is carried out for European Union.

Within the EU SILC survey (year 2011), a sub sample of about 2,133 farm households were identified according to the prevalence of income from self-employed labor in agriculture within the total household income. The income level of the farm households was compared with that of other households using a variety of empirical approaches. Specific efforts were carried out in the comparison to take fully into account all the factors that could influence income differentials among different household groups.

Two main results of the analysis should be stressed. The first shows that on average, considering the whole European Union, a negative differential between farming and non-farming households still exists. The estimated value of the average differential is between 2,000 and 3,000 Euros in 2011, depending on the methodology used in the analysis. The gap is estimated after controlling for differences between farm and nonfarm households in the family size, age, marital status, education, health, rural residence and, last but not least, country. The result is also robust to different empirical approaches and, in our opinion, represents a policy relevant outcome of the study.

The second main achievement of the study refers to the diversification of income differentials across countries and household groups. The direct comparison between the average income of farm households and that of the total population yields different results when looking at different groups of countries. While the average income of farm households is significantly lower in the Central Eastern countries and Mediterranean countries of the EU, in the group of Western Continental countries the incomes of farm families are in line with those of the rest of population. For Western Continental countries the income differential between farm and non-farm households becomes evident only when farmers are compared with a control group of self-employed households that is similar in terms of observable characteristics. Within the group of self-employed households, the farming ones earn on average a lower income also in the Western Continental country, even though self-employed households (including the agricultural ones) on average show an income level *above* the average of the total population. This result holds true not only in 2011 but also along the whole 2008 - 2016 period considered in the analysis.

Taken together the two results shed light on the dual nature of the farm income problem. The comparison with a self-employedgroup opens a space for a genuinely "sector" (agricultural) policy. "Living of agriculture" as a self-employed household in Europe still means earning incomes that are lower than other self-employed households. As the figures in table 5 show, a relevant part of these differentials can be explained by households’ characteristics, such as age or education attainment. Such differences can be partially addressed by policy measures implemented within the second pillar of the CAP (incentives to retirement, support to young farmers, professional training). Such policy seems relevant mainly for Mediterranean and Central Eastern member countries of the EU, where farming families show characteristics that are quite different from those of the rest of the population, even when compared only with the group of the self-employed households. The residual income differential, after controlling for the household characteristics, is likely to depend more on differences between agriculture and other production activities in terms of efficiency in the use of production factors. Policy measures for enhancing productivity in agriculture, such as those traditionally implemented under the "structural" side of the CAP (and currently included in the second pillar) seem the best suited tools to address this component of the farm households’ income problem.

According to our results such sector gap is present within all 27 EU members. But the comparison of farm households’ incomes with those of the total population shows that, at least in some EU members, farming families still belong to the lowest deciles of income distribution, with incomes lower than the whole country average. Such second component of the farm household income problem has been traditionally addressed by CAP income support measures aiming at promoting "...a fair standard of living for the agricultural community". Income support has been one of the major justifications and is included among the goals of the CAP from the very beginning. During the first decades of the common policy, the goal of income support was pursued with "sector" policy measures aimed at increasing "...the earnings of persons engaged in agriculture". For example, the support of agricultural prices improved the economic conditions of farmers but it also leaded to wide impacts on the structure and the competitiveness of the productive sector. Later, the decoupling of support under the first pillar has progressively separated the support to *farms* from the support to *farmers*, making clearer the nature of income support of direct payments received under the CAP. Our results show that a policy justification for supporting the *total* income of farming families still exists in Europe but it is no longer a relevant issue across all EU members and among the whole agricultural households' institutional sector. The upcoming reform of the CAP should address in a more effective way the problem of the targeting of direct payments (and the Regulation proposal recently delivered by the Commission does it). But any area-based determination of payments (whatever the improvements achieved after the reform) is unlikely to completely fix the bias in the distribution of payments towards larger farms (and richer households). From this point of view a renewed debate should be promoted about the opportunity of including *general* social objectives (such as income support of a socially disadvantaged group) among the goals of a *sector* policy like CAP.

The analysis here proposed presents some limits and only partially exploits the information provided by the EU SILC survey. The main limit of this study is the implementation of the full analysis of income differentials for only one year. The choice of studying only 2011 was motivated by the availability of a specific information (gathered *una tantum* within the EU SILC survey) allowing us to test for the endogeneity in the model of income determination used in the analysis. We rejected the hypothesis of endogeneity and we assume that such result can be generalized for the whole available time series. We will develop our study repeating the analysis of income differentials along the whole period 2008-2016. A further development will consider also the possibility of exploiting the panel structure of the EU-SILC survey, taking better into account heterogeneity among member states and individuals. Finally, given the presence of a large set of non-monetary variables on living conditions in the EU-SILC database, we will extend the analysis towards a comprehensive well-being comparison between farm households and the rest of society.

## References


Abadie, A., &Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics*, 29(1), 1-11.

Abadie, A., Drukker, D., Herr, J. L., &Imbens, G. W. (2004). Implementing matching estimators for average treatment effects in Stata. *Stata journal*, 4, 290-311.

Alfons A.& Templ M. (2013). Estimation of Social Exclusion Indicators from Complex Surveys: The R Package laeken. *Journal of Statistical Software* 54(15): 1-25.

Becerril, J., &Abdulai, A. (2010). The impact of improved maize varieties on poverty in Mexico: a propensity score-matching approach. *World Development*, 38(7), 1024-1035.

De Frahan B. H., Dong B. J. and De Blander R (2017), Farm Household Incomes in OECD Member Countries over the Last 30 Years of Public Support. In Mishra AK, Viaggi D, Gomez y Paloma S, *Public Policy in Agriculture. Impact on Labor Supply and Household Income*. London: Routledge.

De Janvry, A., &Sadoulet, E. (2001). Income strategies among rural households in Mexico: The role of off-farm activities. *World Development*, 29(3), 467-480.

El‐Osta, H. S., Mishra, A. K., & Morehart, M. J. (2007). Determinants of economic well‐being among US farm operator households. *Agricultural Economics*, *36*(3), 291-304

Eurostat (2003). Income of agricultural households sector - 2001 Report. Luxembourg: Office for official publications of the European Communities.

Eurostat (2007).*Comparative EU statistics on Income and Living Conditions : Issues and Challenges*, Proceedings of the EU-SILC conference, Helsinki, 6-8 November 2006. Luxembourg: Office for Official Publications of the European Communities.

Gardner, B. L. (1992). Changing economic perspectives on the farm problem.Journal of economic literature,30(1), 62-101.

Gimpelson, V.;& Lukiyanova, A. (2009), Are Public Sector Workers Underpaid in Russia? Estimating the Public-Private Wage Gap, IZA, DP 3941, January.

Glinskaya, E., &Lokshin, M. (2007). Wage differentials between the public and private sectors in India. *Journal of International Development*, 19(3), 333-355.

Haagsma, R. and Koning, N. (2005). Endogenous norms and preferences and the farm income problem. *European Review of Agricultural Economics*, 32(1), 25-49.

Haan, H. de (1994), *In the shadow of the tree: kinship, property, and inheritance among farm families*. Thesis - Landbouwuniversiteit, Wageningen, Wageningen.

Heckman, J., & Navarro-Lozano, S. (2004). Using matching, instrumental variables, and control functions to estimate economic choice models. *Review of Economics and Statistics*, 86(1), 30-57.

Hill B., Bradley B.D. (2015). *Comparison of farmers' incomes in the EU Member States*. Study commissioned by the European Parliament's Committee on Agriculture and Rural Development. IP/B/AGRI/IC/2014-68 June 2015. Available at http://www.ceasc.com/Images/Content/Comparison%20of%20farmers'%20incomes%20EP%20version.pdf

Hill, B. (1995).Farm incomes, wealth and agricultural policy(No. Ed. 2). Avebury, CABI Publishing.

Hopkins, J., and Morehart, M. (2004). Assessing Farm Household Well-Being – Beyond Farmers and Farm Income. *Amber Waves* 2(8): 8.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, 86(1), 4-29.

Kassie, M., Shiferaw, B., & Muricho, G. (2011). Agricultural technology, crop income, and poverty alleviation in Uganda. *World Development*, 39(10), 1784-1795.

Katchova, A. L. (2008). A Comparison of the Economic Well-Being of Farm and Nonfarm Households. *American Journal of Agricultural Economics.*, vol. 90, no. 3, pp. 733–747.2008.

King, G., R. Nielsen. (2016). *Why Propensity Scores Should Not Be Used for Matching. Working Paper*. Available from http://j.mp/1sexgVw.

Lucifora, C., & Meurs, D. (2006). The public sector pay gap in France, Great Britain and Italy. *Review of Income and wealth*, 52(1), 43-59.

Mendola, M. (2007). Agricultural technology adoption and poverty reduction: A propensity-score matching analysis for rural Bangladesh. *Food Policy*, 32(3), 372-393.

Mishra, A. K., El-Osta, H. S., Morehart, M. J., Johnson, J. D., & Hopkins, J. W. (2002).*Income, wealth, and the economic well-being of farm household*s. USDA, Washington. https://www.ers.usda.gov/webdocs/publications/41451/31359_aer812fm.pdf?v=41271

Mishra AK and Sandretto CL, Stability of Farm Income and the Role of Nonfarm Income in U.S. Agriculture. *Review of Agricultural Economics*, vol. 24, no. 1, pp. 208–221, 2002.

Mishra, A. K., & Goodwin, B. K. (1997). Farm income variability and the supply of off-farm labor. *American Journal of Agricultural Economics*, 79(3), 880-887.

OECD (2002). *Farm Household Income Issues in OECD Countries: a Synthesis Report* AGR/CA/APM(2002)11/FINAL, Paris, 29 Oct. 2002.

Peake, W. O., & Marshall, M. I. (2009). *Has the "Farm Problem" Disappeared? A Comparison of Household and Self-Employment Income Levels of the Farm and Nonfarm Self-Employed*. In Southern Agricultural Economics Association Annual Meeting, January 31-February 3, 2009, Atlanta, Georgia(No. 46304). https://ideas.repec.org/p/ags/saeana/46304.html#download

Rocchi B. (2014). Vivere di agricoltura. Le famiglie agricole nella distribuzione del reddito.*Agriregionieuropa*, 36, 31-34.

Rocchi B, Stefani G, Romano D, and Landi C (2012) Are Italian farming households actually poorer than other non agricultural households? An empirical analysis. Are Italian farming households actually poorer than other non agricultural households? An empirical analysis using different sources of microeconomic information," Paper presented at the first AIEAA Conference, Trento (IT) 4-5 June 2012 https://ageconsearch.umn.edu/bitstream/124128/2/Rocchi%20Stefani_Are%20Italian%20farming%20households%20actually%20poorer%20than%20other%20non%20agricultural%20households.pdf

Roodman D. (2011), Fitting fully observed recursive mixed-process models with cmp, *Stata Journal*, 11, (2), 159-206

Rubin, D. B. (1974), ''Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies,'' *Journal of Education Psychology* 66, 688–701.

Severini, S., & Tantari, A. (2014). The contribution of different off-farm income sources and government payments to regional income inequality among farm households in Italy. Bio-based and Applied Economics, 3(2), 119.

Silverman, B. W. (1986). *Density Estimation*. London: Chapman and Hall.

Stefani G., Rocchi B., Romano D. (2012). Does agriculture matter? Revisiting the farm income problem in Italy. Dipartimento di Scienze per l'Economia e l'Impresa. iWorking Paper n. 18/2018. https://www.disei.unifi.it/upload/sub/pubblicazioni/repec/pdf/wp18_2012.pdf

Symes, D.G (1990), Bridging the generations: Succession and inheritance in a changing world, *Sociologia Ruralis*, 30, 280-291.

United Nations (2012). *Statistics on rural development and agricultural household income. The Wye Group Handbook*. Second Edition. Retrieved from http://www.fao.org/economic/ess/ess-capacity/wyegroup/it/

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (Vol. 2). Cambridge, MA: MIT Press.

Yúnez-Naude, A., & Taylor, J. E. (2001). The determinants of nonfarm activities and incomes of rural households in Mexico, with emphasis on education. *World Development*, 29(3), 561-572.

Zagata, L., Sutherland, L.-A. (2015), Deconstructing the 'young farmer problem in Europe': Towards a research agenda, *Journal of Rural Studies*, 38, 39–51.