Impact of Geographical Indications for Food Products on Farm Income in the EU

Kathrin Poetschki

December 21, 2018

Impact of Geographical Indications for Food Products on Farm Income in the EU

Kathrin Poetschki

December 21, 2018

Wageningen University, NL

Registration number 931115 659 110

MSc Management, Economics and Consumer Studies

Course AEP 80430, MSc Thesis Agricultural Economics and Rural Policy

Supervisor and first examiner: dr. ir. Jack Peerlings

University of Bonn, DE

Registration number 302 40 40

MSc Agricultural and Food Economics

Course M401, Master Thesis

Second examiner: prof. dr. Thomas Heckelei

Source image cover page: based on a private photograph taken by the author

Preface

It was not hard for me to find an interesting topic which I wanted to write my Master thesis about. I grew up on the farm of my grandparents, and I look back on many years working on different farms. On the one hand, farming is essential for everyone's life. On the other hand, agriculture is often accused of being responsible for environmental degradation or the violation of human and animal rights. Overfertilisation and its consequences, deforestation, displacement of indigenous peoples, modern forms of slavery, or mistreatment of animals are only some of the popular reproaches. Farmers and their families suffer from missing gratitude and understanding shown by consumers. Expectations, requirements and the administrative burden are increasing, but at the same time, willingness to pay for sustainable and high-quality food seems to be limited. At least, the producer prices are often volatile and too low to make this work and loads paying off. Consequently, the number of farms in Europe is decreasing, while existing farms tend to increase in size to remain competitive. And again, consumers do not hesitate to show their discontent with such a development.

The image on the cover page of this thesis is based on a photograph that has accompanied me during my studies. It reminds me of my home and my roots, and it reminds me of my grandfather who concludes every debate about agriculture and agricultural politics with the same sentence: "Maket es allen recht!". This is Low German and basically means that whatever a person (farmer, politician, consumer) does, it is hard to make everyone happy because there will always be someone complaining. However, it was this recurrent phrase that caught my attention and aroused my curiosity. Throughout my studies, I was searching for the best way to feed the world, the most sustainable way of farming, the most sustainable consumer behaviour and the fairest way to trade and share resources across the globe. I was looking for an answer to the question of what kind of farming is accepted by consumers and allows farmers to earn a living.

Certification schemes such as GIs offer the opportunity to both lend credence to the quality (and sustainability) of the product or production process and differentiate from mass production. Thus, it is also meant to increase farm gate prices. However, I am aware that the jungle of labels and certifications that cover the packing of our products also leads to confusion, frustration and mistrust among consumers. Conversations I had with farmers conveyed the impression that they are not fully convinced by the benefits of certification schemes and GIs either.

This thesis was written as part of a joint Double Degree Master programme of the University of Bonn and the University of Wageningen. At this point, I would like to thank dr. ir. Jack Peerlings for his support, patience, and for the countless meetings and discussions we had that improved my thesis. I also appreciated the help and feedback of prof. dr. Thomas Heckelei, dr. Liesbeth Dries and dr. ir. Koos Gardebroek, as well as the inspiring lectures by dr. ir. Maarten Voors. I am very thankful for having gained experiences in quantitative impact assessment, although it was not always easy and stretched me to my own limits on some days. Last but not least, I owe special thanks to my family for all the support they gave and the opportunity to come home and clear my mind whenever I needed a change of scenery. Finally, after more than five years of studying, I have to admit that my grandfather was right – There is no perfect way of farming, and agricultural politics always have pros and cons. Life is a compromise. This also seems to apply to quantitative impact assessment. Throughout the last months, I found out that it is less perfect and objective than I had expected.

Declaration

I hereby affirm that I have prepared the present paper self-dependently, and without the use of any other tools than the ones indicated. All parts of the text, having been taken over verbatim or analogously from published or not published scripts, are indicated as such. The thesis has not yet been submitted in the same or similar form, or in extracts within the context of another examination.

Place, date of submission ....Wesel, December 21, 2018........

Student's signature ....

Summary

This Master thesis builds up on a quantitative impact assessment of geographical indications (GIs) for food products on farm income in the EU. It focusses on four GI schemes: PDO, PGI, TSG and mountain products. GIs are intended to benefit disadvantaged farms who are unable to compete on the global market. Further, they are expected to stimulate rural development by increasing the viability and resilience of farms in disadvantaged and remote areas. As part of the Strength2Food project, the EU is interested in the effect of GI adoption on farm income, which has not been investigated so far.

The analysis is mainly based on data from the Farm Accountancy Data Network. Few regional characteristics from EUROSTAT were added to the dataset. The impact assessment was done for quality wine specialists and olives specialists both for the years 2014 and 2015. 2014 data considers PDO and PGI labels, whereas 2015 data also covers information about the TSG label and mountain products.

First, potential effects of GIs on farm income are outlined to illustrate that GIs do not necessarily increase farm income. For farmers who produce final PDO or mountain products, income effects are more likely to be positive due to restricted market entry and limited threats to farmers' market power from downstream players of the supply chain. If income effects are negative in the long run, and farmers behave as profit-maximisers, they are expected to stop using GIs.

After a broad discussion of five popular estimation techniques used for impact analysis, an endogenous switching regression (ESR) model was chosen to estimate the income effect by full information maximum likelihood (Stata command *movestay*). Descriptive statistics illustrate the differences between treated and untreated farms. On average, GI olives specialists have a higher farm net income than their non-GI colleagues, although the difference is not significant. For wine specialists, non-GI farms earn significantly more. According to the ESR results, the estimated effect of GIs on farm net income of wine specialists in 2014 is -21303 EUR for treated farms. Untreated farms would have earned 33991 EUR more if they had adopted GIs. While the average treatment effect for GI olives specialists is estimated to be -43196 EUR, the estimated average treatment effect for untreated farms is 1767 EUR. The results confirm self-selection of farms as well as different responses of treated and untreated farms to changes in the control variables (heterogeneous impacts). The chosen estimation technique was able to account for these problems.

The estimates contradict the expectations based on economic theory since adopters are assumed to only adopt GIs if they do not decrease farm profits. However, it is possible that production costs increase relatively more than revenues. Further research could investigate the effect on revenues and costs separately. From a theoretical perspective, it is also not expected that treatment effects for non-adopters are positive and significantly higher than for adopters. Estimation results were compared to those of other estimation techniques to show that estimated effects based on the chosen variables and data are sensitive to the choice of the estimation (and matching) technique. Poor data is seen as one limitation and potential reason for the contradicting and varying estimates. Many farms had to be excluded from the sample because they did not report information about their GI adoption. In addition, the four GI schemes could not be analysed separately. Further, baseline data was not available, so reported data of GI adopters might have been influenced by GI adoption. This compromises the quality of impact estimates. Nevertheless, the thesis gives insights into mechanisms by which GIs can affect farm income. It further elaborates on estimation techniques and ways to improve the reliability of estimated effects.

Keywords: Income effects, product differentiation, geographical indications, PDO, PGI

Table of contents

Figures

Tables

Abbreviations

ATC	Average total costs
ATE	Average treatment effect
ATT	Average treatment effect on the treated
ATU	Average treatment effect on the untreated
AWU	Annual working unit
BH	Base heterogeneity
ESR	Endogenous switching regression
EXT	Externalities, for a description of variables see appendix III
FADN	Farm accountancy data network
FHC	Farm household consumption, for a description of variables see appendix III
FNI	Farm net income, for a description of variables see appendix III
GDPC	Gross domestic product per capita, for a description of variables see appendix III
GI/GIs	Geographical indication/s (here: PDO, PGI, TSG and mountain products)
LFA	Less favoured area, for a description of variables see appendix III
LIA	Liabilities, for a description of variables see appendix III
MA	Mountain area, for a description of variables see appendix III
MACH	Machinery, for a description of variables see appendix III
MC	Marginal cost
MKM	Kilometres of motorway, for a description of variables see appendix III
MKM2	Square of km of motorway, for a description of variables see appendix III
MR	Marginal revenue
OLS	Ordinary least squares
ORG	Organic, for a description of variables see appendix III
OUT	Output, for a description of variables see appendix III
OVER	Overheads, for a description of variables see appendix III
PaRCI	Partial randomization with common impact
PaRVI	Partial randomization with varying impact
PuR	Pure randomization
PDO	Protected designation of origin
PGI	Protected geographic indication
PL	Paid labour, for a description of variables see appendix III
PS	Propensity score
PSM	Propensity score matching
SPC	Specific cost
TSG	Traditional speciality guaranteed
TH	Transitional heterogeneity
UAA	Utilised agricultural area, for a description of variables see appendix III
UL	Unpaid labour, for a description of variables see appendix III

1. Introduction

Agricultural income is lower than the average income in other sectors (European Commission, 2009). On average, public support provides 32% of EU farm income (European Commission, 2017a). This share is larger for small farms and less favoured areas (LFA) (Hill & Brandley, 2015). Geographical indications (GIs) are part of the food quality schemes that have been supported by the EU since 1992 (European Union, 1992). Their goal is to create added value by linking food products to unique physical characteristics, the environment, social ties and/or traditions of their origin (Giovannucci et al., 2009). GIs are part of "the 'quality turn' in the economy [towards] more differentiated, localized and eco-friendly products and forms of economic organization" (Hajdukiewicz, 2014, p. 4). Since it is an alternative to cost-minimizing strategies, it is expected to especially benefit small farms and those in disadvantaged areas, who have difficulties to compete with larger and more efficient producers (Hajdukiewicz, 2014). Moreover, GIs offer opportunities for endogenous development in rural areas if more value added remains at the farm level and, consequently, in rural areas (Gangjee, 2017). Thus, GIs are assumed to improve the relative income position of farmers.

This thesis will focus on four GIs:

- Protected Designation of Origin (PDO): all ingredients must come from and all production steps of a food product need to take place at a specified area (European Commission, 2017b). [1]
- Protected Geographic Indication (PGI): at least one production step (production, processing or preparation) needs to take place at the specified area (European Commission, 2017b).
- Traditional Speciality Guaranteed (TSG): products are produced in a way that is typical or traditional for the specified area, but the product and its ingredients can be produced anywhere (European Commission, 2017b).
- Mountain products: Regulation (EU) No. 665/2014 encompasses detailed descriptions of what kind of products can be called mountain products (European Commission, 2014).

There is an increasing demand for local, traditional and more extensively produced food (Verbeke et al., 2012). Products with GI labels are one answer to these consumption trends. Therefore, GIs present a strategy to increase the economic viability of farm enterprises. GI application is linked to product differentiation strategies, which allow to obtain price premiums (Giovannucci et al., 2009; Van Ittersum, 2002). Product differentiation leads to imperfect competition, which generates market power (sometimes also referred to as pricing or bargaining power) and higher profits for producers (Krugman & Wells, 2013). Firms no longer face a perfectly elastic demand function as with perfect competition of mass-produced goods. GIs allow farmers to produce products that cannot be perfectly substituted. However, production of GI products is sometimes linked to higher production costs, e.g. for registration, application of specifications, marketing and control, which might exceed extra revenues (Hajdukiewicz, 2014). Another potential threat to income gains is too little market power of farmers vis-à-vis downstream stakeholders in the supply chain (traders, processors, retailers), who do not pass on the higher profits that are earned from product differentiation.

This research is part of the 5-year EU-funded project Strength2Food that, amongst other objectives, aims at evaluating the impact of the EU food quality policy and related schemes (Strength2Food, 2016). So far, no research has investigated the impact of GIs on farm income. This thesis will focus on the impact of the above-mentioned GIs on farm income to learn more about their contribution to rural economies in the EU. The hypothesis is that GI application leads to higher farm income. Four questions will help getting insights into the effect of GIs on farm income and rural development:

[1] Source pictures PDO, PGI and TSG labels: https://ec.europa.eu/agriculture/quality/schemes_en

1. What potential effects can GIs have on farm income?
2. What estimation techniques can be used to estimate the impact of GIs on farm income?
3. How do farms producing GI products differ from other farms with respect to farm characteristics?
4. Is there a causal effect of GI uptake on differences in farm income?

A literature review gave insights into potential impacts of GIs on farm income. Further, it helped understand the role of market structures, how GIs can influence market power, and which factors determine the adoption of GIs. Extensive literature review was also conducted to study different estimation techniques that can be applied to assess the impact of GIs. Based on the gained knowledge, an endogenous switching regression model was chosen for the impact evaluation of GIs on farm income.

The quantitative analysis is mainly based on data taken from an unbalanced panel from the Farm Accountancy Data Network (FADN) for the years 2014 and 2015. Three variables about regional characteristics at NUTS2 level were added to this dataset. The respective data was taken from a dataset generated by Van de Pol (2017), which was based on EUROSTAT data. An extensive descriptive data analysis was conducted to discover differences between GI adopters and non-adopter. Based on the literature review, dependent and explanatory variables were chosen from the dataset. For the impact evaluation, the endogenous switching regression was estimated by full information maximum likelihood using the Stata command *movestay*. The model allowed for endogenous self-selection on both observed and unobserved characteristics. The results were also compared to estimated causal effects of other estimation techniques, such as Ordinary Least Squares (OLS) and Propensity Score Matching (PSM).

The thesis is organised as follows: Chapter 2 discusses determinants of income and uptake of GIs based on a literature review. Special attention is paid to the mechanism by which GIs are intended to increase farm income. Chapter 3 briefly introduces to the key problem of impact assessment as well five popular estimation techniques. It starts with the simplest model and builds up to more complex models with stricter assumptions. Chapter 4 introduces the sources of data and samples used for this research. Further, the final estimation technique is chosen and the final model for the impact evaluation with all required variables is specified. Finally, chapter 4 also contains the descriptive statistics. Chapter 5 reports the results of the impact assessment. A general discussion and final conclusions are presented in chapter 6.

2. Theory

The goal of this chapter is to explain how and under what conditions GIs influence farm income. Further, it presents additional categories of variables that affect farm income and those influencing the uptake of GIs. Such variables are relevant for estimating causal effects of GIs on farm income. Section 2.1 gives an overview of determinants of farm income. Section 2.2 discusses the mechanisms by which GIs affect farm income. It also shows why the degree of market power and, consequently, the extent of potential income effects can vary among GI farmers. Section 2.3 elaborates on factors that influence the decision to adopt GIs.

2.1 Determinants of farm income

If you want to estimate the effect of GIs on farm income, you need to know what other factors explain deviation in farm income. The better you control for other determinants, the better the estimated effect of GIs will be. Farm income mainly relies on profits generated from producing and selling agricultural output. For simplicity, taxes and subsidies are ignored for now, so the focus is on revenues and costs linked to agricultural production. Profits are the difference between total revenue and total cost. Equation (2.1) presents a model for short-term profit maximization, where π equals profit, p is the price received for the yield y that is sold, s represents the cost (i.e. shadow price) for quasi-fixed labour (L), n is the shadow interest rate or cost for quasi-fixed capital (C), r is the shadow price for the quasi-fixed land[2] (A), FC refers to fixed cost, and SPC are specific variable production costs (e.g. seeds, fertiliser).

$$\pi = \max_{x,y,L,C}\left(py-(FC+SPC+sL+nC+rA);\, T(y,L,C,A),\, p,\, s,\, n,\, r \gg 0\right) \qquad (2.1)$$

Revenue is determined by production volumes and farm gate prices. Production volumes depend on the amount of inputs used and the efficiency by which they are used or processed to new products that can be sold on the market. Efficiency is influenced by natural or geographical constraints like climate, soil fertility or gradient (Van de Pol, 2017). The amount of inputs used depends on their relative price compared to the expected farm gate price for the final product. Large farms benefit from economies of scale and potential volume discount when buying inputs or paying for services. Thus, larger farm size is negatively correlated with input prices. Apart from real costs, there are also opportunity costs. Farmers are not only profit-maximisers. They also maximise utility, which can put certain constraints on the amount of labour and capital used for farming. A household model can help understand why farmers do not necessarily maximise farm profits only. Farming is not necessarily the only livelihood strategy that contributes to household income. Other productive activities and leisure of household members require labour and capital, which cannot be used to maximize profits earned on the farm. Opportunity costs of working on the farm increase if employment opportunities outside the farm business are offering a higher income or a more attractive work, which is more likely the closer the farm is located to urban areas (Meraner et al., 2015). Thus, labour and capital used for farming are competing with other productive and non-productive activities. Access to capital and interest rates affect the use of capital on the farm (Beckmann & Schimmelpfennig, 2015). A farmer faces price and income volatility, which depends on the (combination of) products he is producing as well as exposure to risks such as weather extremes (Organisation des Nations Unies pour l'alimentation et l'agriculture, 2011). The higher the volatility of a farm's profits are, the more expensive bank loans become as interest rates increase (Organisation des Nations Unies pour l'alimentation et l'agriculture, 2011). This reduces the likelihood that farmers invest in their business. Consequently, they become relatively less efficient compared to those who invest in machinery and innovative production techniques. This reduces their competitiveness and market power. In contrast, farm income is positively affected by the farmers decision to hedge prices or involve in any other form of risk management like insurances, because it reduces volatility in farm profits and interest rates. Land prices influence the affordability of and, consequently, the access to land, which in some cases becomes a limiting factor of production (Beckmann & Schimmelpfennig, 2015). In addition, institutional and legal constraints might pose limitations to profit maximization. For example, farmers who apply for farm payments from the EU must fulfil requirements (i.e. Cross-Compliance and Greening), which are often meant to increase ecological sustainability of farming. These requirements affect farm profits via the amount of

[2] Livestock could play a role as well, I will ignore it here for simplicity.

inputs used for farming. Finally, the quantity of products sold on the market is directly affected by farm household consumption of own products. It reduces the revenue, although it might be welfare improving if it is cheaper to consume own products than buying them in the supermarket (European Commission, 2011).

Total cost basically depends on prices for inputs, quantity of inputs and fixed costs. Specific costs also depend on the amount of inputs used, which is influenced by their price(s) and opportunity costs as outlined above. In addition, the overall infrastructure like roads, railways, harbours, internet and institutions like cooperatives and farmers' associations affect the possible marketing channels and cost of trading both for inputs and outputs, which influence profit maximization. Better infrastructure is therefore positively correlated with competitiveness (lower average unit cost). In general, larger farms tend to benefit from economies of scale which reduce marginal costs of production. The use of machinery affects the efficiency or productivity by which inputs are turned into outputs. Consequently, they influence the unit cost. Assets such as buildings and machinery, but also costs for certification or audits belong to fixed costs. Some certification schemes also impose specific requirements for production processes or inputs used, which are more expensive (Bouamra-Mechemache & Chaaban, 2010). For example, Bouamra-Mechemache and Chaaban (2010) found variable production costs of PDO Brie to be 40% above those for non-PDO Brie.

Finally, farm income is affected by the farm gate price. For small farms or producers of mass products, the price is exogenous. Such farms are price takers. However, there are mechanisms by which farms can increase their market and bargaining power. Farm size, degree of product differentiation, market share, competition from close substitutes and market concentration (both within the sector of interest and up- and downstream players of the supply chain) are relevant factors to think of in relation to market structure and bargaining power, which co-determine the farm gate price. For example, organic production is also usually linked to higher output prices (price premiums), although production costs can be higher as well (Shadbolt et al., 2005). Once, the market structure allows farmers to determine prices, advertising helps convincing people of the special attributes of a certain product and increasing the willingness to pay (Krugman & Wells, 2013). However, advertising is not useful for price takers like firms in a perfect competitive market because for them farm gate price equals marginal cost. However, in a monopolistic competitive market or an oligopoly, producers can additionally benefit from advertising if they have market power to set prices above their marginal cost (Krugman & Wells, 2013). From a consumer's point of view, the affinity to the a specific region, interest in food and quality or origin of food, and a region's attractiveness for tourism (as it is linked to memorability and brand awareness) affect the elasticity of demand and willingness to pay a price premium for products from a specific origin (Van de Pol, 2017). In addition, exchange rates affect long-run profits as they determine the attractiveness of and demand for the product on foreign markets (Beckmann & Schimmelpfennig, 2015). The influence of the market structure and a farm's bargaining power on farm profits are discussed in more detail in section 2.2.

2.2 Theoretical impact of geographical indications on farm income

Usually, farming is a business meant for earning household income. Consequently, adoption of GIs is higher if expected profits from producing (ingredients for) GI products are higher than regular profits. Maximizing economic profits is equal to maximize the difference between total revenue and total cost (both explicit and implicit) (Frank & Cartwright, 2016). In general, farmers operate under perfect competition as there are thousands of farmers in the world who produce the same products. In a perfect competitive market, companies produce standardized products that are perfect substitutes (Krugman & Wells, 2013). Since most agricultural goods are traded on the world market, they can easily be replaced by substitutes from all over the world. As a result, farmers do not have any market power. They are price takers. In the long run, economic profits are zero because farms produce until marginal cost equals the exogenous price, which is the marginal revenue that firms can obtain (Krugman & Wells, 2013). Only farms with relatively low average total costs, e.g. by applying modern technology or benefitting from economies of scale, can make profits in the short run. Small farms and farms in disadvantaged areas tend to be the least efficient farms with highest average total costs (Meraner et al., 2015). If the exogenous price is below their marginal cost, they make negative profits. Finding a way out of perfect competition allows them to stay in the business. With imperfect competition, producers gain market power, so they are no longer price takers.

Three questions have to be answered: First, how to achieve a market structure with imperfect competition? Second, why do profits increase with imperfect competition? Third, how and under what conditions can GIs turn the market structure from perfect to imperfect competition?

How to achieve imperfect competition?

Oligopoly and monopolistic competition are the two important prevalent market structures of imperfect competition that can result from GI uptake. The first situation is given when only few firms produce the same product (Krugman & Wells, 2013). When many competing producers offer a range of similar but differentiated products, and entry into or exit from that market are free in the long run, one speaks about monopolistic competition (Krugman & Wells, 2013). In both cases, pricing power allows firms to earn higher profits than with perfect competition, although it can be limited by competition owing to the existence of imperfect substitutes (Krugman & Wells, 2013).

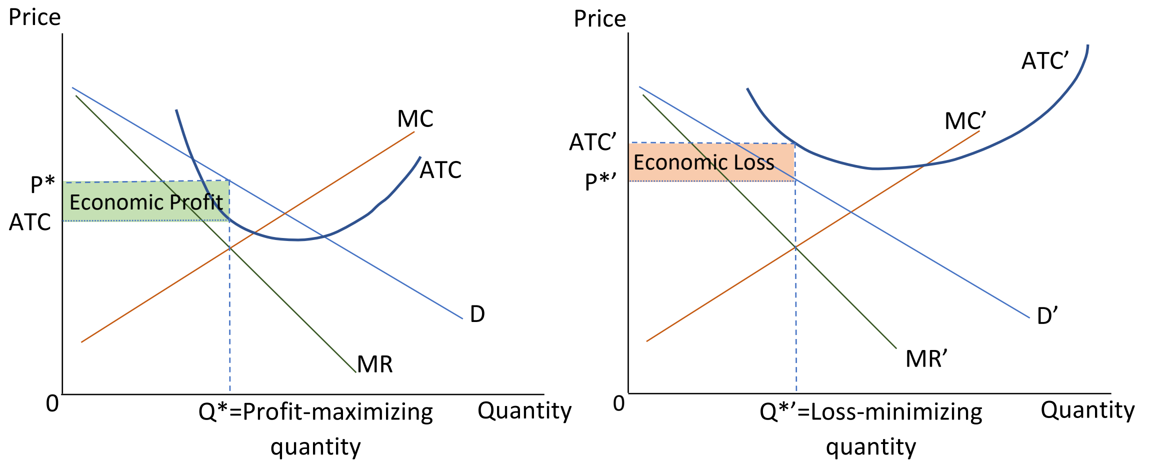
Consumers do not have the same tastes and preferences. Hence, producing several varieties of a product with diverse attributes pays off for producers (Estrin et al., 2008). It reduces competition intensity (Krugman & Wells, 2013). GIs certify a unique quality that is linked to the product's origin. GI labels make this differentiation clear to consumers. Product differentiation allows producers to make profits from selling a specific product, which other firms are not allowed, willing or able to perfectly copy (Varian, 2014). Product differentiation is the "attempt bs a firm to convince buyers that its product is different from the products of other firms in the industry" (Krugman & Wells, 2013). The demand curve is no longer perfectly elastic because people are willing to pay more for the special attributes of the differentiated product (Varian, 2014). This gives some market power to producers depending on the competition from rivals who produce imperfect (but maybe close) substitutes (Krugman & Wells, 2013). If the relative price of a differentiated product is too high (because of higher price premiums and/or higher production cost) compared to the imperfect substitutes offered on the market, consumers switch to one of these relatively cheaper products. This depends on the elasticity of demand both with respect to own prices and prices of (imperfect) substitutes.

What happens to profits when there is imperfect competition?

Whenever there is imperfect competition, demand is no longer perfectly elastic (demand curve is no longer a horizontal line). The steeper the demand curve, the less elastic is the demand. With monopolistic competition or an oligopoly, a firm maximizes its profits by producing the quantity at which marginal cost equals marginal revenue, just like in a monopoly. Figure 1 shows two firms in a monopolistic competitive market. The firm on the left side earns positive economic profits as its average total costs (ATC) at the profit-maximizing output quantity Q* are below the price P*, which consumers are willing to pay. The firm produces as much until marginal revenue (MR) equals marginal cost (MC), which is Q*. For quantity Q*, consumers are willing to pay price P* as shown by the demand function (D). The firm on the right side earns negative economic profit (losses) as its ATC curve lies above the demand curve (D'). Again, firms produce the quantity for which MR' equal MC', but consumers' willingness to pay for that quantity Q*' lies below the ATC' for that quantity. Consequently, the demand curve must cross the average total cost curve to allow a firm to make positive economic profits in the short run (Krugman & Wells, 2013). The long-run equilibrium is characterized by zero profits because more firms will enter the market as long as firms make positive profits and market entry is free (Krugman & Wells, 2013). However, in the case of PDO and mountain products, market entry is limited since production and processing are linked to a certain area, so even ingredients need to have the local origin.

How and under which conditions do GIs lead to higher profits?

For simplification, I assume that each farm produces only one product. In addition, I assume that consumers are convinced that the product is different, and they are willing to pay more for the special attributes. Further, a specific GI product (like the PDO Prosciutto di Parma) can be produced by one or several farms. In the latter case, farms produce perfect substitutes that are not further differentiated, e.g. by product packaging. If there was only one producer of that GI product, he operates under monopolistic competition. This is illustrated in scenario (a) of Figure 2. If there are many differentiated products and the differentiated GI product is produced by several



Source: Author's sketch based on Krugman & Wells (2013)

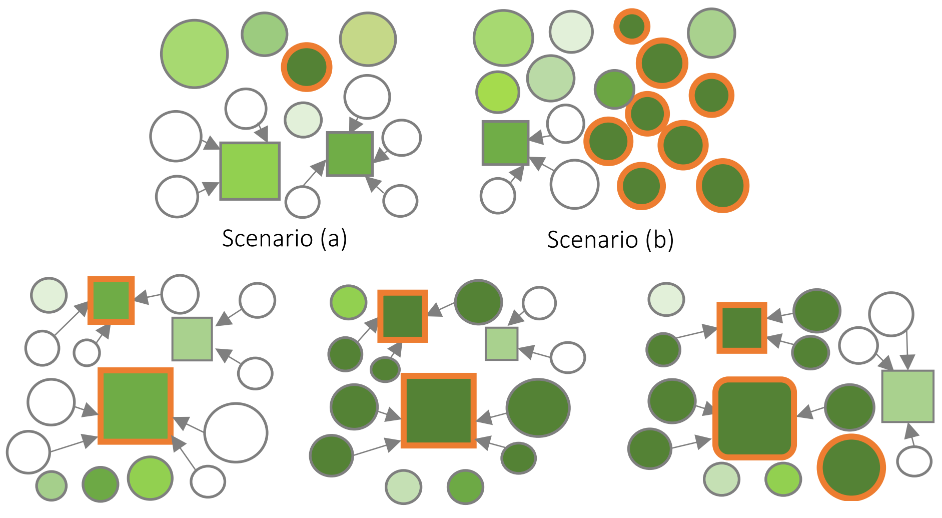
Figure 1: Profit-maximization and loss-minimization with imperfect competition

firms, such as shown in scenario (b) of Figure 2, the producers of this GI product operate in a homogenous oligopoly, with few farms producing perfect substitutes and facing competition from close (non-GI) substitutes. A GI certification that is shared by several producers can function as a collective brand strategy, like Borg and Gratzer (2013) argue for the case of PDO products. The more producers enter, the closer the market structure will be to perfect competition as more and more farms produce perfect substitutes.

The model for GI market structures becomes even more complex when considering that a farmer, who is involved in a value chain of a specific GI product, can take up two distinct positions: Either she processes her own raw products to produce the GI product herself like in scenarios (a) and (b), or she delivers the ingredients for the GI product to a processing company. Scenario (c) shows the case where several farmers are producing ingredients for a GI product. The production of this specific GI product does not require ingredients from a specific origin (e.g. PGI or TSG). Therefore, the output of farmers who are involved in the GI value chain can be easily substituted by ingredients offered on the world market. Thus, these farmers do not have any market or pricing power as they face perfect competition, although they produce ingredients for a GI product.

In contrast, farmers gain market power if geographic attributes of their raw products such as their origin are appreciated by consumers and somehow differentiate them from the output of farmers in the rest of the world. PDO and mountain products usually have strict specifications with respect to their ingredients' origin, while ingredients for PGI and TSG products can theoretically be sourced from all over the world. Thus, income effects might differ depending on which GI scheme is applied. Scenario (d) shows the case where the output of farmers, who participate in the GI value chain, differs from output of other farmers. Ingredients for the dark green coloured GI product cannot be sourced from other farmers than the dark green coloured farmers. If there is only one farmer supplying the necessary ingredient, he is a monopolist. It is more realistic to think of several farmers who fulfil the GI specifications. Consequently, GI farmers operate under a homogenous oligopoly and have some market power. Since there is a limited number of farms that can offer ingredients with the required origin, it is unlikely that these differentiated farms end up in perfect competition.

Scenario (e) adds two new components. First, there are both farmers who produce final GI products and those who produce ingredients with a specific origin for the GI product. Second, there is not only one independent processor of ingredients, but also a cooperative-driven processor (square with orange contour and rounded edges). For example, *Royal Friesland Campina* is owned by member farms of the cooperative *Zuivelcoöperatie Friesland Campina*. The milk price payed to member farmers includes issues of member bonds. Interest on member bonds additionally affects farm income (Friesland Campina, 2018). If the product differentiation leads to a high price premium paid by consumers, farmers can either benefit from higher prices paid for their milk or via their member bonds that become more profitable the more profitable the company is. Independent processors might not forward the price premium paid by consumers to farmers. If cooperative-driven processors did the same, farmers could at least benefit from member bonds. Shortcomings of the model are that it assumes that a processor only produces the GI product, and that all members of the cooperative deliver ingredients for



Note: Each circle represents a farm. All squares with sharp edges are independent processors. Squares with rounded edges represent cooperative-driven processors. Arrows signal delivery of ingredients from a farm to a processor. The size of the geometrical form reflects a firm's economic size. Green colors represent differentiated products. GI products are indicated by the orange contour. The white color is used for ingredients that are not differentiated and, consequently, can be substituted by any other ingredient from the world market. The dark green color represents the value chain of a GI product whose ingredients must originate from a specific area (PDO or mountain product).

Source: Author's sketch

Figure 2: Market structures for GI farmers

this product. In real life, however, the processing company produces several products both with and without GI labels. Gains in total profits are shared, although not all members have been involved in the product differentiation that was responsible for the increase in the processor's profits (personal communication, June 16, 2018). If the share of the GI product is relatively low, income effects for farmers are also low and maybe insignificant.

Figure 2 shows that GI farmers can face competition from other farmers and/or processors who produce the same GI ingredients or the same GI product. There can be efficiency gaps among producers of the same GI ingredient or product. Huang and Zhang (2018) analysed the effect of technological gaps in an oligopoly between "advanced" and "backward" firms of unequal size and operating costs on their profits given that all firms produce a similar product. They found that the more efficient firms are likely to determine prices (price leadership), while the backward firms have less market power and behave as price-takers. Consequently, market power still depends on farmers' relative position in the GI market with respect to efficiency and market share. However, Huang and Zhang conclude that despite of the efficiency gap, both types of firms earn higher profits because of imperfect competition. They claim that this may even be the result of collusive behaviour among advanced and backward firms, which is difficult to uncover. However, it is likely that producers of a certain GI product feel connected and collaborate such as in the case of a collective brand strategy (Borg & Gratzer, 2013).

To sum up, market power of GI farms depends on the price elasticity of demand, competition from imperfect but close substitutes, the number of farms producing the same GI product, the market share and competitiveness of the farm with respect to colleagues/competitors who produce the same GI product. Further, Figure 2 has shown that it makes a difference whether a farm produces a final GI product or ingredients for a GI product. In the latter case, ingredients can be easily substituted by agricultural products bought on the world market, if they do not need to be sourced from a specific origin, which decreases the farms' market power.

I explained why the type of processor (cooperative-driven or independent) influences a GI ingredient supplier's profits. What has not been considered so far is the market structure and market power of downstream players in general. In their models, Krugman and Wells ignore the complexity of modern supply chains. Figure 2 distinguished between farms who produce final products and those who produce ingredients. The latter is linked to a more complex supply chain as ingredients are processed by another level of the supply chain, whose market structure affects the market power of the GI farms. If processing of GI ingredients is controlled by few firms, they form an oligopoly that can increase its revenues by limiting the production of GI products (Krugman & Wells, 2013). This reduces processors' demand for GI ingredients. Consequently, producers of GI ingredients do not have any market power and in a price-taking position if the demand for their GI ingredients is (artificially) limited. Further, downstream players such as processors and retailers often have large market shares because these levels of the supply chain are highly concentrated. Therefore, even farms producing (ingredients for) differentiated products can end up without any market power if downstream players are powerful and dictate prices (personal communication, June 16, 2018).

This section has discussed the market structure that GI producers are facing. It has shown that GI uptake can help maximising profits by gaining (some) market power which allows producers of GI products to operate under a higher price. However, it has also highlighted that positive income effects are not guaranteed and depend on which GI scheme is applied. Farmers who produce final products with PDO or mountain product certification are expected to benefit from more market power and gains in profits than TSG or PGI farms, especially if they only product ingredients for the GI product.

2.3 Determinants of the decision to adopt geographical indications

In principle, GI adoption is meant to lead to product differentiation, which again is intended to increase market power of producers by decreasing the price elasticity of demand for the labelled product. Thus, farm gate prices and profit margins are expected to be higher, which positively affects farm income. Therefore, farms with little pricing power (price-takers) and low farm income are assumed to have higher expected benefits from GI adoption. Consequently, factors determining market power and farm income (especially those determining efficiency, competitiveness and farm gate prices) influence a farmer's decision to adopt GIs. The higher the expected gains of GIs with respect to efficiency, competitiveness or bargaining power, the higher the probability of uptake. However, only few studies about specific determinants of GI adoption have been published. In contrast, other certification schemes like organic farming and farm diversification in the form of agri-tourism are studied more often.

According to Bouamra-Mechemache and Chaaban (2010), a larger size of enterprises is negatively correlated with PDO certification. Farms located in less favoured areas with natural constraints such as high gradients, low soil fertility, harsh climate or weather extremes tend to be less efficient and more likely to engage in farm diversification such as GI production (Van de Pol, 2017). The better the soil fertility, the more efficient is the production, which allows farmers to earn economic profits in competitive markets. The larger the distance to urban areas and the larger the role of agriculture in terms of employment and economic activity in a region is, the less off-farm employment opportunities exist. This increases the probability of farm diversification such as the uptake of food quality schemes (Meraner et al., 2015). At the same time, proximity to main roads or hubs has a positive influence on farm diversification as it facilitates marketing and trading (Meraner et al., 2015). Full-time farmers are more likely to use GIs (Van de Pol, 2017). On the one hand, GDP is expected to be lower in regions with more registered products, as they are meant to increase farm profits of farmers in remote areas (Van de Pol, 2017). On the other hand, national or regional GDP per capita influences consumption and willingness to pay for quality attributes, which is also reflected in the slope of the demand curve. A resulting hypothesis is that a higher GDP per capita allows for a steeper demand curve for GI products, which results in more market power and higher profit margins for farmers. Beckmann and Schimmelpfennig (2015) found a positive effect of GDP on the uptake of PDO labels.

Giaccio et al. (2018) investigate determinants of agri-tourism income. The authors find that access to subsidies and advice from external institutions have a positive effect on the uptake. Experiences of "a long history of using trademarks or other quality assurance schemes" has a positive effect on GI uptake, too (Van de Pol, 2017, p. 25). Meraner et al. (2015) found that in the Netherlands, mixed farm types diversify the most. There are more fruits, vegetables and cereals, cheeses and meat products certified as PDO or PGI product than for example oils and fats or fish (Van de Pol, 2017). New employment opportunities for family members might play a role, particularly if farmers are producing final GI products as PDO production is more labour intensive than conventional production (Giaccio et al., 2018; Bouamra-Mechemache & Chaaban, 2010). In the Netherlands, larger families in general are more likely to engage in farm diversification (Meraner et al., 2015). GI uptake is less profitable if a large share of produced goods is consumed by the farm household itself (Van de Pol, 2017).

Further, younger farmers are more likely to diversify (Giaccio et al., 2018; Meraner et al., 2015). Farmers' attitudes and attachment to the "maintenance of rural lifestyles and the preservation of cultural heritage, especially as related to local food production" also determines GI uptake (Giaccio et al., 2018, p. 219). The involvement in a GI value chain requires some motivation, engagement and organisation. Therefore, additional determinants can be farmers' managerial skills, market orientation and social capital in terms of networks among farmers within a region (Van de Pol, 2017). This can be linked to farmers' educational level, but correlation was found to be insignificant in previous research (Giaccio at al., 2018). Direct sales are another way to increase added value on the farm. It also indicates market orientation and farmers' willingness to engage in product differentiation. Therefore, direct sales are assumed to be positively correlated with GI uptake.

3. Estimation techniques

The goal of this chapter is to provide a theoretical background about estimation techniques which are frequently used for impact assessment. Knowing more about strengths and weaknesses of different estimation techniques is helpful in selecting a proper method and model for the impact evaluation of GIs for food products on farm income. Section 3.1 introduces to the general problem of impact assessment: missing data. True counterfactuals never exist. Section 3.2 deals with estimation techniques for cases where treated sample units (here GI users) were chosen randomly, that is treatment is an exogenous variable. In contrast, the estimation techniques presented in section 3.3 assume treatment to be the outcome of a decision taken by each sample unit, which leads to endogenous self-selection.

3.1 The problem of impact assessment

The main problem of impact assessment is one of missing data. What would have happened to GI farms if they had not used GIs? This is the question we pose when we intend to measure the impact of GI adoption. Those farms who are producing only (ingredients for) GI products are further called GI farms, (GI) adopters or treated farms. Similarly, we could also ask what effect GIs would have had on non-adopters (also called non-GI farms, control group or untreated farms) if they had adopted GIs.

If Y_i^1 indicates the outcome of person i with adoption of GIs, while Y_i^0 represents the outcome of person i without adoption, the average treatment effect (ATE) is the average change in Y (ΔY_i), so

$$ATE= E(\Delta Y_i)=E(Y_i^1)-E(Y_i^0). \tag{3.1}$$

However, both the outcome for GI farms in case of no GI adoption and the outcome of non-GI farms with adoption are not observed. The goal of impact assessment is, therefore, to estimate a valid counterfactual that best represents the situation of treated farms if they had not been treated (no GI uptake) and vice versa. The average treatment effect on the treated (ATT) can be written as

$$\begin{aligned} ATT &= E(\Delta Y_i|T_i=1) = E\left(Y_i^1-Y_i^0\,|\,T_i=1\right) \\ &= E\left(Y_i^1\,|\,T_i=1\right)-E(Y_i^0\,|\,T_i=1). \end{aligned} \tag{3.2}$$

T_i=1 indicates that the farm was indeed treated (GI adoption), while T_i=0 refers to farms that were not treated. We do not observe $E(Y_i^0|T_i=1)$, but we do observe $E(Y_i^0|T_i=0)$. Comparing the outcomes of adopters and non-adopters then leads to

$$\begin{aligned} \beta_{naive} &= E\left(Y_i^1\,|\,T_i=1\right)- E\left(Y_i^0\,|\,T_i=0\right) \\ &= E\left(Y_i^1\,|\,T_i=1\right)-E\left(Y_i^0\,|\,T_i=1\right)+ E\left(Y_i^0\,|\,T_i=1\right)- E\left(Y_i^0\,|\,T_i=0\right) \\ &= E\left(Y_i^1-Y_i^0\,|\,T_i=1\right)+[E\left(Y_i^0\,|\,T_i=1\right)- E\left(Y_i^0\,|\,T_i=0\right)] \\ &= ATT + \text{Selection Effect.} \end{aligned} \tag{3.3}$$

The selection effect arises if treatment and control group differ even before or without the treatment. In such cases $E\left(Y_i^0\,|\,T_i=1\right)$ is not equal to $E\left(Y_i^0\,|\,T_i=0\right)$.

3.2 Exogenous programme placement

It is relatively easy to find a valid counterfactual when the treated units are randomly selected into the treated group, while untreated units are randomly selected into the control group. This can happen purely random by flipping a coin for each unit. An alternative is that treated units are selected randomly conditional on some observable characteristics. This is especially common if the treatment or programme is designed to have an impact on a specific kind or group of units in the sample, e.g. those who are relatively poor or those with less favourable land (Khandker et al., 2010).

Pure randomization (PuR)

A counterfactual is easily found when treatment and control group are on average identical except for the intervention, that means they score similarly with respect to observed and unobserved characteristics and they are exposed to the same trends. This key assumption is called *mean independence* (Khandker et al., 2010). With mean independence, equations (3.4a) and (3.4b) apply:

$$E\left(Y_i^1 \middle| T_i=1\right)=E\left(Y_i^1 \,|\, T_i=0\right)=E(Y_i^1) \tag{3.4a}$$

$$E\left(Y_i^0 \middle| T_i=1\right)=E\left(Y_i^0 \,|\, T_i=0\right)=E(Y_i^0). \tag{3.4b}$$

Again, the term Y_i^1 in equation (3.4a) is the outcome for a farm with treatment and Y_i^0 is the outcome for a farm without treatment. Equations (3.4a) and (3.4b) usually only apply when the impact evaluation is based on an experimental design. Pure randomization allows to assume that the outcome of GI adopters would have been the same as that of non-adopters if both had not received any treatment (no GI adoption) and vice versa. Then $E\left(Y_i^1 \middle| T_i=1\right)$ would be equal to $E\left(Y_i^1 \,|\, T_i=0\right)$. Therefore, non-adopters are a valid counterfactual and the selection effect equals zero, so $\hat{\beta}_{naive}$ equals the ATT.

In such a setting, to estimate $\hat{\beta}_{naive}$ a simple Ordinary Least Squares (OLS) regression can be run. Equation (3.5) shows such a Random Control Trial (RCT) model for impact evaluation for the case of GI adoption, where Y_i is the outcome of farm income, α is the mean outcome of farm income for non-adopters (T_i=0), $\alpha+\beta$ is the mean outcome for GI adopters (T_i=1), and ε_i is the error term (Khandker et al., 2010). Consequently, β indicates the treatment effect, which is the difference in mean outcome between adopters and non-adopters.

$$Y_i=\alpha+\beta T_i+\varepsilon_i \tag{3.5}$$

The evaluation of income effects of GIs is, however, based on an observational study instead of an experimental design. Consequently, pure randomization is not a reasonable assumption in this case. Therefore, model (3.5) is also called the "naïve" impact estimation model. Although the estimated treatment effect is likely to be biased, this model is useful for testing whether the income of treated and untreated farms differ. However, the potential difference should not be interpreted as causal effect of GI adoption.

Partial randomization with common impact (PaRCI)

Next to pure randomization, there are also cases of conditional exogeneity of programme placement. If this would be true for the case of GIs, farms were randomly assigned to the control or treated group conditional on some observable characteristics X_i (Khandker et al., 2010). Indeed, GIs were introduced to offer an alternative to cost-minimizing strategies, which was intended to benefit small farms in disadvantaged areas, who have difficulties to compete with larger and more efficient producers (Hajdukiewicz, 2014). If only those farms adopted GIs that were intended to use them, it is likely that their average outcome before GI adoption was not the same as the average outcome of non-adopters. Rather, $E\left(Y_i^0 \middle| T_i=1\right) < \left(Y_i^0 \,|\, T_i=0\right)$. The average income of the treated group is expected to be lower than the average income of the control group if none of them adopts GIs. Controlling for some observables leads to model (3.6), where Y_i is the outcome of farm income, α^C is the mean outcome of farm income for non-adopters (T_i=0) conditional on the covariates X_i, α^T-α^C is the deviation of farm income for GI adopters (T_i=1) from those of non-adopters conditional on X_i, and ε_i is the error term. This model assumes that both adopters and non-adopters react similarly with respect to changes in the variables included in the vector X_i. The treatment effect can be estimated by the coefficient of the treatment dummy variable α^T-α^C (Khandker et al., 2010).

$$Y_i=\alpha^C+(\alpha^T-\alpha^C)T_i+\beta X_i+\varepsilon_i \tag{3.6}$$

Control variables must cover all relevant differences between adopters and non-adopters, so they have the same potential outcomes conditional on these variables (Duflo et al., 2006). "Controlling for baseline values of covariates likely to influence or predict the outcome does not affect the expected value of an estimator of [the treatment effect], but it can reduce its variance" (Duflo et al., 2006, p. 34). In contrast, control variables with little or any effect on the variation in the outcome reduce degrees of freedom and increase standard errors (Duflo et al., 2006). Including covariates that are influenced by the treatment leads to biased estimates because part of the treatment effect is then embraced by the coefficients of these variables. Baseline values ensure that covariates are unaffected by treatment (Duflo et al., 2006). Finally, X_i can encompass dummies for all categories that one intends to control for, as well as interactions (Duflo et al., 2006). This estimation technique assumes that all relevant covariates are observed to solve for selection bias.

Partial randomization with varying impact (PaRVI)

Based on the PaRCI model, we can add another assumption of varying impact of covariates between adopters and non-adopters. Therefore, another estimation technique is to start with two income equations for the treated and untreated groups separately. Equation (3.7a) is the model estimated for the sample treated farms (T_i=1), while equation (3.7b) reflects the model used to estimate farm income for the control group (T_i=0). Interpretation is similar to equation (3.5), with μ_i^T and μ_i^C being the error terms for the adopters and non-adopters respectively.

$$Y_i^T=\alpha^T+\beta^T X_i+\mu_i^T \quad \text{if } T_i=1, i=1, \ldots, n \tag{3.7a}$$

$$Y_i^C=\alpha^C+\beta^C X_i+\mu_i^C \quad \text{if } T_i=0, i=1, \ldots, n \tag{3.7b}$$

Multiplying equation (3.7a) by T_i and equation (3.7b) by $(1-T_i)$ allows to merge both models leading to equation (3.8):

$$Y_i=\alpha^C+(\alpha^T-\alpha^C)T_i+\beta^C X_i+\left(\beta^T-\beta^C\right)X_i T_i+\varepsilon_i, \tag{3.8}$$

where β^C refers to the coefficients of non-adopters with respect to the covariates X_i, while β^T-β^C represents the difference between coefficients of adopters and non-adopters (Khandker et al., 2010). The average treatment effect on the treated (ATT) equals $E(Y_i|T_i=1, X_i)=E[\alpha^T-\alpha^C+X_i\left(\beta^T-\beta^C\right)]$. The estimate is consistent if there is no selection bias apart from the bias corrected for by the covariates, that is if $E(\mu_i^T|X_i, T=t)=E\left(\mu_i^C|X_i,T=t\right)=0$, t={0, 1} (Khandker et al., 2010; Duflo et al., 2006). While the PaRVI model assumes that coefficients for control variables vary between the treated and untreated groups of farms, the PaRCI model simply assumes that $\beta^T=\beta^C$. However, the more variables one intends to control for and the more continuous variables are intended to be included in the vector of control variables of the previous models, the more difficult it becomes to form similar treated and untreated groups that fulfil the assumption of conditional independence (Duflo et al., 2006). Therefore, the PaRCI and PaRVI models have their limitations when applied to the case of GIs.

3.5 Endogenous self-selection

Estimating a valid counterfactual outcome becomes more difficult when it is not obvious why some units are treated, and others are not. In the case of GI adoption, it is not fully clear which factors influence the decision to adopt or not adopt GIs. It is not exogenously chosen which farms adopt GIs and which farms do not. Therefore, selection bias is not the result of programme placement. Rather, it is caused by self-selection. It is likely to be based on individually expected benefits or the expected utility of GI adoption, which is positive for adopters and zero or negative for non-adopters. Consequently, application of the previously discussed techniques potentially leads to biased estimates. What is needed is an estimation technique that corrects for self-selection based on observed characteristics and/or unobserved characteristics.

Observed self-selection

One possible approach is the use of all observed determinants of treatment to estimate the probability for an individual to be in the treated group, which is called the propensity score (PS). The impact estimation technique linked to that is called Propensity Score Matching (PSM). "[O]ne tries to develop a counterfactual or control group that is as similar to the treatment group as possible in terms of observed characteristics" (Khandker et al., 2010, p. 54). It has two core assumptions: Conditional independence and common support (Khandker et al., 2010). Similar to the previous estimation techniques, conditional independence or unconfoundedness assumes that conditional on the control variables, farms randomly select into the control or treatment group (Verbeek, 2012). Consequently, the potential outcome of farm income is independent of treatment conditional on the control variables. For PSM to produce unbiased estimates of treatment effects, the assumption of strongly ignorable treatment assignment must hold, which means that both GI application and potential outcomes of farm income are independent conditional on the observed variables used as control variables (Shadish & Steiner, 2010). The control variables need to influence both treatment (GI adoption) and pre-treatment outcomes (farm income before/without GI adoption) (Shadish & Steiner, 2010; Caliendo & Kopeinig, 2008). Variables that only affect the outcome can be included too, but it is crucial that none of the covariates is affected by treatment (Caliendo & Kopeinig, 2008). Over-parametrization should be avoided, so adding too many covariates does not improve the estimation of treatment effects (Caliendo & Kopeinig, 2008). In addition, PSM estimates are only unbiased if uptake of GIs is solely affected by observed characteristics that can be controlled for. As before, it is a very strong assumption that no unobservables play a role in GI adoption.

Propensity scores are estimated by a binary choice (probit) model with adoption (T) as the dependent variables and several covariates as explanatory variables, which are (theoretically) correlated with adoption and the outcome. The predicted values of adoption from the selection equation are equal to the propensity score $P(x_i)$, which reflects the probability of GI uptake conditional on the covariates. These covariates can be summarized in a vector x_i. The PS can be written as

$$p(x_i)= Pr\{T_i=1|x_i\}. \tag{3.9}$$

McFadden R^2 and Count R^2 are indicators for how well the model predicts GI adoption. The area of common support is the range of estimated PS of adopters for which non-adopters with same PS exist, so that matches between adopters and non-adopters can be formed within this range. Balancing tests are used to check whether treated and untreated groups score similarly with respect to the covariates after matching. There are different matching techniques to check for the robustness of estimates. For example, given a treated farm with a specific PS, nearest-neighbour (NN) matching takes the untreated farm with the nearest PS as a match and compares the expected outcomes for both (Caliendo & Kopeinig, 2008). This can also be done using a specified number of nearest neighbours, either with or without replacement where non-adopters are (not) allowed to be used several times for matching. While replacement reduces the bias, but it increases the variance, using more than one nearest neighbour increases the bias and reduces the variance (Caliendo & Kopeinig, 2008). Next, radius matching only uses non-adopters for matching that lie within a specified range of PS. This might lead to more non-adopters not being used for matching, which increases the chance of sampling bias (Khandker et al., 2010). The variance of estimates also goes up with fewer matches that can be made (Caliendo & Kopeinig, 2008). It is difficult to define a reasonable tolerance (Caliendo & Kopeinig, 2008). Again, this method can be used with or without replacement. The advantage of radius matching over caliper matching is that all comparison group members within the defined tolerance level can be used as match, which “allows for usage of extra (fewer) units when good matches are (not) available” (Caliendo & Kopeinig, 2008, p. 42). Compared with NN matching, variance tends to be lower (Caliendo & Kopeinig, 2008). The last example is kernel matching, which uses all non-adopters as matches for each of the treated farms. Each non-participant is weighted depending on the distance of its PS compared to the GI adopter with whom he/she is matched. Since more information is used by this matching algorithm, the variance of estimates tends to be lower than with NN (Caliendo & Kopeinig, 2008). However, bias increases when rather bad matches are made.

With PSM, the average treatment effect (ATE) is the “mean difference in outcomes across these two groups” (Khandker et al., 2010, p. 53). Sampling bias can be a probem if a nonrandom subset of the GI farms has to be excluded from the analysis because no matches can be found. According to Khandker et al. (2010), the ATE for cross-section data within the common support can be written as

$$ATE_{PSM}=\frac{1}{N_T}\left[\sum_{i\in T} Y_i^T - \sum_{j\in C} \omega(i,j)Y_i^C\right] \tag{3.10}$$

where N_T represents the number of GI farms i, Y_i^T is the outcome of farm income for GI farms, Y_i^C is the outcome of farm income for non-GI farms, and $\omega(i,j)$ is the weight assigned to the matched non-GI farms j (Khandker et al., 2010).

Two main problems can arise when using PSM. First, PSM only corrects for self-selection based on observable characteristics. In the case of GI adoption, it is likely that there are selectivity effects due to unobserved differences between adopters and non-adopters. Second, PSM requires valid matches to exist.

Unobserved self-selection and heterogenous effects

An alternative technique is the Endogenous Switching Regression (ESR) model. It solves for selectivity effects caused by observed and unobserved differences between the control and treated group (Di Falco et al., 2011). It also allows for heterogeneity, that is different impacts of GI adoption on treated and control group. Like PSM, the ESR model assumes the adoption decision to be endogenous (Di Falc et al., 2011). For this model, a two-stage framework is applied. First, a selection equation estimated by a probit model is needed with a dependent variable that reflects whether the expected benefits of GI adoption for a farmer are positive or not, and with determinants of adoption as explanatory variables (Khonje et al., 2015):

$$T_i^*=Z_i\alpha+\varepsilon_i \text{ with } T_i=\begin{cases} 1 \text{ if } T_i^*>0 \\ 0 \text{ otherwise} \end{cases} \tag{3.11}$$

T_i is 1 for all farms who expect that they gain from GI adoption, while it is zero for all other farms. Z_i is a vector which covers variables that influence the adoption decision (Di Falco et al., 2011). It should include a selection instrument that affects the adoption decision but not farm income (Khonje et al., 2015). A simple falsification test helps identifying valid instruments. They must be jointly statistically significant in the selection equation, but jointly statistically insignificant in the outcome equation.

The second stage of this model consists of two outcome regression equations which are estimated by an OLS regression with selectivity correction (Khonje et al., 2015). Model 3.12a is used for all adopters, while model 3.12b is used for non-adopters, where X_{1i} and X_{0i} are vectors of exogenous covariates, β_1 and β_0 are the respective vectors of parameters, and w_{1i} and w_{0i} are random disturbance terms.

$$Y_i^1=\beta_1 X_{1i}+w_{1i} \text{ if } T_i=1 \tag{3.12a}$$

$$Y_i^0=\beta_0 X_{0i}+w_{0i} \text{ if } T_i=0 \tag{3.12b}$$

The error term of equation 3.11 is correlated with the error terms of equations (3.12a) and (3.12b), so the expected values of w_{1i} and w_{0i} conditional on ε_{i_i} are nonzero. ESR models have been applied in several other impact assessments in previous research where the mathematical background of the model specification has been discussed in detail (Di Falco et al., 2011; Khonje et al., 2015; Lokshin & Sajaia, 2004). Full information maximum likelihood estiamtion is said to be an efficient method to estimate the ESR model, which can be done by the Stata command *movestay*. This estimation technique provides four relevant estimates: the expected outcome of GI farms with GI adoption (eq. 3.13a), the expected outcome of non-GI farms without GI adoption (eq. 3.13b), and both counterfactuals, so the expected outcome of GI farms without GI adoption (eq. 3.13c) and the expected outcome of non-GI farms with GI adoption (eq. 3.13d). These four cases are summarized in Table 1, with

$$E\left(Y_i^1 \middle| T_i=1\right)= \beta_1 X_{1i}+\sigma_{1\varepsilon}\lambda_{1i} \tag{3.13a}$$

$$E\left(Y_i^0 \middle| T_i=0\right)= \beta_0 X_{0i}+\sigma_{0\varepsilon}\lambda_{0i} \tag{3.13b}$$

$$E\left(Y_i^0 \middle| T_i=1\right)= \beta_0 X_{0i}+\sigma_{0\varepsilon}\lambda_{1i} \tag{3.13c}$$

$$E\left(Y_i^1 \middle| T_i=0\right)= \beta_1 X_{1i}+\sigma_{1\varepsilon}\lambda_{0i}, \tag{3.13d}$$

where $\sigma_{1\varepsilon}$ and $\sigma_{0\varepsilon}$ denote the covariances of w_{1i} and ε_i and w_{0i} and ε_i, respectively. If their estimates are statistically significant, the decision of GI adoption and farm income are correlated, so the hypothesis of no sample selectivity bias is rejected. Further, $\lambda_{1i}=\frac{\varphi(Z_i\alpha)}{\Phi(Z_i\alpha)}$ and $\lambda_{0i}=\frac{\varphi(Z_i\alpha)}{1-\Phi(Z_i\alpha)}$, with $\phi(.)$ being the standard normal probability density function and $\Phi(.)$ being the standard normal cumulative density function. The average treatment effect of GI adoption for GI farms is then

$$\begin{aligned} ATT &= E\left(Y_i^1 \middle| T_i=1\right)- E\left(Y_i^0 \middle| T_i=1\right) \\ &= (\beta_1-\beta_0)X_{1i}+(\sigma_{1\varepsilon}-\sigma_{0\varepsilon})\lambda_{1i}. \end{aligned} \tag{3.14}$$

Table 1: Conditional expectations and treatment effects ESR model

	Decision stage		
	Adoption	No adoption	Treatment effects
GI farms	(a) $E\left(Y_i^1 \middle| T_i=1\right)$	(c) $E\left(Y_i^0 \middle| T_i=1\right)$	ATT
Non-GI farms	(d) $E\left(Y_i^1 \middle| T_i=0\right)$	(b) $E\left(Y_i^0 \middle| T_i=0\right)$	ATU
Heterogeneity effects	BH_1	BH_0	TH

Source: Di Falco, Veronesi, & Yesuf (2011), p. 837
Note: (a) and (b) are observed, while (c) and (d) are counterfactuals
T_i=1 if the farm produces (ingredients for) GI products, T_i=0 if the farm does not use any GI label
Y_i^1: farm income if farm adopted GIs
Y_i^0: farm income if farm did not adopt GIs
ATT: treatment effect on the treated (GI farms)
ATU: treatment effect on the untreated (non-GI farms)
BH_t: effect of base heterogeneity for farms that adopted GIs (T=1), and those who did not adopt GIs (T=0)
TH = (ATT-ATU), i.e. transitional heterogeneity

For non-adopters, the expected average treatment effect is

$$\begin{aligned} ATU&=E\left(Y_i^1 \middle| T_i=0\right)- E\left(Y_i^0 \middle| T_i=0\right) \\ &=(\beta_1-\beta_0)X_{0i}+(\sigma_{1\varepsilon}-\sigma_{0\varepsilon})\lambda_{0i}. \end{aligned} \tag{3.15}$$

In both equations 3.14 and 3.15, the second term ((.)λ) "is the selection term that captures all potential effects of the differences in unobserved variables" (Khonje et al., 2015).

In addition, effects of base heterogeneity (BH) can be measured. For example, the food quality schemes policy is intended to help farms with difficulties to compete on the global market, so it is assumed that GI farms have had a lower farm income than non-adopters before they started using GIs (BH<0). For adoption, BH is

$$\begin{aligned} BH_1&=E\left(Y_i^1 \middle| T_i=1\right)- E\left(Y_i^1 \middle| T_i=0\right) \\ &=\beta_1(X_{1i}-X_{0i})+\sigma_{1\varepsilon}(\lambda_{1i}-\lambda_{0i}). \end{aligned} \tag{3.16}$$

For non-adoption, BH is equal to

$$\begin{aligned} BH_0&= E\left(Y_i^0 \middle| T_i=1\right)- E\left(Y_i^0 \middle| T_i=0\right) \\ &=\beta_0(X_{1i}-X_{0i})+\sigma_{0\varepsilon}(\lambda_{1i}-\lambda_{0i}). \end{aligned} \tag{3.17}$$

Finally, the transitional heterogeneity (TH) indicates whether the effect of GI adoption is the same for both adopters and non-adopters (if they would adopt GIs).

$$TH = ATT-ATU \tag{3.18}$$

If it is positive, non-adopters would not gain as much from adoption as the actual adopters do.

4. Empirical model

This chapter specifies the model that is used to estimate the impact of GIs on farm income. Section 4.1 introduces to the data that is used for the impact assessment. Section 4.2 deals with the choice of a proper estimation technique. Further, it includes the selection of variables. Therefore, the theoretical background on determinants of farm income and GI adoption as discussed in chapter 2 is linked with the information provided by the data sources. Section 4.3 presents the descriptive statistics which show how GI adopters and non-adopters differ with respect to the variables of interest.

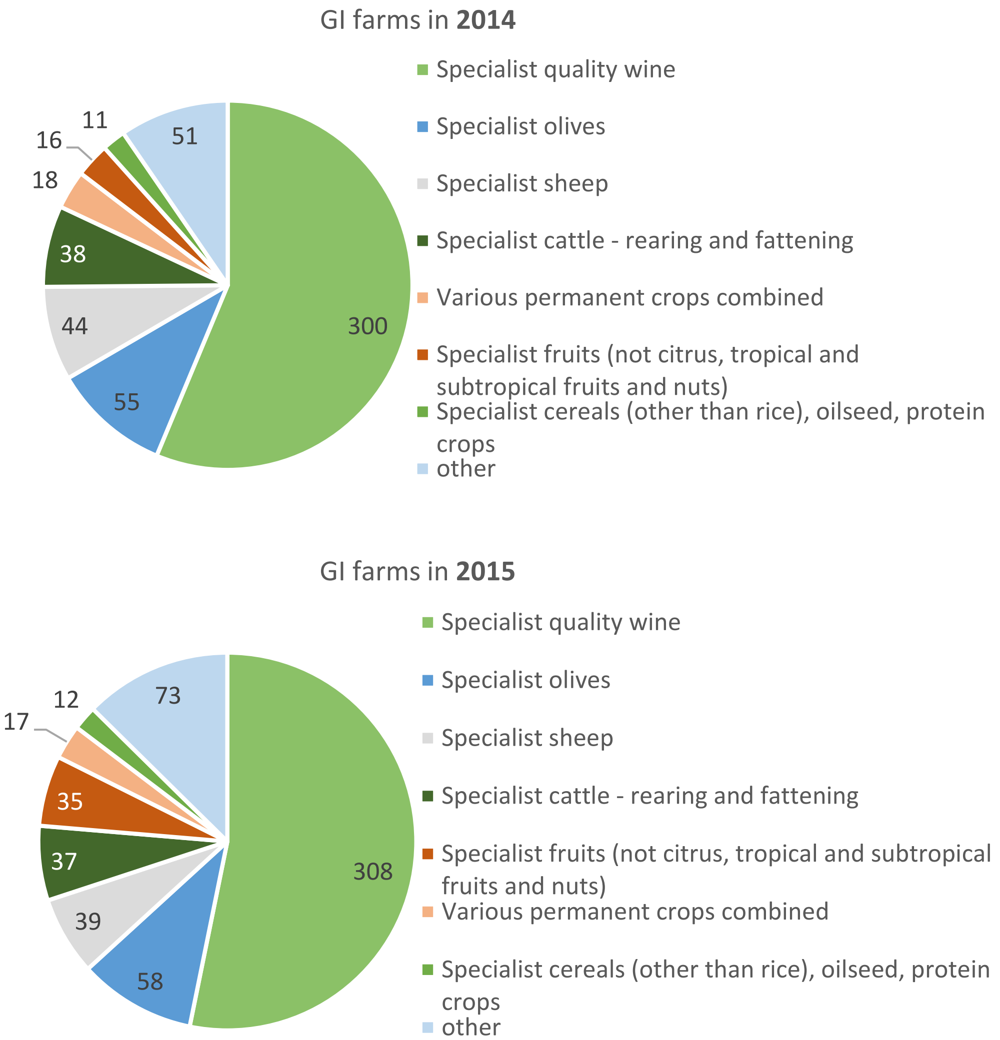
4.1 Data source

According to the DOOR database of the European Commission (2018), there are 635 registered PDO products (including those in non-EU countries), 737 registered PGI products and 58 TSG products. Within the EU, Italy has the most registered PDO, PGI or TSG products with a total of 296 registrations, followed by France with 247 registrations and Spain with 195 registrations. An overview of all registrations for all 28 EU member states of the EU is provided in appendix I.

For this research, the EU provided data from the Farm Accountancy Data Network (FADN) containing information about individual farms across the 28 member states of the European Union in the years 2014 and 2015. In addition, a dataset that was prepared by van de Pol (2017) and which is mainly based on EUROSTAT data adds some characteristics at NUTS2 and national level (EU27, excluding Croatia). The FADN sample is taken as a basis for this research. FADN uses three criteria for stratification to ensure that specific categories of farms are sufficiently represented in the sample: region, economic size and type of farming (European Commission, 2016). This ensures that the FADN sample covers the heterogeneity of farms in the European Union (EU28). Since the dataset of van de Pol does not contain information about Croatia, I excluded Croatia from the sample for this impact evaluation. The relevant question about GI application was not asked in France, Germany, Lithuania, Luxembourg, Latvia, and Slovakia. These countries are excluded from the impact evaluation too. In Ireland the question was only asked in 2015. Therefore, it cannot be considered in the impact evaluation for 2014. There are also some NUTS2 regions within the remaining EU member states where none of the farms has responded to the GI question. Consequently, impact evaluation can also not be representative for these regions: BE21, BE22, BE23, BE24, BE25, AT32, AT33, AT34, and all NUTS2 regions in the UK except for Northern Ireland.

Overall, this leads to the exclusion of about 30% of all farms from the sample, both for 2014 and 2015. The main reason for that is missing data with respect to the GI question. Croatia would have added 13 GI farms in both 2014 and 2015. 1% of the farms that are left in the dataset (called sample set 2 in appendix II) are GI farms, 90% are non-GI farms, and 4% produce some GI ingredients or products both in 2014 and 2015. The remaining 5% of the farms left in the dataset have missing data (code number 0), but different from the above-mentioned countries and NUTS2 regions, the data for these farms seems to be missing randomly. As I only want to compare GI farms with non-GI farms, the remaining farms with missing data on GI uptake and those with some GI production (code number 3) are excluded. This leads to 52133 farms left in the 2014 sample and 52606 farms left in the 2015 sample, with 1% GI farms and 99% non-GI farms respectively. I refer to this sample as sample set 3. Appendix II shows how the FADN dataset is modified step by step to end up with the samples used for the impact assessment.

The last step is categorising farms into different farm types and comparing only farms of similar farm type. It prevents comparing farms that are exposed to totally different value chains, production processes (and related use of inputs and costs) and average output prices. Finding a reasonable counterfactual for GI farms requires comparing them with farms of a similar farm type. Figure 3 shows all GI farms within the EU categorized by farm type in 2014 and 2015. It is based on the sample excluding Croatia and the countries and NUTS2 regions listed above, with Ireland being part of the reduced dataset in 2015, but not in 2014. The statistical power of impact



Source: Author's sketch based on FADN sample set 3[3]

Figure 3: Prominent farm types among GI farms

assessment decreases with smaller sample size (Nuzzo, 2016). A smaller sample makes it harder to measure significant impacts and not making a type II error (not rejecting the hypothesis of no impact, although there is one), especially if the effect is very small. Consequently, conducting an impact evaluation only makes sense for farm types with a reasonable number of GI farms. Figure 3 makes clear that quality wine specialists are most useful for the impact evaluation. In 2014, there were 300 GI and 937 non-GI quality wine specialists. In 2015, there were 308 GI and 912 non-GI quality wine specialists. The second largest GI farm group is found among olives specialists. In 2014, there were 55 GI and 981 non-GI olives specialists. One year later, 58 olives specialists were GI farms, while 934 did not use any GI labels.

4.2 Model specification

Chapter 3 presented different popular estimation techniques for impact assessment. In the case of GI adoption, we can neither speak of pure randomization nor about partial randomization. GI adoption is endogenous as farms

[3] See appendix III. SET 3 = all GI farms and non-GI farms from the FADN dataset, excluding HRV, FRA, DEU, LTU, LUX, LVA, SVK, IRE in 2014 (IRE is included in 2015), NUTS2 regions BE21, BE22, BE23, BE24, BE25, AT32, AT33, AT34, and all NUTS2 regions of the UK except for UKN0.

select themselves into the group of treated farms. Therefore, PSM or ESR techniques are expected to be appropriate estimation techniques. As the two datasets do not provide any information about personal characteristics of the farmer or the workers on the farms, it is likely that self-selection occurs not only based on observables, but also based on unobserved characteristics such as age, education and attitudes (e.g. towards own efforts to create a marketing strategy via product differentiation). Consequently, PSM is expected to give biased estimates, while an ESR model can account for selectivity based on both observed and unobserved characteristics. Hence, the impact of GIs for food products on farm income is estimated by an ESR model using the *movestay* command with Stata.

For the specification of the ESR model there are four categories of variables needed. First, the outcome variable has to be specified, which indicates the farm income. Next, a variable is needed that indicates whether a farm is only producing (ingredients for) GI products. Third, exogenous control variables that affect the outcome variable and the adoption decision must be chosen. Finally, instrumental variables must be found that affect the decision to adopt GIs, but not farm income itself. The following paragraphs will specify the variables chosen from the two datasets. A list of all selected variables including their abbreviations and descriptions is provided in appendix III.

Income variable

I chose farm net income (FNI) to be the indictor of farm income for which treatment effects are measured. It reflects the renumeration to fixed factors of production and the farmer's risks in the accounting year (loss/profit) in Euro. In the FADN dataset, FNI is calculated as indicated by equation (4.1), where total intermediate consumption equals the sum of specific costs and farming overheads, while external factors cover wages, rent and interest paid.

FNI = Total output[4] – Total intermediate consumption + Balance current subsidies and taxes – Depreciation + Balance subsidies and taxes on investments – Total external factors (4.1)

Treatment variable

The FADN dataset included a variable of GI uptake. However, it was different for both years. In 2014, farmers were asked whether they are involved in the production of (ingredients for) PDO and/or PGI products. This question was answered by "no PDO/PGI production" (code number 1), "only PDI/PGI production" (code number 2) or "some PDO/PGI production" (code number 3). In 2015, TSG and mountain products were added to the list. Therefore, the results of the impact assessment of both years cannot be compared as it measures the impact of different GIs. I conducted the impact evaluation for both years separately. A treatment variable T was generated with T=1 for GI adopters (code number 2) and T=0 for the remaining farms. I only compare farms that *only* use GIs (code number 2) with farms that do not use any GIs at all (code number 1). The category "some GI products" is too imprecise to guarantee that the impact estimates are unbiased. The given dataset only allowed estimating joint effects. It was not possible to estimate treatment effects for each GI scheme separately as the GI variable in both 2014 and 2015 encompasses several GIs. This poses a possible risk of downward biased estimates in case that positive income effects for one GI label are cancelled out by negative (or less positive) impacts of other GI labels.

Control variables

The following variables were used as covariates: total output (OUT), specific costs (SPC), overheads (OVER), total utilised agricultural area (UAA), paid labour in annual working units (PL) and unpaid labour in annual working units (UL), liabilities (LIA), external factors (EXT), machinery (MACH), the Gross Domestic Product per capita in the NUTS2 region (GDPC), farm household consumption (FHC), as well as dummies for organic production (ORG=1 if the holding applies only organic farming), less favoured area (LFA=1 if the majority of the holding is situated in LFA) and mountain area (MA=1 if the majority of the holding is situated in MA).

As discussed in chapter 2, profits (farm net income) can be said to mainly depend on revenues and costs. Farm revenues are equal to production volumes times the farm gate price. In the FADN dataset, revenues are called total output, reflected by the variable OUT. Production costs are affected by specific costs for the production of

[4] According to FADN, total output is equal to the production volume times farm gate prices, which is the same as revenues. In this thesis, I will stick to the FADN wording.

crops or animal products (SPC), overhead costs (OVER), as well as costs for labour, land and capital (summarized by the variable EXT). The latter is also affected by the amount of labour, land[5] and capital that is used. Therefore, UL, PL, UAA and LIA are added. The Gross Domestic Product per capita (GDPC) is an indicator for the elasticity of demand. Consumers with higher income are able to spend more money on luxury products and added value such as a specific origin of a product. In contrast, farmers who sell their products in an environment with relatively poor households are expected to have less market power or opportunities to increase their profit margins. Output prices for organic products are usually higher (Nieberg & Offerman, 2003). Hence, ORG is added as control variable to the outcome equation. Finally, less favoured or mountain area and the amount of machinery used on the farm are indicators for the competitiveness and production efficiency of a farm. Consequently, LFA, MA and MACH are added to the list of control variables. FHC is added because households with relatively high consumption of own products can sell less products on the market. Expected benefits of GI adoption also decrease with lower production volumes sold on the market.

Selection instruments

Selection instruments affect the decision to adopt GIs, but they do not have any effect on the outcome variable FNI of non-adopters (Di Falco et al., 2011). Since GIs are intended to increase profit margins and farm income, it is likely that determinants of farm income influence the decision to adopt GIs. For example, if UAA had a positive impact on FNI, farms with less UAA would have less FNI. Consequently, they need to seek higher profit margins, e.g. by producing (ingredients for) GI products, which increases their probability to adopt GIs. However, UAA would not be a valid instrument in this case as it affects both FNI and the adoption decision.

Information about the farms' FNI in the past would not affect current FNI, but it would affect the decision to adopt GIs. However, such a lagged variable is not part of the dataset. According to the European Commission (2011), the economic size of the farm is the best indicator for "small" or disadvantaged farms that are intended to benefit from GIs. However, economic size turned out to be no valid instrument, also when using it in combination with other variables. Several combinations of potential instruments have been tested to find variables that are not jointly significant in the outcome equation of non-adopters but jointly significant in the selection equation.

Transportation costs are often neglected in economic theories (like trade theory), so differences in the number of kilometres of motorway per 1000 km² (MKM) and its square (MKM2) are expected to have no or a minor effect on FNI. However, according to Van de Pol (2017), a certain minimum of MKM is required for farms to be interested in adopting GIs, whereas too much MKM (high MKM2) reduces the probability of uptake. The reason for this is that farms with excellent access to markets and relatively cheap transport costs have a competitive advantage and can better compete on the bulk market (without product differentiation via GIs or other means). Indeed, MKM and MKM2 fulfil the requirements for valid instruments for the sub-samples of quality wine specialists (both 2014 and 2015) and olives specialists in 2014 (see appendix IV for the test of joint significance). The expected correlations between all the above-mentioned variables of interest and both FNI and GI adoption are summarised in Table 2. Based on equations 3.11, 3.12a and 3.12b, the following models are specified for the impact assessment:

Outcome equation for GI farms:

$$FNI^1=\beta_0^T+\beta_1^T OUT+\beta_2^T SPC+\beta_3^T OVER+\beta_4^T UL+\beta_5^T PL+\beta_6^T UAA+\beta_7^T LIA+\beta_8^T EXT+\beta_9^T LFA+\beta_{10}^T MA+ \beta_{11}^T ORG+\beta_{12}^T MACH+\beta_{13}^T GDPC+\beta_{14}^T FHC+w_{1i} \quad (4.2a)$$

Outcome equation for non-GI farms:

$$FNI^0=\beta_0^T+\beta_1^T OUT+\beta_2^T SPC+\beta_3^T OVER+\beta_4^T UL+\beta_5^T PL+\beta_6^T UAA+\beta_7^T LIA+\beta_8^T EXT+\beta_9^T LFA+\beta_{10}^T MA+ \beta_{11}^T ORG+\beta_{12}^T MACH+\beta_{13}^T GDPC+\beta_{14}^T FHC+w_{0i} \quad (4.2b)$$

Selection equation:

$$T=\alpha_0+\alpha_1 OUT+\alpha_2 SPC+\alpha_3 OVER+\alpha_4 UL+\alpha_5 PL+\alpha_6 UAA+\alpha_7 LIA+\alpha_8 EXT+\alpha_9 LFA+\alpha_{10} MA+ \alpha_{11} ORG+\alpha_{12} MACH+\alpha_{13} GDPC+\beta_{14}^T FHC + \alpha_{15} MKM+\alpha_{16} MKM2+\varepsilon_i \quad (4.3)$$

[5] Since the impact assessment is only done for quality wine specialists and olives specialists, livestock does not play any role.

Table 2: Expected correlation of control variables and FNI and GI adoption respectively

	Relation to GI adoption	Relation to FNI
Total output (OUT)	Higher output, higher productivity (volumes) and/or more bargaining power (prices), so less need to adopt GIs	Higher output, higher FNI
Total specific cost (SPC)	Higher specific costs, more GI adoption (farm needs to generate higher output prices)	Higher specific costs, lower FNI
Overheads (OVER)	Higher overheads, more GI adoption (farm needs to generate higher output prices)	Higher overheads, lower FNI
Total utilised agricultural area per holding (UAA)	less UAA, more GI adoption (Hajdukiewicz, 2014)	More UAA, higher FNI
Paid labour input (PL) in annual working units (AWU)	More PL, higher GI adoption if labour costs increase relatively more than productivity; otherwise more PL leads to less GI adoption	More PL, lower FNI
Unpaid labour input (UL) in AWU	More UL, more GI adoption (Meraner et al., 2015)	More UL, higher FNI
Liabilities incl. long-, medium- and short-term loans (LIA)	Higher LIA, more GI adoption (farm needs to generate higher output prices)	Higher LIA, lower income
External factors (EXT)	Higher cost for labour, capital and land, more GI adoption (farm needs to generate higher output prices)	Higher EXT, lower FNI
Dummy for mainly less favoured area (LFA)	LFA increases attractiveness of GIs as a mean to increase FNI by product differentiation (Van de Pol, 2017)	LFA, lower FNI for non-GI farms, but higher FNI for GI farms
Dummy for mainly mountain area (MA)	Mountain area increases the attractiveness of GIs to differentiate products (Van de Pol, 2017)	Mountain area, lower FNI (higher production cost)
Dummy for applying only organic production (ORG)	Less GI adoption among organic farms; relatively low marginal benefit; or more GI application among organic farms because these farms have difficulties to compete with large, conventional farms and/or because these farms are more open-minded towards product differentiation and farm diversification	Organic production, higher FNI (Crowder & Reganold, 2015)
Machinery (MACH)	Less MACH, higher GI adoption (product differentiation to be competitive)	More MACH, higher FNI
Gross Domestic Product per inhabitant in the region of the holding (GDPC)	Higher GDPC, more GI adoption (Van de Pol, 2017); according to Beckmann and Schimmelpfennig (2015) more uptake in poorer regions (farms need to generate higher output prices)	Higher GDPC, higher FNI (willingness and ability to pay for high quality food is expected to be higher)
Farm household consumption (FHC)	Higher FHC of own products, less GI application (Van de Pol, 2017)	Higher FHC, lower FNI (European Commission, 2011)
Km of motorway per 1000 km² (MKM)	The more MKM, the more GI application; Some basic access to infrastructure is required for farms to apply GIs (Van de Pol, 2017)	More MKM, higher FNI; but expected to be insignificant (used as instrument)
Square of MKM (MKM2)	The larger MKM2, the less GI application (more GI application in relatively remote areas) (Meraner et al., 2015; van de Pol, 2017)	More MKM2, higher FNI; but expected to be insignificant (used as instrument)

4.3 Descriptive statistics

Appendix V presents four tables, which show the mean of quality wine specialists and olives specialists for both 2014 and 2015 with respect to the variables of interest. LFA and MA were not considered in 2015 as the question about LFA and MA has not been answered by any farm in 2015.

What is most interesting with respect to the assessment of income effects of GI labels is to what extent adopters and non-adopters differ. Therefore, Table 3 summarizes these differences, which are calculated by taking the mean of non-adopters less the mean of adopters. Among olives specialists, adopters have a higher average FNI than non-adopters. If adoption would be purely random, the differences of 4187 EUR in 2014 and 827 EUR in 2015 could be interpreted as effect of GI adoption. However, as outlined in chapter 3, farmers select themselves into the schemes. In addition, the differences are not significant. For quality wine specialists, the differences are positive and significant, which means that non-adopters earn higher FNI than GI adopters. Again, it would be naïve to immediately conclude that GI farms would be better off if they stopped using GI labels.

For the other variables, no clear patterns can be observed. SPC, UAA, LFA and GDPC are the only variables where differences in all four samples have the same sign, although not all of them are significant. In other cases, signs change either between years (such as with UL, PL or FHC) or between farm types (such as in the case of LIA, EXT, ORG, MA, MACH, MKM and MKM2). On average, GI adopters in these four samples tend to have lower SPC, less UAA and more LFA than non-adopters, and they live in regions which tend to have lower GDPC. While GI wine farms are located in areas with significantly higher MKM and MKM2 than their non-GI colleagues, GI olives specialists live in NUTS2 regions with less MKM and MKM2 than their non-GI colleagues, although the coefficient for MKM2 is not significant. In 2015, an average GI wine farm had one paid AWU more and about five unpaid AWU less than a non-GI wine farm, whereas it has had nearly one paid AWU more than non-adopters in 2014. MA was lower for GI olive farms than for non-GI olive farms. The monetary value of machinery was significantly higher for GI wine farms and significantly lower for GI olive farms in contrast to their non-GI colleagues.

Table 3: Differences between non-adopters and adopters

	Quality Wine Specialists 2014	Quality Wine Specialists 2015	Olives specialists 2014	Olives specialists 2015
FNI	28314.93***	23437.85**	-4186.59	-827.17
OUT	36717.01*	20296.75	-668.96	9415.79
SPC	11247.83	3332.08	1755.08	4435.97**
OVER	219.01	-421.95	1511.12	3458.24
UAA	4.84**	0.03	1.31	7.13
UL	0.06*	-1.08***	-0.09*	-0.01
PL	-0.85***	5.01**	0.09	0.37*
LIA	-25111.22***	-11346.12	3279.99	4959.28
EXT	-1948.93	-3955.00	2845.52	6491.52*
ORG	-0.02**	-0.01	0.02	0.08**
LFA	-0.003	-	-0.17***	-
MA	-0.01	-	0.10*	-
GDPC	3.14***	2.95***	0.55	0.36
MACH	-17683.87***	-16756.01***	7109.89**	8602.09**
FHC	32.21	-17.14	262.61**	-61.07
MKM	-4.79***	-3.77***	3.16**	6.21***
MKM2	-444.87***	-362.51***	87.01	138.70

Note: Difference = mean(non-adopters) - mean(adopters); *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

5. Results

This chapter reports the estimated results for the chosen model. Further, it discusses the reliability and robustness of estimated income effects. For this purpose, it also compares results of the endogenous switching regression model to those of other estimation techniques that were outlined in chapter 3.

Results of the endogenous switching regression estimated by full information maximum likelihood are summarised in Table 4. The Stata command *movestay* was used to estimate the ESR model. The first column presents the estimates of the selection equation (4.3). It is estimated by a probit function with GI adoption (T) being the dependent variable and all the other variables discussed in chapter 4 (including instruments) being explanatory variables. The second and third column report the estimates of the outcome equations for adopters (4.2a) and non-adopters (4.2b) respectively. These models are estimated by an OLS regression with farm net income (FNI) being the dependent variable. Here, explanatory variables do not encompass the chosen instruments. No results were found for wine specialists and olives specialists in the year 2015[6]. The original Stata output of the ESR is shown in appendix VI.

First, I will look at the estimated selection equation. According to the reported estimates for wine specialists in 2014, farms are more likely to use GIs the more overheads, paid labour and liabilities, and the less output, land and farm household consumption they have. This is in line with my initial expectations. For mountain area, the estimated coefficient is positive as expected, but not significant. I expected MACH to have a positive influence on the efficiency and competitiveness of a farm, which I assumed to be negatively correlated with GI adoption. However, the probability to adopt GIs also seems to be higher for farms with more machinery. The coefficients for specific and external costs are negative (although not significant), whereas they were expected to be positive. Organic production seems to positively affect GI adoption, while welfare in the NUTS2 region (indicated by GDPC) decreases the probability of GI uptake. The latter is in line with what Beckmann and Schimmelpfennig (2015) argue. MKM and MKM2 are jointly significant as shown in appendix IV. Both coefficients are positive, although higher MKM2 was expected to decrease the probability of adoption.

Olives specialists in 2014 were more likely to use GIs the more paid labour and LFA they declared to have. The same seems to apply to land and output, which is counterintuitive, as farms with higher output theoretically face less pressure to increase their revenue. However, farms with larger output (e.g. due to higher input of land) also benefit more from gains in farm gate prices than farms with small production volumes. GI adoption and certification might also be relatively expensive and riskier for smaller farms with less output. Specific costs, overheads and costs for labour, land and capital (EXT) seem to have a negative effect on GI adoption, which is also not in line with what I expected. The more machinery a farm used and the more of its own products it consumed, the less likely GI adoption. As expected, the coefficient for the organic production dummy is negative and the coefficients for unpaid labour, liabilities and mountain area are positive, although none of them is significant. Again, MKM and MKM2 are jointly significant, although their coefficients have opposite signs than expected (negative for MKM and positive for MKM2). This is not in line with findings of Van de Pol (2017) who concluded that some basic level of infrastructure is needed to ensure access to the market, while too much infrastructure tends to decreases expected benefits from GI adoption.

The next step is to look at estimated income effects of GI adoption. Column (2) and (3) report the estimates of the outcome equations (4.2a) and (4.2b) with FNI being the dependent variable. Given the result of the likelihood-ratio test for joint independence of the three equations, the outcome equations of adopters and non-adopters are significantly different. Heterogenous effects occur when treated and control group are differently affected by control variables. For example, for both samples, paid labour had a significantly negative effect on FNI for adopters and a significantly positive effect on FNI for non-adopters (see Table 4). Consequently, any kind

[6] I stopped Stata from running the *movestay* command when no results were found after 3000 iterations. The *movestay* command was very sensitive to the choice of variables. Adding more or dropping specific explanatory variables as well as changing the instruments (although they were not valid) enabled Stata to report estimates. However, I stuck to the theoretical model presented in chapter 4.

Table 4: Endogenous switching regression results

	Quality wine specialists 2014					
	(1)		(2)		(3)	
			T = 1		T = 0	
Dep. Variable	T		FNI		FNI	
OUT	-2.4e-06***	(5.5e-07)	1.031***	(0.014)	0.976***	(0.005)
SPC	-3.7e-06	(2.4e-06)	-0.670***	(0.094)	-0.931***	(0.006)
OVER	8.9e-06***	(3.1e-06)	-1.662***	(0.091)	-0.903***	(0.030)
UL	-0.016	(0.073)	-868.921	(2248.921)	-677.855	(901.398)
PL	0.084***	(0.021)	-3775.723***	(495.182)	1549.671***	(274.235)
UAA	-0.015***	(0.002)	337.328***	(58.481)	3.523	(25.512)
LIA	1.4e-06***	(5.0e-07)	-0.048***	(0.013)	0.002	(0.008)
EXT	-3.4e-06	(2.7e-06)	-0.887***	(0.059)	-1.320***	(0.029)
LFA	-0.019	(0.103)	-606.249	(3058.379)	654.874	(1496.878)
MA	0.125	(0.102)	1173.752	(3312.389)	2168.352	(1328.234)
ORG	0.591**	(0.239)	49.598	(6372.195)	2877.903	(3661.693)
MACH	4.5e-06***	(1.1e-06)	-0.082***	(0.018)	-0.107***	(0.014)
GDPC	-0.058***	(0.009)	-364.639	(251.417)	-600.935***	(107.381)
FHC	-1.8e-04**	(7.9e-05)	5.119**	(2.211)	-0.714	(1.028)
MKM	0.003	(0.007)				
MKM2	2.0e-04**	(8.9e-05)				
const.	0.669***	(0.216)	24539.970***	(5994.594)	20010.59***	(2873.082)
σ_i			23269.65**	(1529.517)	16722.97**	(465.848)
ρ_i			-0.7083713**	(0.065)	0.8210964**	(0.027)
N	1237		300		937	

	Olives specialists 2014					
	(1)		(2)		(3)	
			T = 1		T = 0	
Dep. var.	T		FNI		FNI	
OUT	8.9e-06***	(2.3e-06)	1.156***	(0.040)	1.061***	(0.021)
SPC	-6.3e-05**	(2.9e-05)	-0.639*	(0.332)	-0.727***	(0.106)
OVER	-2.2e-05*	(1.2e-05)	-0.744***	(0.255)	-1.672***	(0.082)
UL	0.084	(0.139)	1646.588	(1985.197)	3816.295**	(1582.966)
PL	1.916***	(0.371)	-18537.020***	(7135.293)	12587.280***	(1568.462)
UAA	0.004	(0.004)	-13.599	(41.860)	326.040***	(37.387)
LIA	2.8e-06	(2.1e-06)	-0.149	(0.095)	-0.027**	(0.013)
EXT	-2.0e-04***	(3.0e-05)	0.025	(0.627)	-1.166***	(0.077)
LFA	0.306*	(0.166)	-693.197	(2307.963)	93.734	(1950.510)
MA	0.169	(0.189)	1834.965	(2441.849)	2177.679	(1809.565)
ORG	-0.094	(0.186)	4328.610*	(2510.138)	927.272	(2124.726)
MACH	-7.5e-06**	(3.5e-06)	-0.142**	(0.055)	-0.070***	(0.026)
GDPC	-0.011	(0.022)	426.178	(372.9314)	-455.142**	(187.062)
FHC	-3.9e-04***	(1.2e-04)	-2.637	(1.966)	-2.121**	(0.832)
MKM	-0.016*	(0.009)				
MKM2	1.4e-04	(1.0e-04)				
const.	-0.250	(0.216)	-6551.09	(7139.163)	6985.474*	(4048.694)
σ_i			5113.386**	(598.425)	22406.59**	(643.143)
ρ_i			-0.1308662	(0.618)	0.949151**	(0.047)
N	1036		55		981	

Note: Stata command *movestay FNI OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC, select(T=MKM MKM2)* was used. Standard errors in parentheses. σ_i denotes the square root of the variance of the error terms w_{1i} and w_{0i} in the outcome equation (4.2a) and (4.2b), respectively; ρ_i indicates the correlation coefficient between the error term of the selection equation (4.41) and the error term of the outcome equations (4.2a) and (4.2b), respectively. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

of model that assumes common impacts (such as the PaRCI model described in chapter 3) is likely to give biased estimates because there is heterogeneity in the sample.

Again, I first look at wine specialists in 2014. FNI tends to be higher the more output and land the farm has. In contrast, it decreases with increasing specific costs, overheads and costs for labour, land and capital (EXT). This is in line with initial predictions from economic theory. Overheads seem to play a more important (negative) role for GI farms, whereas the coefficient of external costs is larger for non-adopters. Paid labour has a significantly negative marginal effect of -3776 EUR on FNI of GI farms. For non-adopters, it has a significant marginal effect of +1550 EUR. It seems that paid labour is less productive on GI farms. The coefficients of unpaid labour are nearly the same for treated and untreated farms, although they are not significant. Interestingly, their sign is also negative, which means that the use of unpaid inputs tends to decrease FNI. Land does not have a significant impact on farm income of non-adopters. Even the coefficient is very low. For GI farms, however, each hectare of land increases farm income by about 337 EUR. In contradiction to its expected effect, machinery turns out to negatively influence FNI. Further, liabilities decrease FNI of adopters, whereas GDP per inhabitant in the NUTS2 region decreases FNI of non-adopters. Further, consumption of own products by GI farms is said to have a positive impact, which is not in line with initial predictions based on economic theory. T-statistics for LFA, MA and ORG are too low to show significant estimates. However, it is interesting that LFA is estimated to be negative for GI adopters and positive for non-adopters.

The latter also applies to olives specialists in 2014. Like in the case of wine specialists, and in accordance with my expectations, higher output increases FNI, whereas an increase in specific costs and/or overheads decreases FNI of both treated and untreated farms. The coefficient of unpaid labour is only significantly positive for untreated farms. Same applies to land, where FNI of GI farms tends to be even lower, the more land is cultivated. However, this negative coefficient is not significant. Liabilities, external costs and farm household consumption are negatively correlated with FNI, but only significant for non-adopters. Machinery has a negative impact for both treated and control group, although I predicted it to be positively correlated with FNI. Higher GDP per inhabitant seems to significantly decrease farm net income of non-adopters, while the coefficient for adopters is nearly of the same magnitude but positive (although not significant). This could be linked to the willingness (and ability) to pay for quality products with GI labels (Van de Pol, 2017). LFA and MA do not have any significant effect. One last and very striking point to mention is the effect of paid labour. For treated farms, adding one paid AWU decreases farm net income by -18537 EUR. For untreated farms, the effect is +12587 EUR. Both coefficients are significant at 1% level. A similar pattern can be observed for wine specialists, but the difference in magnitude is much larger in the case of olives specialists.

Table 4 also reports estimates for rho$_1$ (ρ_i for GI farms) and rho$_2$ (ρ_i for non-GI farms), which are correlation coefficients between the error term of the selection equation (4.3) and the error terms of the outcome equations (4.2a) and (4.2b) respectively. If they are significant, there is a problem of unobserved self-selection that would cause other estimation techniques to generate biased estimates. Indeed, rho$_2$ is significantly positive for wine specialists and olives specialists in 2014. This suggests that non-adopters earn less than a random farm would have earned when not applying GIs (Lokshin & Sajaia, 2004). Rho$_1$ is only significantly negative for wine specialists. According to Lokshin and Sajaia (2004), it indicates that GI wine specialists earn higher FNI than a random farm from the wine sample would earn when adopting GIs. An insignificant rho means that the respective group (treated or untreated) does not earn more or less than a random farm would earn with the adoption status.

Finally, Table 5 presents the estimated treatment effects for treated (ATT) and untreated farms (ATU). The estimated sample means of treated ($E(Y_i^1 | T_i=1)$) and untreated farms ($E(Y_i^0 | T_i=0)$) are close to the true sample means. For quality wine specialists in 2014, farm net income was 20774 EUR (ESR estimate is 20375 EUR) for treated farms and 49089 EUR (ESR estimate is 48201 EUR) for untreated farms. Counterfactuals were estimated to be 41678 EUR for treated farms if they had not used GIs, and 82192 for untreated farms if they had used GIs. This leads to an estimated treatment effect for treated farms of -21303 EUR. Such a significantly negative ATT is counterintuitive. It is rather unlikely that farms adopt food quality schemes such as GIs if they lead to such a decrease in farm net income. A slightly negative effect might be expected for the first year(s) after the adoption of GIs if investments have to be made. However, a 50% decrease in farm income due to GI adoption is unlikely

Table 5: Income effects of GIs based on the ESR[7]

Quality wine specialists 2014			
	Decision stage		
	Adoption	No adoption	Treatment effects
GI farms	(a) $E(Y_i^1 \mid T_i=1)$ = 20374.75	(c) $E(Y_i^0 \mid T_i=1)$ = 41677.94	ATT = -21303.19
Non-GI farms	(d) $E(Y_i^1 \mid T_i=0)$ = 82192.13	(b) $E(Y_i^0 \mid T_i=0)$ = 48201.48	ATU = 33990.65
Heterogeneity effects	BH_1 = -61817.38	BH_0 = -6523.54	TH = -55293.84

Olives specialists 2014			
	Decision stage		
	Adoption	No adoption	Treatment effects
GI farms	(a) $E(Y_i^1 \mid T_i=1)$ = 27563.52	(c) $E(Y_i^0 \mid T_i=1)$ = 70759.95	ATT = -43196.43
Non-GI farms	(d) $E(Y_i^1 \mid T_i=0)$ = 22453.74	(b) $E(Y_i^0 \mid T_i=0)$ = 20686.36	ATU = 1767.38
Heterogeneity effects	BH_1 = 5109.78	BH_0 = 50073.59	TH = -44963.81

Format based on: Di Falco et al. (2011), p. 837
Note: (a) and (b) are observed, while (c) and (d) are counterfactuals
T_i=1 if the farm produces (ingredients for) GI products, T_i=0 if the farm does not use any GI label
Y_i^1*:* farm income if farm adopted GIs
Y_i^0: farm income if farm did not adopt GIs
TT: treatment effect on the treated (GI farms)
TU: treatment effect on the untreated (non-GI farms)
BH_t: effect of base heterogeneity for farms that adopted GIs (T=1), and those who did not adopt GIs (T=0)
TH = (TT-TU), i.e. transitional heterogeneity

and would stop farms from producing (ingredients for) GI products. In contrast, the ATU is estimated to be positive. According to the results, non-adopters would earn about 34000 EUR more if they would use GIs. This would be a significant increase in farm net income, which would convince non-adopters to start using GIs unless market entry is not restricted, or additional GI products would not lead to an overall decrease in price premiums. According to the authors of the *movestay* Stata command (Lokshin & Sajaia, 2004), the significant rho_1 and rho_2 for the sample of quality wine specialists indicate, the average income of non-adopters is lower compared to a random farm given that it does not apply GIs, whereas adopters earn more than a random farm would earn if it applied GIs. The estimates, however, suggest that non-adopters always earn more than adopters.

For olives specialists in 2014, the true sample mean of FNI was 27564 EUR for adopters and 23377 EUR for non-adopters. The ESR model estimated these parameters to be 27564 EUR and 20686 EUR respectively. According to the ESR results, GI olives specialists would have earned 70760 EUR if they had not used GIs, which is 43196 EUR more than they effectively earned and 50074 EUR more than FNI of true non-adopters. Here, the interpretation of rho_1 is in line with the estimated figures. In contrast, untreated olives specialists would have earned 22454 EUR if they had adopted GIs, which is 1767 EUR more than they in fact earned. If all olives specialists would use GIs, initial adopters would earn about 5110 EUR more than the initial control group.

To sum up, the results indicate that both unobserved self-selection and heterogeneous effects between treated and control group do play a role. Both can be accounted for by the ESR model. Therefore, ESR seems to be the most promising estimation technique compared to the other techniques presented in chapter 3. However, it is not realistic that GI adopters accept such a decline in income caused by the adoption of GIs, while non-adopters do not immediately start using GIs if their income is estimated to increase (especially for wine farms). To get at least few significant coefficients and income effects, I did not use robust standard errors, although this would give more reliable results in terms of the significance of estimates. It is also important to realise that the estimates

[7] Stata command *movestay FNI OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC, select(T=MKM MKM2)* was used. For 2015, LFA and MA had to be excluded because of missing data. No results were found for 2015. I stopped when no results were found after 3000 iterations. The complete Stata output is presented in appendix VI.

for 2014 refer to the joint effect of PDO and PGI labels. The reported ESR results do not provide any information about the effect of TSG and mountain product labels. The next paragraph will discuss several results of impact assessments based on alternative estimation techniques to show how sensitive estimates are to the chosen model and estimation technique.

Robustness of the impact estimates

As outlined in chapter 3, different estimation techniques can be used to estimate effects of specific programmes or policies. Appendix VII contains the Stata output for these alternative estimations. To put the ESR results into perspective, I will briefly present these results. First, the PuR model is a simple t-test for farm net income (see descriptive statistics). It suggests that GI wine specialists in 2014 and 2015 earn significantly less than their non-GI colleagues, whereas GI olives specialists on average earned more than their non-GI colleagues. However, the differences between GI and non-GI olives specialists in 2014 and 2015 are not significant. Because GI adoption is not randomly assigned to farms, the naïve t-test is not a reliable estimate of the income effects of GIs. The next simplest way to estimate income effects is to run a simple OLS model with common impacts, so same coefficients for adopters and non-adopters (PaRCI model). Using the same variables as in the ESR model (not including the instruments used for the selection equation) results in an estimated average treatment effect of 810 EUR for wine specialists in 2014 and -2774 EUR for olives specialists in 2014. However, both estimates are not significant. The ESR model has shown that there are heterogenous impacts of explanatory variables between treated and untreated farms. The PaRVI model also allows treated and untreated farms to be affected differently by the explanatory variables. Indeed, coefficients of many interaction terms between explanatory variables and the treatment variable turn out to be significant in the PaRVI model, especially for the sample of quality wine producers. The estimated income effect is 687 EUR for wine specialists in 2014 and -2753 EUR for olives specialists in 2014. As outlined in chapter 3, PSM accounts for self-selection based on observables. For PSM to produce unbiased results, GI adoption must not be affected by unobserved characteristics. Further, explanatory variables must not be affected by GI adoption. In addition, good matches must be available. For the PSM model, I used the same variables as for the selection equation of the ESR model. For wine specialists in 2014, PSM estimates the ATT to be -9907 EUR when using the nearest neighbour matching method. There are treated and untreated farms with similar propensity scores. Most differences between the two groups are no longer significant after matching. The hypotheses of differences being equal to zero is only rejected for less favoured area and mountain area, although differences were not significant before matching. Using the 5-nearest-neighbours matching and kernel matching technique results in estimated ATTs of -1534 EUR and 1421 EUR respectively. This shows how sensitive impact estimations are to the choice of the matching technique. For olives specialists, PSM estimates the ATT to be 12170 EUR with nearest neighbour matching, 2304 EUR with 5-nearest-neighbours matching, and 4896 EUR with kernel matching. Overall, propensity scores are estimated to be rather low for treated and untreated farms. A reason why ESR was finally preferred over PSM was that ESR accounts for unobserved heterogeneity, which PSM does not. Unknown attitudes, age and education of farmers likely affect the decision to adopt GIs. Interestingly, the estimated treatment effects for wine specialists in 2015 only range between 1238 EUR (PaRCI model) and 3878 EUR (PSM, NN(1) matching), except for the naïve ATT (t-test) of -23438 EUR. The estimates for olive specialists in 2015 are as diverse as those of 2014 and vary between -5500 EUR (PSM, kernel matching) and 10811 EUR (PSM, NN(1) matching). Although the ESR model with full information maximum likelihood estimation seems to be the most appropriate estimation technique, the results differ enormously from initial expectations and intuition. They are also far more negative than the PSM estimates. In addition, the suggested interpretation of rho1 and rho2 was not in line with estimated figures of the same Stata output. Apart from potential errors with respect to the application of the *movestay* Stata command, there are also shortcomings with respect to the available data which might explain some of the variation in estimates and the unexpected magnitudes and signs. They are discussed in more detail in the last chapter to give some recommendations for further research.

6. Conclusion and discussion

This thesis aimed at investigating the effect of geographical indications (GIs) for food products on farm income within the EU. Four sub-questions were asked, which were answered step by step throughout the thesis. This chapter summarizes the findings to give some concluding and questioning remarks on the research, including a general discussion about potential pitfalls and shortcomings of the impact assessment.

Firstly, potential effects of GIs on farm income were discussed in chapter 2. Adoption of GIs is linked to product differentiation, which is intended to increase farm gate prices. Market power of GI farms depends on the price elasticity of demand, competition from imperfect but close substitutes, the number of farms producing the same GI product, and the market share and competitiveness of the farm with respect to colleagues/competitors who produce the same GI product. Farmers who produce final products within the schemes of Protected Designation of Origin (PDO) or mountain products are expected to benefit the most from higher market power and gains in profits. Market entry for farms producing products within the schemes for Traditional Speciality Guaranteed (TSG) or Protected Geographic Indication (PGI) is less restricted, which reduces market power. In addition, farms that only produce ingredients for the products are expected to face market power from downstream players of the supply chain (e.g. processors or retailers), which decreases their market power. To conclude, GI adoption does not necessarily lead to higher farm profits. However, if farmers behave as profit-maximisers, they only adopt GIs if adoption does not decrease their profits in the long run.

Secondly, five popular estimation techniques used for impact analysis were outlined in chapter 3. The simplest model is that of pure randomisation (PuR), which assumes treatment (GI adoption) to be purely random. It is equal to a t-test for the difference in farm income between treated and untreated farms. I expected the results to be biased as GI adoption is not randomly assigned. With partial randomisation models estimated by Ordinary Least Squares (OLS) regressions, GI adoption is assumed to be exogenously assigned conditional on some observable characteristics. One OLS model was introduced that assumes the same impacts of covariates for treated and untreated farms (PaRCI model with common impact). Another model allows for varying impacts of covariates between treated and control group (PaRVI model for heterogeneous effects). Propensity Score Matching (PSM) assumes treatment (GI adoption) to be endogenous. However, it can only account for observed differences between treated and untreated farms. Finally, the endogenous switching regression (ESR) model was introduced, which also allows for unobserved heterogeneity. Indeed, farms are not exogenously assigned to the group of GI adopters or the control group. Therefore, PSM and ESR models were preferred. Since the given datasets did not contain any information about the farmer, his/her attitude towards and experiences with GIs, farm diversification or product differentiation, and his/her age or education, some potential determinants of GI adoption where unobserved. Therefore, the impact assessment was chosen to be based on an ESR model. Other estimation techniques were used to illustrate the variation in estimated income effects.

Thirdly, I was interested in differences between GI adopters and non-adopters. For the analysis, data from the Farm Accountancy Data Network and EUROSTAT was combined. To reduce the risk of estimation bias, income effects were only measured comparing farms of a specialised farm type that only produce (ingredients) for GI products to farms of the same specialised farm type who do not use any GIs. Quality wine specialists and olives specialists were the specialised farm types with the largest numbers of GI adopters. On average, GI olives specialists had a higher farm net income than their non-GI colleagues, although the difference was not significant. For wine specialists, non-GI farms earned significantly more. For the other variables, no clear patterns were found. GI adopters tended to have lower specific costs, be more exposed to less favoured area, have less land input, and live in regions with less gross domestic product per inhabitant (not all differences being significant). While GI wine farms were located in areas with significantly more kilometres of motorway than their non-GI colleagues, GI olives specialists lived in NUTS2 regions with significantly less kilometres of motorway than their non-GI colleagues. Mixed evidence was also found for less favoured area, farm household consumption of own products, labour input and use of machinery. Since no baseline data was available and it is unknown for how many years adopters are already applying GIs, it is not clear to what extent non-adopters and adopters differed before the latter start using GIs.

Finally, the impact of GIs on farm income was assessed by using an endogenous switching regression model. Results were estimated by full information maximum likelihood (Stata command *movestay*). For both farm types, the impact assessment was done for 2014 and 2015. Given the chosen model, Stata did not provide results for the samples of 2015. 2014 data only considered PDO and PGI labels, so the estimated effects do not provide information about the impact of TSG and mountain product labels. For wine specialists in 2014, the effect of GIs on farm net income was estimated to be -21303 EUR for treated farms. Untreated farms would have earned 33991 EUR more if they had adopted GIs. While the average treatment effect for GI olives specialists was estimated to be -43196 EUR, the average treatment effect for untreated farms was estimated to be 1767 EUR. These estimates contradict the theory since adopters are assumed to only adopt GIs if they do not decrease their farm profits. From a theoretical perspective, it is also not expected that treatment effects for untreated farms are significantly higher than for treated farms. The results were also compared to those of other estimation techniques. PSM estimates of the income effect for olives specialists tended to be positive for both years. The same applies to the joint effect of PDO, PGI, TSG and mountain product labels for wine specialists, whereas the PSM results suggested a negative joint effect of PDO and PGI labels on farm income of GI wine farms. In contrast, Ordinary Least Squares regression estimated positive income effects for wine specialists, and negative income effects for olives specialists.

General discussion

There are some general shortcomings with respect to the chosen model and data, which might have caused estimates to be biased or Stata having problems to estimate any effects for 2015 when using the *movestay* command. First, it is not possible to distinguish between farms who produce final GI products and those who only produce ingredients for GI products. Chapter 2 has shown that expected gains in market power are higher if the GI farmer produces final products instead of ingredients for GI products. Further, it is unknown whether those farms who only produce ingredients for GI products are always aware that their raw products are finally sold as GI product. If they are not aware that their raw products are later turned into a GI product (reporting to be non-adopters), they end up in the control group instead of the treated group. In 2014, the question about GIs explicitly asked for PDO and PGI labels. In 2015, also TSG and mountain products were encompassed. It is unknown, which GI scheme was exactly applied by adopters. The results for 2014 can only be interpreted as joint effect of PDO and PGI labels. As outlined in chapter 3, the PDO scheme is expected to increase market power of farmers more than PGI labels because market entry is more restricted, and ingredients cannot be substituted from all over the world. Therefore, the income effect of PDOs alone might be higher than the estimates of the ESR model that are reported in chapter 5. Interpretation of the results for 2015 is even harder as they refer to the joint effect of four GI schemes.

The FADN data also contained farmers with missing data, e.g. with respect to the GI question or the question about LFA and MA. The latter was not answered by any farmer in 2015. This might have caused Stata not being able to estimate the ESR model for 2015 since results were very sensitive to adding or dropping variables. In addition, there some wine specialists reported to produce only (ingredients for) GI products, but a further question about whether the majority of his/her vineyards were used to produce GI products was answered by no. The same applies to olives specialists. Such errors or inconsistencies in the data likely produce biased and inconsistent estimates. I considered all farms to be GI farms who indicated to only produce (ingredients for) GI products, independent of their answer to the specific GI questions about vineyards or olives. Follow-up research could even exclude farms with inconsistent data. 2330 and 2555 farms who reported to produce some (ingredients for) GI products in 2014 and 2015 respectively were excluded from the analysis because it was unknown how much of the production was affected by the GI scheme. For the same reason I also preferred analysing income effects for specialists only. Otherwise, the farm income would depend on different branches of farming industry. Since I only compared GI wine specialists with non-GI wine specialists and GI olives specialist with non-GI specialists, I ensured that I do not compare farms that are not able to produce the same product due to different climatic or geographic characteristics. Follow-up research could control for NUTS2 regions to account for the different numbers of GIs that are already published and used in the different NUTS2 regions. However, this would not allow to control for MKM, MKM2 and GDPC because these variables are measured at NUTS2 level. Adding control variables for NUTS2 regions to the applied model might lead to multicollinearity.

The given data did not allow to differentiate between quantity and prices both of input and output. A high output could result from high production volumes or high farm gate prices. This makes it even harder to observe (the reason of) differences between adopters and non-adopters. Even more problematic, it is to what extent control variables of GI farms are affected by GI adoption. All the discussed estimation techniques require control variables to be independent of treatment (Khandker et al., 2010). For example, assume that small farms with much unpaid labour adopted GIs before 2014 to gain market power. Due to adoption, farm gate prices, output and income increased over time. Such farms invested in both land and (paid) labour to increase their production and earn even higher profits. Assume that these farms end up with the same or higher average amount of land and paid labour as non-adopters, who tended to be larger and using more paid labour anyway. Looking at 2014 data only, land and paid labour would be either estimated to have no effect on GI uptake (because averages of these variables are nearly the same for treated and untreated farms in 2014), or the effect would be estimated to be positive. Consequently, the model would estimate higher probabilities of GI adoption for farms with more land and paid labour, although it is a small farm with less paid labour (and more unpaid labour) that is more likely to adopt GIs. The estimated causal effects would be biased.

To prevent such misinterpretation, it is best to use baseline data (Yao et al., 2017). Baseline data reflects the conditions under which some farms decided to use GIs in the future, so determinants of uptake can clearly be identified. Consequently, the impact assessment based on non-baseline data must be treated with caution as it cannot be guaranteed that farm specific characteristics are independent of treatment. Descriptive statistics for baseline data would give reliable information about differences between treated and control group before GIs are adopted by the treated group. The cheapest and fastest way to get baseline data would be to collect retrospective data for those farms who are part of the treated and control groups of the sub-samples used for the impact assessment conducted in chapter 5. Such baseline data helps investigating the relevance of the different variables for the GI adoption. Further, questions about why farms decided (not) to use GIs could add information about determinants of GI adoption. Instead of collecting data from a representative sample for the EU, case studies could be an alternative and relatively quick and cheap way to assess the impact of GIs on farm income, such as Bouamra-Mechemache and Chaaban (2010) have shown for the case of PDO Brie. Case studies could investigate whether changes in farm gate prices and production costs of adopters are caused by GI adoption or general trends/changes in farm gate prices, or input prices and quantities (e.g. higher cost for irrigation in years with low precipitation). It could also focus on farms who are producing (ingredients for) the same GI product to determine whether there are differences in income effects.

Finally, it is debateable whether I selected all (and only) relevant variables from the dataset to assess the income effect. For example, land could also be separated into leased land and land that is owned by the farm. Van de Pol (2017) also added determinants such as food culture or population density at NUTS2 level. Instead of farm income, research could also deal with effects on revenues and costs separately. The estimated effect on farm income could be negative (or neutral) because production costs increase relatively more than (or as much as) revenues. If this was true, policy makers could think about subsidising certification or investments for GI adopters to allow their gains in revenues to exceed the increase in their production costs. Information about the pre-treatment characteristics of adopters could also answer the question whether GIs are adopted by those farms who are intended to benefit (less competitive farms).

In conclusion, this research on the impact of geographic indications on farm income does not provide reliable figures about the monetary effect. However, I elaborated on the theoretical background and potential mechanisms by which GIs can influence farm income, as well as different estimation techniques and their pros and cons with respect to the impact assessment of GIs. The broad discussion of shortcomings and potential pitfalls can help to better prepare and/or improve future research. Previous research mainly focussed on consumer behaviour and the impact of GIs on the willingness to pay for food products with GI labels. According to Skuras and Vakrou (2002), Greek consumers are willing to pay more for GI wine. However, attitudes and claims are often not in line with behaviour. Even if consumers are found to (be willing to) pay more for GI products, it does not provide information about whether and what kind of farms benefit from these price premiums. This information is needed to assess the effectiveness of the food quality policies of the EU. I did not find impact assessments based on farm accountancy data. Therefore, the idea to use farm accountancy data to investigate effects of GIs on farm income (or revenues and costs separately) is worth to be pursued.

References

Beckmann, J., & Schimmelpfennig, D. (2015). Determinants of farm income. *Agricultural Finance Review, 75*(3), pp. 385-402.

Borg, E. A., & Gratzer, K. (2013). Collective brand strategy, entrepreneurship, and regional growth: The role of a protected designation of origin (PDO). *Journal of World Economic Research, 2*(3), pp. 26-38.

Bouamra-Mechemache, Z., & Chaaban, J. (2010). Determinants of adoption of protected designation of origin label: Evidence from the French Brie cheese industry. *61*(2), pp. 225-239.

Caliendo, M., & Kopeinig, S. (2008). Implementation of propensity score matching. *Journal of Economic Surveys, 22*(1), pp. 31-72.

Crowder, D. W., & Reganold, J. P. (2015). Financial competitiveness of organic agriculture on a global scale. *Processing of the National Academy of Sciences, 112*(24), pp. 7611-7616.

Di Falco, S., Veronesi, M., & Yesuf, M. (2011). Does adaptation to climate change provide food security? A micro-perspective from Ethiopia. *American Journal of Agricultural Economics, 93*(3), pp. 829-846.

Duflo, E., Glennerster, R., & Kremer, M. (2006). *Using randomization in development economic research: A toolkit.* NBER Technical Working Paper No. 333. doi:10.3386/t0333

Estrin , S., Laidler, D., & Dietrich, M. (2008). *Microeconomics* (5th ed.). Harlow: Pearson Education.

European Commission. (2009). *Why do we need a Common Agricultural Policy? Discussion paper by DG Agriculture and Rural Development.*

European Commission. (2011). *EU agricultural economic briefs. What is a small farm? Brief N° 2.* Agriculture and Rural Development.

European Commission. (2014, June 20). *Commission publishes rules for "mountain product" label and clarifies other elements of EU quality labels*. Retrieved March 07, 2018, from https://ec.europa.eu/agriculture/newsroom/164_en

European Commission. (2016, September 1). *Farm Accountancy Data Network*. Retrieved August 14, 2018, from Sample selection: ec.europa.eu/agriculture/rica/methodology2_en.cfm

European Commission. (2017a). *CAP post-2013: Key graphs & figures. Share of direct payments and total subsidies in agricultural factor income.* Retrieved May 07, 2018, from https://ec.europa.eu/agriculture/sites/agriculture/files/cap-post-2013/graphs/graph5_en.pdf

European Commission. (2017b). *EU agricultural product quality policy*. Retrieved March 07, 2018, from https://ec.europa.eu/agriculture/quality_en

European Commission. (2018). *DOOR database*. Retrieved July 25, 2018, from http://ec.europa.eu/agriculture/quality/door/list.html?locale=en

European Union. (1992). *Council Regulation (EEC) No. 2081/92 of 14 July 1992 on the protection of geographical indications and designations of origin for agricultural products and foodstuffs*. Retrieved April 04, 2018, from http://www.wipo.int/wipolex/en/text.jsp?file_id=126806

Frank, R. H., & Cartwright, E. (2016). *Microeconomics and behavior* (Second ed.). London: McGraw-Hill Education.

Friesland Campina. (2018). *Annual report 2017.* Retrieved July 05, 2018, from https://www.frieslandcampina.com/app/uploads/sites/2/2018/02/FrieslandCampina-Annual-Report-2017.pdf

Gangjee, D. S. (2017). Proving provenance? Geographical indications certification and its ambiguities. *World Development*, pp. 12-24.

Giaccio, V., Giannelli, A., & Mastronardi, L. (2018). Explaining determinants of agri-tourism income: Evidence from Italy. *Tourism review, 73*(2), pp. 216-229.

Giovannucci, D., Josling, T., Kerr, W., O'Conner, B., & Yeung, M. T. (2009). *Guide to geographical indications: Linking products and their origins.* International Trade Centre, Geneva.

Hajdukiewicz, A. (2014). European Union agri-food quality schemes for the protection and promotion of geographical indications and traditional specialities: An economic perspective. *Folia Horticulturae, 26*(1), pp. 3-17.

Hill, B., & Brandley, B. D. (2015). *Comparison of farmers' incomes in the EU member states.* Report for European Parliament, DG Internal Policies.

Huang, W., & Zhang, Y. (2018). Technological gap and heterogeneous oligopoly. *The Quarterly Review of Economics and Finance, 67*, pp. 1-7.

Khandker, S. R., Koolwal, G. B., & Samad, H. A. (2010). *Handbook on impact evaluation. Quantitative methods and practices.* Washington: The World Bank.

Khonje, M., Manda, J., Alene, A. D., & Kassie, M. (2015). Alanysis of adoption and impacts of improved maize varieties in Eastern Zambia. *World Development, 66*, pp. 695-706.

Krugman, P., & Wells, R. (2013). *Microeconomics* (Third ed.). New York: Worth Publishers.

Lokshin, M., & Sajaia, Z. (2004). Maximum likelihood estimation of endogenous switching regression models. *The Stata Journal, 4*(3), pp. 282-289.

Meraner, M., Heijman, W., Kuhlman, T., & Finger, R. (2015). Determinants of farm diversification in the Netherlands. *Land Use Policy, 42*, pp. 767-780.

Nieberg, H., & Offerman, F. (2003). The profitability of organic farming in Europe. Organic agriculture: sustinability, markets and policies. *OECD workshop on organic agriculture* (pp. 141-151). Washington: CABI Publishing.

Nuzzo, R. L. (2016). Statistical power. *Pm&r, 8*(9), pp. 907-912.

Organisation des Nations Unies pour l'alimentation et l'agriculture. (2011). *Safeguarding food security in volatile global markets.* A. Prakash (Ed.). Rome: Food and Agricultural Organisation.

Shadbolt, N., Kelly, T., & Holmes, C. (2005). Organic dairy farming: Cost of production and profitability. *Australian Farm Business Management Journal, 2*(2), p. 136.

Shadish, W. R., & Steiner, P. M. (2010). A primer on propensity score analysis. *Newborn and Infant Nursing Reviews, 10*(1), pp. 19-26.

Skuras, D., & Vakrou, A. (2002). Consumers' willingness to pay for origin labelled wine: a Greek case study. *British Food Journal, 104*(11), pp. 898-912.

Strength2Food. (2016). Retrieved March 16, 2018, from Aims & Objectives: https://www.strength2food.eu/aims-objectives/

Van de Pol, L. (2017). *Explaining the spatial distribution in the uptake of PDO and PGI in Europe.* Wageningen: Wageningen University.

Van Ittersum, K. (2002). *The role of region of origin in consumer decision-making and choice.* Wageningen: Mansholt Graduate School.

Varian, H. R. (2014). *Intermediate microeconomics* (9th ed.). New York: W. W. Norton & Company.

Verbeek, M. (2012). *A guide to modern econometrics.* Chichester: John Wiley & Sons Ltd.

Verbeke, W., Pieniak, Z., Guerrero, L., & Hersleth, M. (2012). Consumers' awareness and attitudinal determinants of European Union quality label use on traditional foods. *Bio-based and Applied Economics, 1*(2), pp. 213-229.

Yao, X. I., Wang, X., Speicher, P. J., Hwang, E. S., Cheng, P., Harpole, D. H., Berry, M. F., Schrag, D., & Pang, H. H. (2017). Reporting and guidelines in propensity score analysis: A systematic review of cancer and cancer surgical studies. *Journal of the National Cancer Institute, 109*(8), p. djw323.

Appendix I: GI registrations EU28

Table 6: Registered PDO, PGI and TSG products per EU member state; Source: DOOR database (July, 2018)

Country	PDO registrations	PGI registrations	TSG registrations	Total
ITA	167	127	2	296
FRA	104	124	1	247
ESP	103	89	4	195
POR	64	74	1	139
ELL	76	30	0	106
DEU	12	78	0	90
UKI	26	41	4	71
POL	8	22	9	39
CZE	6	23	5	34
SVN	8	13	3	24
HRV	10	9	0	19
BEL	3	11	5	19
SVK	2	10	7	19
OST	10	6	1	17
HUN	6	8	1	15
NED	6	5	4	15
SUO	5	2	3	10
SVE	3	3	2	8
IRE	3	4	0	7
LTU	1	4	2	7
DAN	0	7	0	7
BGR	0	2	5	7
CYP	1	4	0	5
LVA	1	1	3	5
LUX	2	2	0	4
ROU	1	3	0	4
EST	0	0	0	0
MLT	0	0	0	0

Appendix II: Sample selection

Sample Set 1 = full FADN dataset with 28 member states of the EU

Sample Set 2 = FADN dataset excluding Croatia and those NUTS2 regions where all farms have missing data, so excluding DEU, FRA, LTU, LUX, LVA, SVK, BE21, BE22, BE23, BE24, BE25, AT32, AT33, AT34, and all of UKI except for UKN0 for 2014 and 2015, and IRE for 2014 only

Sample Set 3 = based on SET 2, but excluding all farms with code number 3 (some GI) or 0 (missing data) for the GI variable; the sample of quality wine specialists and those of olives specialists are used for the impact assessment; they are sub-samples taken from sample set 3

Table 7: From raw data to the final sample

		2014		2015	
SET 1	GI	546	1%	592	1%
	non-GI	52795	64%	53285	65%
	some GI	2344	3%	2566	3%
	GI=0	26762	32%	25780	31%
		82447	100%	82223	100%
SET 2	GI	533	1%	579	1%
	non-GI	51600	90%	52027	90%
	some GI	2332	4%	2555	4%
	GI=0	2980	5%	2847	5%
		57445	100%	58008	100%
SET 3	GI	533	1%	579	1%
	non-GI	51600	99%	52027	99%
		52133	100%	52606	100%

Appendix III: List of variables

Table 8: List of variables, abbreviations and descriptions

Variable Code	Original Code	Source	Description	Scale	Level
T	A_CL_150_C	FADN	**2014:** indication for whether the holding produces agricultural products and/or foodstuffs protected by a **PDO or PGI** or whether it produces agricultural products which are known to be used to produce foodstuffs protected by PDO/PGI within the meaning of Council Regulation (EC) No 510/2006 **2015:** indication for whether the holding produces agricultural products and/or foodstuffs protected by a **PDO/PGI/TSG/mountain product** or whether it produces agricultural products which are known to be used to produce foodstuffs protected by PDO/PGI/TSG/mountain product within the meaning of Council Regulation (EC) No 1151/2012 **Code numbers**: 1) no 2) only 3) some First, all farms with A_CL_150_C=3 were excluded from the sample (those with missing data as well). Next, a dummy variable was generated with T=1 if A_CL_150_C=2 and T=0 if A_CL_150_C=1.	Dichotomous (yes/no)	Farm
FNI	SE420	FADN	**Farm net income** in EUR; remuneration to fixed factors of production (work, land, capital) and remuneration to the entrepreneurs' risk (loss/profit) in the accounting year; = Total output (SE131) – Total intermediate consumption *(=Total specific costs + Total farming overheads*; SE275*)* + Balance current subsidies and taxes (SE600) – Depreciation (SE360) + Balance subsidies and taxes on investments (SE405) – Total external factors *(= Wages paid + Interest paid + Rent paid*, SE365*)*; If unpaid (family) labour > 0, FNI = Family Farm Income (FFI)	Continuous	Farm
OUT	SE131	FADN	**Total output** in EUR	Continuous	Farm
SPC	SE281	FADN	**Total specific costs** in EUR	Continuous	Farm
OVER	SE336	FADN	**Total farming overheads** in EUR; supply costs linked to production activity but not linked to specific lines of production	Continuous	Farm
PL	SE020	FADN	**Paid labour** input in annual working units (AWU)	Continuous	Farm
UL	SE015	FADN	**Unpaid labour** input in AWU	Continuous	Farm
LIA	SE485	FADN			
UAA	SE025	FADN	**Total utilised agricultural area** in ha; does not include areas used for mushrooms, land rented or less than one year on an occasional basis, woodland and other farm areas (roads, ponds, non-farmed areas, etc.); it consists of land in owner occupation, rented land and land in share-cropping; it includes agricultural land temporarily not under cultivation for agricultural reasons or being withdrawn from production as part of agricultural policy measures	Continuous	Farm
EXT	SE365	FADN	**Total external factors** in EUR; remuneration of inputs (labour, land, capital) which are not the property of the holder; includes wages, rent and interest paid	Continuous	Farm
LFA	Based on A_CL_160_C	FADN	**Less favoured area**; A_CL_160_C has the following code numbers: In 2014: 1) majority of the UAA of the holding is not situated in a less favoured area, 2) majority of the UAA of the holding is situated in a LFA, 3) majority of the UAA of the holding is situated in a mountainous area, 4) the areas are so small and numerous in these member states that the information is not significant	Dichotomous (yes/no)	Farm

			A dummy variable was generated with LFA=1 if A_CL_160_C=2 and LFA=0 for the rest.		
MA	Based on A_CL_160_C	FADN	**Mountainous area**; See LFA; A dummy variable was generated with MA=1 if A_CL_160_C=3 and MA=0 for the rest.	Dichotomous (yes/no)	Farm
ORG	Based on A_CL_140_C	FADN	**Organic production**; A_CL_140_C has the following code numbers: 1) holding does not apply organic production methods, 2) holding applies only organic production methods, 2) holding applies both organic and other production methods, 4) holding is converting to organic production methods A dummy variable was generated with ORG=1 if A_CL_140_C=2 (only organic)	Dichotomous (yes/no)	Farm
MACH	SE455	FADN	Value of **machinery** in EUR	Continuous	Farm
FHC	SE260	FADN	**Farm household consumption** in EUR	Continuous	Farm
GDPC	GDP_inh	Van de Pol (2017)	**GDP per inhabitant** in EUR in a specific NUTS2 region in 2013	Continuous	NUTS2
MKM	Mw_km2	Van de Pol (2017)	**Km of motorway** per 1000km² in a specific NUTS2 region in 2013	Continuous	NUTS2
MKM2	Based on MW_km2	Van de Pol (2017)	**Square of MKM**	Continuous	NUTS2

Appendix IV: Specification tests for instrumental variables

Table 9: Specification tests to find valid instruments

Quality Wine Specialists 2014		
Dependent variable:	GI adoption 1/0	FNI for non-adopters
MKM	0.003 (0.008)	127.054 (108.247)
MKM2	0.0003 (0.0001)	-1.411 (1.698)
Wald test on MKM and MKM2	χ^2=97.55***	F-stat.=0.95
Pseudo R^2	0.161	
Adj. R^2		0.993
Sample size	1237	937
Quality Wine Specialists 2015		
Dependent variable:	GI adoption 1/0	FNI for non-adopters
MKM	0.003 (0.008)	143.165 (197.416)
MKM2	0.0003 (0.0001)	-2.245 (3.123)
Wald test on MKM and MKM2	χ^2=66.53***	F-stat.=0.27
Pseudo R^2	0.1362	
Adj. R^2		0.9802
Sample size	1220	912
Olives specialists 2014		
Dependent variable:	GI adoption 1/0	FNI for non-adopters
MKM	-0.038 (0.012)	-70.364 (145.068)
MKM2	0.0003 (0.0001)	0.956 (1.611)
Wald test on MKM and MKM2	χ^2=10.88***	F-stat.= 0.18
Pseudo R^2	0.092	
Adj. R^2		0.856
Sample size	1036	981
Olives specialists 2015		
Dependent variable:	GI adoption 1/0	FNI for non-adopters
MKM	-0.058 (0.012)	78.755 (163.621)
MKM2	0.0005 (0.0001)	1.635 (1.954)
Wald test on MKM and MKM2	χ^2=23.15***	F-stat.=3.22**
Pseudo R^2	0.1304	
Adj. R^2		0.902
Sample size	992	934

Note: GI adoption 1/0 is the dependent variable of a probit model (Stata command: *probit T OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC MKM MKM2*). FNI for non-adopters is the dependent variable of an OLS regression among non-adopters only (Stata command: *reg FNI OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC MKM MKM2)*. For the samples from the year 2015, LFA and MA were excluded. Standard errors in parentheses. Parameters for the other variables are not reported for simplicity. **Significant at 5% level; ***Significant at 1% level.

Table 9 presents the results of the specification tests which were conducted to test whether the chosen instruments are valid instruments. To be valid instruments, variables must jointly influence the decision to adopt GIs, but they must not jointly affect the outcome (FNI) of non-adopters (Di Falco et al., 2011). Khonje et al. (2015) even test for joint effects on the outcome variable (FNI) for the whole sample. In this case, the impact of MKM and MKM2 was not jointly significant for wine specialists in 2014 and 2015, and for olives specialists in 2014 (results are not reported here). According to the results, MKM and MKM2 are valid instruments for all samples except for olives specialists in 2015.

As mentioned earlier, GIs were introduced to offer opportunities to generate price premiums (higher farm gate prices that allow farmers to earn higher profit margins via product differentiation strategies). From a theoretical point of view, a variable for previous FNI (lagged FNI) might be a valid instrument as it does not affect current FNI, but it affects the decision to adopt GIs as farms with relatively low FNI might be interested in generating higher farm gate prices. Similarly, farms with smaller economic size (indication for what farms potentially could earn; potential production volumes minus variable costs based on averages of the previous years) or farms with less favoured area and mountain area might face a higher probability to adopt GIs. However, these variables also affect FNI and respective specification tests revealed them as no valid instruments for the samples for which the impact assessment was done.

Appendix V: Descriptive statistics tables

Table 10: Descriptive statistics quality wine specialists 2014

	Total mean (std. dev.) *n=1237*	Mean adopters (std. dev.) *N=300*	Mean non-adopters (std. dev.) *N=937*	Difference (T Statistic; P Value)
FNI	42222.21 (164766.8)	20774.28 (116755.4)	49089.21 (176918.4)	28314.93*** (2.5966; 0.0048)
OUT	97812.4 (418577.4)	70000.09 (135028.7)	106717.1 (474564.5)	36717.01* (1.3227; 0.0931)
SPEC	19981.49 (220840.2)	11461.51 (29938.44)	22709.34 (253150)	11247.83 (0.7677; 0.2214)
OVER	13765.22 (36593.87)	13599.32 (37588)	13818.34 (36290.11)	219.0144 (0.0902; 0.4641)
UAA	21.64592 (33.12625)	17.97913 (34.00895)	22.81991 (32.77043)	4.840781** (2.2063; 0.0138)
UL	1.079487 (0.635126)	1.033601 (0.5985943)	1.094178 (0.646005)	0.0605772* (1.4384; 0.0753)
PL	1.024277 (4.035492)	1.667182 (6.171711)	0.8184369 (3.026925)	-0.8487455*** (-3.1822; 0.0007)
LIA	15018.05 (87481.29)	34039.23 (128047.6)	8928.019 (68666.52)	-25111.22*** (-4.3585; 0.0000)
EXT	16762.1 (63523.11)	18238.37 (51280.96)	16289.44 (66989.08)	-1948.926 (-0.4624; 0.3220)
ORG	0.0242522 (0.1538935)	0.04 (0.1962866)	0.0192102 (0.1373366)	-0.0207898** (-2.0391; 0.0208)
LFA	0.2643492 (0.4411644)	0.2666667 (0.4429555)	0.2636073 (0.4408241)	-0.0030594 (-0.1045; 0.4584)
MA	0.2530315 (0.4349247)	0.2633333 (0.4411776)	0.2497332 (0.4330897)	-0.0136001 (-0.4712; 0.3188)
GDPC	23.390222 (6.121261)	21.00867 (6.297804)	24.15272 (5.865967)	3.144054*** (7.9344; 0.0000)
MACH	25309.56 (72912.16)	38704.7 (113578)	21020.83 (53135.18)	-17683.87*** (-3.6746; 0.0001)
FHC	180.3728 (558.403)	155.9709 (624.2835)	188.1855 (535.7288)	32.2146 (0.8696; 0.1924)
MKM	27.37914 (15.02412)	31.00667 (18.75899)	26.21772 (13.41982)	-4.788951*** (-4.8486; 0.0000)
MKM2	975.1593 (1084.254)	1312.14 (1609.139)	867.2679 (822.9254)	-444.8721*** (-6.2806; 0.0000)

Note: Difference = mean(non-adopters) - mean(adopters); *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

Table 11: Descriptive statistics quality wine specialists 2015

	Total mean (std. dev.) *n=1220*	Mean adopters (std. dev.) *n=308*	Mean non-adopters (std. dev.) *n=912*	Difference (T Statistic; P Value)
FNI	42269.61 (183665.9)	24748.86 (39683.37)	48186.71 (210875.5)	23437.85** (1.9385; 0.0264)
OUT	91368.65 (258088)	76196 (133915.6)	96492.75 (288065.8)	20296.75 (1.1935; 0.1165)
SPEC	15426.39 (82779.37)	12935.52 (34808.17)	16267.6 (93584.46)	3332.077 (0.6106; 0.2708)
OVER	12936.33 (36855.33)	13251.76 (29413.94)	12829.81 (39063.34)	-421.9527 (-0.1737; 0.4311)
UL	1.075437 (0.6514622)	1.056446 (0.6852196)	1.08185 (0.6399217)	0.0254041 (0.5916; 0.2771)
PL	1.14044 (4.529599)	1.947362 (6.731757)	0.8679265 (3.447918)	-1.079436*** (-3.6341; 0.0001)
UAA	22.10208 (37.47015)	18.35666 (34.44832)	23.36698 (38.37247)	5.010329** (2.0316; 0.0212)
LIA	22655.66 (151120.5)	31137.35 (131799.4)	19791.23 (157073.4)	-11346.12 (-1.1394; 0.1274)
EXT	16431.02 (59871.29)	19387.55 (54347.21)	15432.54 (61621.25)	-3955.003 (-1.0024; 0.1582)
ORG	0.0204918 (0.1417334)	0.025974 (0.1593166)	0.0186404 (0.1353254)	-0.0073337 (-0.7850; 0.2163)
GDPC	23.04238 (6.196475)	20.83766 (6.195823)	23.78695 (6.0202)	2.949289*** (7.3788; 0.0000)
MACH	24976.13 (74034.24)	37501.93 (109256)	20745.92 (56926.01)	-16756.01*** (-3.4496; 0.0003)
FHC	217.6369 (606.8585)	230.4515 (642.2087)	213.3091 (594.7499)	-17.14238 (-0.4285; 0.3342)
MKM	27.37705 (15.28812)	30.19481 (18.53163)	26.42544 (13.90776)	-3.769367*** (-3.7613; 0.0001)
MKM2	983.0377 (1108.04)	1254.032 (1591.317)	891.5175 (869.637)	-362.5149*** (-5.0133; 0.0000)

Note: Difference = mean(non-adopters) - mean(adopters); *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

Table 12: Descriptive statistics olives specialists 2014

	Total mean (std. dev.) *n=1036*	Mean adopters (std. dev.) *n=55*	Mean non-adopters (std. dev.) *n=981*	Difference (T Statistic; P Value)
FNI	23599.35 (54197.15)	27563.68 (45832.61)	23377.09 (54639.74)	-4186.589 (-0.5573; 0.2887)
OUT	39116.47 (89784.33)	39749.92 (49751.8)	39080.96 (91527.19)	-668.9584 (-0.0537; 0.4786)
SPEC	6914.125 (19117.73)	5252.218 (5028.691)	7007.301 (19607.21)	1755.082 (0.6623; 0.2540)
OVER	7861.482 (20967.46)	6430.582 (6290.878)	7941.705 (21494.32)	1511.124 (0.5199; 0.3016)
UAA	22.28405 (31.05598)	21.03964 (21.94421)	22.35382 (31.49566)	1.314186 (0.3053; 0.3801)
UL	0.8771098 (0.4575304)	0.959076 (0.4383051)	0.8725144 (0.458365)	-0.0865617* (-1.3659; 0.0861)
PL	0.5762083 (1.226083)	.4865304 (0.5181691)	0.5812361 (1.253944)	0.0947057 (0.5572; 0.2887)
LIA	5478.731 (127990.3)	2372.873 (9992.474)	5652.861 (131509.8)	3279.989 (0.1849; 0.4267)
EXT	10211 (27273.26)	7516.545 (7268.374)	10362.06 (27968.46)	2845.517 (0.7528; 0.2259)
ORG	0.1611969 (0.3678901)	0.1454545 (0.355808)	0.1620795 (0.3687117)	0.016625 (0.3260; 0.3723)
LFA	0.3474903 (0.4764031)	.5090909 (.504525)	.3384302 (.473417)	-0.1706607*** (-2.5923; 0.0048)
MA	0.3880309 (0.487537)	.2909091 (.4583678)	.393476 (.4887701)	0.102567* (1.5192; 0.0645)
GDPC	18.43571 (3.866037)	17.91455 (3.081271)	18.46493 (3.904588)	0.5503882 (1.0274; 0.1522)
MACH	21044.76 (29349.15)	14312.33 (16855.82)	21422.21 (29855.87)	7109.885** (1.7500; 0.0402)
FHC	644.7432 (888.6621)	396.0727 (449.1598)	658.685 (905.1297)	262.6123** (2.1363; 0.0164)
MKM	22.25097 (12.2354)	19.25455 (13.96718)	22.41896 (12.11715)	3.164415** (1.8687; 0.0310)
MKM2	644.666 (1053.375)	562.2727 (509.843)	649.2854 (1075.708)	87.0127 (0.5959; 0.2757)

Note: Difference = mean(non-adopters) - mean(adopters); *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

Table 13: Descriptive statistics olives specialists 2015

	Total mean (std. dev.) *n=992*	Mean adopters (std. dev.) *n=58*	Mean non-adopters (std. dev.) *n=934*	Difference (T Statistic; P Value)
FNI	34476.52 (67723.68)	35255.33 (65228.99)	34428.16 (67909.02)	-827.1716 (-827.1716; 0.4641)
OUT	55147.5 (123882.3)	46282.23 (66540.74)	55698.02 (126590.5)	9415.794 (0.5615; 0.2873)
SPEC	7479.388 (19182.46)	3302.776 (3131.15)	7738.749 (19725.39)	4435.974** (1.7106; 0.0437)
OVER	10626.1 (38370.56)	7370.052 (7127.457)	10828.29 (39497.12)	3458.239 (0.6658; 0.2528)
UAA	25.48302 (49.07708)	18.77379 (17.63368)	25.89966 (50.36188)	7.125864 (1.0731; 0.1418)
UL	0.9120108 (0.4675665)	0.9257095 (0.3562048)	0.9111601 (0.4737561)	-0.0145494 (-0.2298; 0.4091)
PL	0.7515294 (1.934834)	0.4052727 (0.5887318)	0.7730314 (1.986759)	0.3677588* (1.4053; 0.0801)
LIA	6080.475 (131465.9)	1411.155 (7874.361)	6370.433 (135471.3)	4959.277 (0.2786; 0.3903)
EXT	12843.46 (36241.08)	6731.483 (9509.492)	13223 (37243.44)	6491.519* (1.3242; 0.0929)
ORG	0.1280242 (0.3342853)	0.0517241 (0.2234038)	0.1327623 (0.3394995)	0.0810382** (1.7934; 0.0366)
GDPC	17.93135 (3.280677)	17.58966 (3.110068)	17.95257 (3.291395)	0.3629144 (0.8173; 0.2070)
MACH	24101.23 (32570.45)	16002.09 (20563.82)	24604.18 (33115.19)	8602.092** (1.9545; 0.0255)
FHC	455.0847 (720.0238)	512.5862 (734.0791)	451.5139 (719.3908)	-61.07229 (-0.6266; 0.2655)
MKM	21.72581 (11.33801)	15.87931 (14.88326)	22.08887 (10.98872)	6.209555*** (4.0790; 0.0000)
MKM2	600.4315 (888.5078)	469.8448 (527.8681)	608.5407 (905.7443)	138.6959 (1.1537; 0.1244)

Note: Difference = mean(non-adopters) - mean(adopters); *Significant at 10% level; **Significant at 5% level; ***Significant at 1%.

Appendix VI: Stata output ESR model

Quality wine specialists 2014

```
. movestay FNI OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC, select(T=MKM MKM2)
```
[8]

```
Iteration 4035:log likelihood = -14238.075

Endogenous switching regression model            Number of obs    =      1237
                                                 Wald chi2(14)    =   8120.72
Log likelihood = -14238.075                      Prob > chi2      =    0.0000
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
FNI_1						
OUT	1.031206	.0143083	72.07	0.000	1.003162	1.05925
SPC	-.6703817	.0939664	-7.13	0.000	-.8545524	-.4862109
OVER	-1.662055	.0913058	-18.20	0.000	-1.841011	-1.483099
UL	-868.9204	2248.921	-0.39	0.699	-5276.725	3538.884
PL	-3775.723	495.1816	-7.62	0.000	-4746.261	-2805.185
UAA	337.3281	58.48145	5.77	0.000	222.7066	451.9497
LIA	-.0481939	.0132621	-3.63	0.000	-.0741871	-.0222008
EXT	-.8869956	.0593553	-14.94	0.000	-1.00333	-.7706614
LFA	-606.2494	3058.379	-0.20	0.843	-6600.562	5388.064
MA	1173.752	3312.389	0.35	0.723	-5318.412	7665.916
ORG	49.59818	6372.196	0.01	0.994	-12439.68	12538.87
MACH	-.082116	.0183013	-4.49	0.000	-.1179858	-.0462462
GDPC	-364.6386	251.4165	-1.45	0.147	-857.4059	128.1286
FHC	5.119466	2.210877	2.32	0.021	.7862264	9.452705
_cons	24539.97	5994.594	4.09	0.000	12790.78	36289.16
FNI_0						
OUT	.9764058	.0047262	206.59	0.000	.9671426	.985669
SPC	-.9313677	.0058614	-158.90	0.000	-.9428558	-.9198797
OVER	-.9033217	.0302375	-29.87	0.000	-.9625861	-.8440573
UL	-677.8546	901.3977	-0.75	0.452	-2444.562	1088.852
PL	1549.671	274.235	5.65	0.000	1012.18	2087.162
UAA	3.522793	25.51176	0.14	0.890	-46.47933	53.52492
LIA	.0019755	.0077426	0.26	0.799	-.0131996	.0171506
EXT	-1.319538	.0291799	-45.22	0.000	-1.376729	-1.262347
LFA	654.874	1496.878	0.44	0.662	-2278.954	3588.702
MA	2168.352	1328.234	1.63	0.103	-434.9389	4771.643
ORG	2877.903	3661.693	0.79	0.432	-4298.883	10054.69
MACH	-.1070556	.0137869	-7.77	0.000	-.1340773	-.0800338
GDPC	-600.9349	107.3806	-5.60	0.000	-811.397	-390.4728
FHC	-.7143821	1.0281	-0.69	0.487	-2.729421	1.300657
_cons	20010.59	2873.082	6.96	0.000	14379.45	25641.73
T						
OUT	-2.43e-06	5.47e-07	-4.45	0.000	-3.51e-06	-1.36e-06
SPC	-3.71e-06	2.40e-06	-1.55	0.122	-8.41e-06	9.87e-07
OVER	8.85e-06	3.08e-06	2.88	0.004	2.82e-06	.0000149
UL	-.0157281	.072543	-0.22	0.828	-.1579099	.1264536
PL	.0838817	.0211709	3.96	0.000	.0423874	.1253759
UAA	-.0151526	.002008	-7.55	0.000	-.0190882	-.0112171
LIA	1.38e-06	4.97e-07	2.78	0.005	4.07e-07	2.36e-06
EXT	-3.40e-06	2.72e-06	-1.25	0.211	-8.73e-06	1.93e-06
LFA	-.0194534	.1028373	-0.19	0.850	-.2210107	.1821039
ORG	.5908514	.2392038	2.47	0.014	.1220207	1.059682
MACH	4.53e-06	1.05e-06	4.33	0.000	2.48e-06	6.58e-06
GDPC	-.0581644	.0085488	-6.80	0.000	-.0749198	-.0414091
FHC	-.000181	.0000786	-2.30	0.021	-.0003351	-.0000269
MA	.1245336	.1022065	1.22	0.223	-.0757874	.3248546
MKM	.0025266	.0068702	0.37	0.713	-.0109388	.0159921
MKM2	.0001954	.0000887	2.20	0.027	.0000217	.0003692
_cons	.669177	.2156592	3.10	0.002	.2464926	1.091861

[8] The first time I ran this command, the results were already found after iteration 2626. The output looked the same except for few minor differences (e.g. d was estimated to be 82192.08 instead of 82192.13). Descriptive statistics did not change, so I could not find the root of the problem.

```
       /lns1 |   10.05491   .0657302   152.97   0.000     9.926076    10.18373
       /lns2 |   9.724539   .0278568   349.09   0.000      9.66994    9.779137
         /r1 |  -.8839072   .1308466    -6.76   0.000    -1.140362   -.6274526
         /r2 |   1.160174   .0833625    13.92   0.000      .996786    1.323561
-------------+----------------------------------------------------------------
     sigma_1 |   23269.65   1529.517                      20456.92    26469.11
     sigma_2 |   16722.97   465.8484                      15834.41     17661.4
       rho_1 |  -.7083713   .0651891                     -.8145359   -.5562957
       rho_2 |   .8210964   .0271596                       .760241    .8676668
------------------------------------------------------------------------------
LR test of indep. eqns. :            chi2(1) =     94.88   Prob > chi2 = 0.0000

. mspredict a, yc1_1

. mspredict b, yc2_2

. mspredict c, yc2_1

. mspredict d, yc1_2

. sum a

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
           a |        300    20374.75    113950.8   -1044998    1429689

. sum b

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
           b |        937    48201.48    176775.5  -184470.1    3446531

. sum c

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
           c |        300    41677.94    106654.2  -907355.3    1451897

. sum d

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
           d |        937    82192.13    246214.2  -146834.9    5424111
```

Olives specialists 2014

```
. movestay FNI OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC, select(T=MKM MKM2)
Iteration 7:   log likelihood = -11861.272
Iteration 8:   log likelihood = -11861.272

Endogenous switching regression model             Number of obs   =       1036
                                                  Wald chi2(14)   =    4320.68
Log likelihood = -11861.272                       Prob > chi2     =     0.0000
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
FNI_1						
OUT	1.156258	.0403245	28.67	0.000	1.077223	1.235292
SPC	-.639275	.3321135	-1.92	0.054	-1.290205	.0116555
OVER	-.7440233	.2545217	-2.92	0.003	-1.242877	-.2451699
UL	1646.588	1985.197	0.83	0.407	-2244.327	5537.504
PL	-18537.02	7138.293	-2.60	0.009	-32527.81	-4546.219
UAA	-13.59926	41.8599	-0.32	0.745	-95.64316	68.44464
LIA	-.1492194	.0945123	-1.58	0.114	-.3344601	.0360213
EXT	.0249483	.6267364	0.04	0.968	-1.203432	1.253329
LFA	-693.1973	2307.963	-0.30	0.764	-5216.721	3830.327
MA	1834.965	2441.849	0.75	0.452	-2950.972	6620.902
ORG	4328.61	2510.138	1.72	0.085	-591.1692	9248.39
MACH	-.1419131	.0550029	-2.58	0.010	-.2497168	-.0341093
GDPC	426.1775	372.9314	1.14	0.253	-304.7547	1157.11
FHC	-2.636877	1.966058	-1.34	0.180	-6.490279	1.216525
_cons	-6551.09	7139.163	-0.92	0.359	-20543.59	7441.412
FNI_0						
OUT	1.060528	.0209255	50.68	0.000	1.019515	1.101541
SPC	-.727107	.1062012	-6.85	0.000	-.9352576	-.5189565
OVER	-1.672482	.0816809	-20.48	0.000	-1.832574	-1.51239
UL	3816.295	1582.966	2.41	0.016	713.7395	6918.851
PL	12587.28	1568.462	8.03	0.000	9513.153	15661.41
UAA	326.0396	37.38703	8.72	0.000	252.7624	399.3169
LIA	-.0265521	.0127917	-2.08	0.038	-.0516233	-.0014809
EXT	-1.166228	.076514	-15.24	0.000	-1.316192	-1.016263
LFA	93.7341	1950.51	0.05	0.962	-3729.194	3916.662
MA	2177.679	1809.565	1.20	0.229	-1369.003	5724.361
ORG	927.2721	2124.726	0.44	0.663	-3237.113	5091.658
MACH	-.0704491	.0260325	-2.71	0.007	-.1214719	-.0194263
GDPC	-455.1417	187.0621	-2.43	0.015	-821.7767	-88.50666
FHC	-2.120604	.8320186	-2.55	0.011	-3.75133	-.4898776
_cons	6985.474	4048.694	1.73	0.084	-949.8198	14920.77
t						
OUT	8.85e-06	2.71e-06	3.27	0.001	3.54e-06	.0000142
SPC	-.0000629	.0000296	-2.12	0.034	-.0001209	-4.84e-06
OVER	-.0000221	.0000128	-1.73	0.084	-.0000471	2.93e-06
UL	.083541	.1393262	0.60	0.549	-.1895333	.3566154
PL	1.915923	.3714704	5.16	0.000	1.187854	2.643992
UAA	.0036618	.0044427	0.82	0.410	-.0050457	.0123694
LIA	2.82e-06	2.10e-06	1.34	0.179	-1.29e-06	6.93e-06
EXT	-.000199	.0000302	-6.58	0.000	-.0002583	-.0001397
MA	.1690114	.1894556	0.89	0.372	-.2023146	.5403375
ORG	-.0940468	.1861373	-0.51	0.613	-.4588692	.2707757
MACH	-7.45e-06	3.54e-06	-2.11	0.035	-.0000144	-5.15e-07
GDPC	-.0105596	.0216437	-0.49	0.626	-.0529804	.0318612
FHC	-.000385	.0001287	-2.99	0.003	-.0006373	-.0001328
LFA	.3061579	.165578	1.85	0.064	-.0183691	.6306848
MKM	-.0160503	.0091151	-1.76	0.078	-.0339155	.001815
MKM2	.0001422	.0001066	1.33	0.182	-.0000668	.0003511
_cons	-.2504908	.4195924	-0.60	0.551	-1.072877	.5718952

```
        /lns1 |   8.539617    .117031    72.97   0.000     8.31024    8.768994
        /lns2 |   10.01711   .0287033   348.99   0.000    9.960853    10.07337
          /r1 |  -.1316211   .6289207    -0.21   0.834   -1.364283    1.101041
          /r2 |   1.823144   .4792419     3.80   0.000    .8838472    2.762441
--------------+----------------------------------------------------------------
      sigma_1 |   5113.386   598.4249                      4065.29    6431.697
      sigma_2 |   22406.59   643.1427                     21180.86    23703.26
        rho_1 |  -.1308662   .6181498                    -.8773827    .8008726
        rho_2 |    .949151   .0474988                     .7083414     .992059
-------------------------------------------------------------------------------
LR test of indep. eqns. :            chi2(1) =    31.24   Prob > chi2 = 0.0000
-------------------------------------------------------------------------------

. mspredict a, yc1_1

. mspredict b, yc2_2

. mspredict c, yc2_1

. mspredict d, yc1_2

. sum a

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
           a |        55    27563.52    45544.74  -2216.811   302582.1

. sum b

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
           b |       981    20686.36    51525.34  -69071.98   816400.6

. sum c

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
           c |        55    70759.95    46726.53   24033.89   323252.3

. sum d

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
           d |       981    22453.74    56526.64    -158330   912268.9
```

Quality wine and olives specialists 2015

I used the same command as for the 2014 samples. LFA and MA were excluded because FADN data for the year 2015 does not provide information about these variables. For both samples, no results were found after 3000 iterations.

Appendix VII: Impact estimation with other estimation techniques

Quality wine specialists 2014

(1) PaRCI model for wine 2014: ATE= 810

```
. reg FNI T OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC
```

Source	SS	df	MS
Model	3.3171e+13	15	2.2114e+12
Residual	3.8400e+11	1221	314493963
Total	3.3555e+13	1236	2.7148e+10

Number of obs	=	1237
F(15, 1221)	=	7031.62
Prob > F	=	0.0000
R-squared	=	0.9886
Adj R-squared	=	0.9884
Root MSE	=	17734

FNI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
T	810.3933	1236.4	0.66	0.512	-1615.311	3236.098
OUT	.9970615	.0041668	239.29	0.000	.9888866	1.005236
SPC	-.9603345	.0056022	-171.42	0.000	-.9713255	-.9493436
OVER	-1.122263	.0259663	-43.22	0.000	-1.173207	-1.07132
UL	-827.5774	875.199	-0.95	0.345	-2544.638	889.4832
PL	-1370.219	195.6651	-7.00	0.000	-1754.096	-986.3418
UAA	35.97933	22.45731	1.60	0.109	-8.079865	80.03852
LIA	-.0007135	.0063453	-0.11	0.910	-.0131623	.0117354
EXT	-1.224921	.0231569	-52.90	0.000	-1.270353	-1.17949
LFA	-417.2486	1366.657	-0.31	0.760	-3098.506	2264.009
MA	1426.128	1253.324	1.14	0.255	-1032.779	3885.035
ORG	2696.911	3307.518	0.82	0.415	-3792.136	9185.959
MACH	-.0730134	.0093859	-7.78	0.000	-.0914276	-.0545992
GDPC	-439.3163	96.08136	-4.57	0.000	-627.8192	-250.8135
FHC	1.109544	.9409462	1.18	0.239	-.736507	2.955594
_cons	12806.35	2600.662	4.92	0.000	7704.092	17908.62

(2) PaRVI model for wine 2014: ATT= 8968.73-8281.43[9]= 687.30

```
. reg FNI T OUT OUT_T SPC SPC_T OVER OVER_T UL UL_T PL PL_T UAA UAA_T LIA LIA_T EXT EXT
> _T LFA LFA_T MA MA_T ORG ORG_T MACH MACH_T GDPC GDPC_T FHC FHC_T
```

Source	SS	df	MS
Model	3.3239e+13	29	1.1462e+12
Residual	3.1598e+11	1207	261790876
Total	3.3555e+13	1236	2.7148e+10

Number of obs	=	1237
F(29, 1207)	=	4378.21
Prob > F	=	0.0000
R-squared	=	0.9906
Adj R-squared	=	0.9904
Root MSE	=	16180

FNI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
T	8968.729	5526.285	1.62	0.105	-1873.464	19810.92
OUT	.9836394	.0047012	209.23	0.000	.9744161	.9928628
OUT_T	.0223582	.0115235	1.94	0.053	-.00025	.0449665
SPC	-.9331262	.0057891	-161.19	0.000	-.944484	-.9217683
SPC_T	.2838926	.0736996	3.85	0.000	.1392991	.4284861
OVER	-.8845044	.030519	-28.98	0.000	-.9443807	-.8246282
OVER_T	-.6817011	.0782794	-8.71	0.000	-.83528	-.5281223
UL	-1185.053	912.4953	-1.30	0.194	-2975.306	605.2
UL_T	357.1744	2009.011	0.18	0.859	-3584.367	4298.716
PL	1412.953	275.6893	5.13	0.000	872.0699	1953.837
PL_T	-4386.041	468.0977	-9.37	0.000	-5304.416	-3467.665
UAA	38.62449	26.32147	1.47	0.143	-13.01642	90.2654
UAA_T	113.7571	48.37162	2.35	0.019	18.85526	208.6589
LIA	-.0067718	.0080944	-0.84	0.403	-.0226525	.0091089
LIA_T	-.0109162	.0123757	-0.88	0.378	-.0351965	.013364
EXT	-1.368245	.0291823	-46.89	0.000	-1.425498	-1.310991
EXT_T	.3681526	.0522045	7.05	0.000	.2657309	.4705742

[9] ATT=E(Y_j|T_i=1, X_i)=E[α^T-α^C+X_i(β^T-β^C)]; α^T-α^C is the coefficient of T; to calculate X_i(β^T-β^C), I multiplied the mean of each variable with the coefficient of the respective interaction term (such as: mean of OUT for adopters* coefficient of OUT_T)

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
LFA	1137.273	1556.948	0.73	0.465	-1917.353	4191.898
LFA_T	-3799.335	2872.272	-1.32	0.186	-9434.535	1835.866
MA	1761.407	1362.08	1.29	0.196	-910.8991	4433.714
MA_T	-1199.352	2977.237	-0.40	0.687	-7040.487	4641.782
ORG	-247.98	3871.468	-0.06	0.949	-7843.535	7347.575
ORG_T	7259.4	6212.426	1.17	0.243	-4928.953	19447.75
MACH	-.1346861	.0141833	-9.50	0.000	-.1625127	-.1068595
MACH_T	.0612911	.0196961	3.11	0.002	.0226487	.0999335
GDPC	-286.0508	111.3203	-2.57	0.010	-504.4537	-67.64799
GDPC_T	-363.1549	223.3422	-1.63	0.104	-801.3369	75.0271
FHC	.092265	1.05899	0.09	0.931	-1.985402	2.169932
FHC_T	4.241009	2.016772	2.10	0.036	.2842406	8.197777
_cons	8126.21	2954.543	2.75	0.006	2329.599	13922.82

(3) PSM (NN(1) matching)[10] for wine 2014; ATT= -9906.52

```
. psmatch2 T OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC MKM MKM2, outcome(FNI) neighbor (1) ate common

Probit regression                               Number of obs   =       1237
                                                LR chi2(16)     =     220.45
                                                Prob > chi2     =     0.0000
Log likelihood = -575.03544                     Pseudo R2       =     0.1609
```

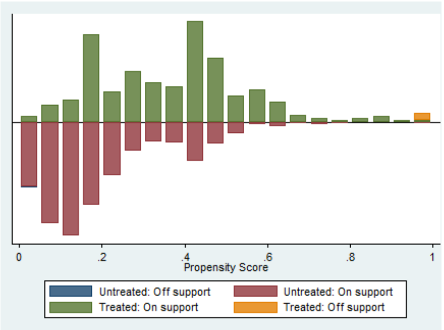
T	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
OUT	-8.82e-07	6.11e-07	-1.45	0.148	-2.08e-06	3.14e-07
SPC	-3.25e-06	2.94e-06	-1.10	0.269	-9.01e-06	2.51e-06
OVER	3.68e-06	2.82e-06	1.30	0.192	-1.85e-06	9.20e-06
UL	.041287	.0777283	0.53	0.595	-.1110577	.1936316
PL	.0046113	.0160152	0.29	0.773	-.026778	.0360006
UAA	-.0087602	.0020432	-4.29	0.000	-.0127649	-.0047555
LIA	1.07e-06	5.59e-07	1.91	0.056	-2.87e-08	2.16e-06
EXT	4.18e-06	2.61e-06	1.60	0.109	-9.37e-07	9.29e-06
LFA	-.0758189	.106861	-0.71	0.478	-.2852626	.1336248
MA	.264883	.1093354	2.42	0.015	.0505895	.4791765
ORG	.5744954	.2490573	2.31	0.021	.0863522	1.062639
MACH	3.09e-06	1.09e-06	2.84	0.005	9.55e-07	5.23e-06
GDPC	-.0874042	.0090332	-9.68	0.000	-.1051089	-.0696995
FHC	-.0001463	.0000838	-1.75	0.081	-.0003104	.0000179
MKM	.0034192	.008249	0.41	0.679	-.0127487	.019587
MKM2	.0003487	.0001065	3.27	0.001	.0001399	.0005575
_cons	.8093245	.2247686	3.60	0.000	.368786	1.249863

Note: 1 failure and 0 successes completely determined.

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	20774.2774	49089.2052	-28314.9279	10904.7962	-2.60
	ATT	26572.8364	36479.3603	-9906.52391	23808.0567	-0.42
	ATU	44876.0791	80245.8908	35369.8116	.	.
	ATE			24474.0431	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	3	934	937
Treated	4	296	300
Total	7	1,230	1,237



[10] 1-nearest-neighbour matching

. pstest OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC MKM MKM2, both treatment (T)

Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
OUT	U	70000	1.1e+05	-10.5		-1.32	0.186	0.08*
	M	67766	74613	-2.0	81.4	-0.40	0.691	0.25*
SPC	U	11462	22709	-6.2		-0.77	0.443	0.01*
	M	10580	9768	0.5	92.8	0.43	0.665	2.21*
OVER	U	13599	13818	-0.6		-0.09	0.928	1.07
	M	12153	10096	5.6	-839.1	1.00	0.316	3.93*
UL	U	1.0336	1.0942	-9.7		-1.44	0.151	0.86
	M	1.0463	.9779	11.0	-12.9	1.40	0.161	0.99
PL	U	1.6672	.81844	17.5		3.18	0.001	4.16*
	M	1.3707	1.5219	-3.1	82.2	-0.30	0.766	0.59*
UAA	U	17.979	22.82	-14.5		-2.21	0.028	1.08
	M	17.117	19.153	-6.1	57.9	-0.90	0.367	2.54*
LIA	U	34039	8928	24.4		4.36	0.000	3.48*
	M	27143	25056	2.0	91.7	0.22	0.825	0.47*
EXT	U	18238	16289	3.3		0.46	0.644	0.59*
	M	14807	14730	0.1	96.0	0.02	0.980	0.74*
LFA	U	.26667	.26361	0.7		0.10	0.917	.
	M	.27027	.34459	-16.8	-2329.4	-1.96	0.050	.
MA	U	.26333	.24973	3.1		0.47	0.638	.
	M	.26689	.19932	15.5	-396.8	1.95	0.052	.
ORG	U	.04	.01921	12.3		2.04	0.042	.
	M	.04054	.05405	-8.0	35.0	-0.77	0.440	.
MACH	U	38705	21021	19.9		3.67	0.000	4.57*
	M	28269	24900	3.8	81.0	0.66	0.507	1.00
GDPC	U	21.009	24.153	-51.7		-7.93	0.000	1.15
	M	21.133	21.097	0.6	98.8	0.08	0.940	1.25
FHC	U	155.97	188.19	-5.5		-0.87	0.385	1.36*
	M	158.08	199.45	-7.1	-28.4	-0.90	0.370	1.68*
MKM	U	31.007	26.218	29.4		4.85	0.000	1.95*
	M	31.334	33.159	-11.2	61.9	-1.15	0.249	0.89
MKM2	U	1312.1	867.27	34.8		6.28	0.000	3.82*
	M	1328.9	1489	-12.5	64.0	-1.25	0.213	1.14

* if variance ratio outside [0.80; 1.25] for U and [0.80; 1.26] for M

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.161	220.45	0.000	15.3	11.4	72.1*	0.46*	69
Matched	0.021	17.16	0.376	6.6	5.8	34.2*	0.82	62

* if B>25%, R outside [0.5; 2]

Differences between the treated group and the control group with respect to PL, UAA, LIA ORG, MACH, GDPC, MKM and MKM2 are no longer significant after matching. Therefore, also the matched pseudo R^2 is lower (0.021 instead of 0.161). However, treated and control group differ with respect to LFA and MA after matching.

(5) PSM (NN(5) matching)[11] for wine 2014; ATT= -1533.85

```
. psmatch2 T OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC MKM MKM2, outcome(FNI) neighbor (5) ate common
Note: 1 failure and 0 successes completely determined.
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	20774.2774	49089.2052	-28314.9279	10904.7962	-2.60
	ATT	26572.8364	28106.6848	-1533.84837	12138.5546	-0.13
	ATU	44876.0791	58849.3395	13973.2603	.	.
	ATE			10241.4683	.	.

Note: S.E. does not take into account that the propensity score is estimated.

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.161	220.45	0.000	15.3	11.4	72.1*	0.46*	69
Matched	0.011	8.79	0.922	5.5	6.6	24.6	0.61	69

* if B>25%, R outside [0.5; 2]

Results of the probit model and the number of untreated and treated farms off/on support are the same as for NN(1). Differences between treated and control group are no longer significant after matching (results of the balancing test/*pstest* command are not fully reported here). Pseudo R^2 has decreased to 0.011.

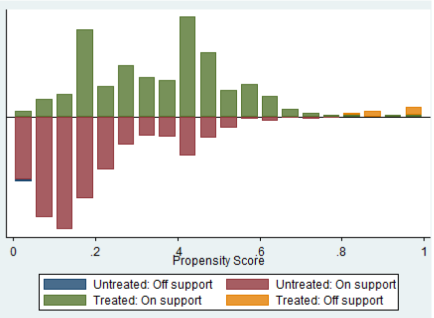
(6) PSM (kernel matching) for wine 2014; ATT= 1421.45

```
. psmatch2 T OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC MKM MKM2, outcome(FNI) kernel ate common
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	20774.2774	49089.2052	-28314.9279	10904.7962	-2.60
	ATT	27119.0637	25697.6142	1421.44948	9317.18466	0.15
	ATU	44876.0791	56640.146	11764.0668	.	.
	ATE			9300.73545	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	3	934	937
Treated	8	292	300
Total	11	1,226	1,237



Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.161	220.45	0.000	15.3	11.4	72.1*	0.46*	69
Matched	0.010	8.32	0.939	4.3	2.9	24.1	0.72	62

* if B>25%, R outside [0.5; 2]

Results of the probit model are the same as for NN(1). Again, differences between treated and control group are no longer significant after matching (results of the *pstest* command are not fully reported here). Pseudo R^2 has decreased to 0.010.

[11] 5-nearest-neighbours matching

Olives specialists 2014

(1) PaRCI model for olives 2014; ATE= -2773.72

```
. reg FNI T OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC
```

Source	SS	df	MS
Model	2.6172e+12	15	1.7448e+11
Residual	4.2295e+11	1020	414654127
Total	3.0401e+12	1035	2.9373e+09

Number of obs =	1036
F(15, 1020) =	420.78
Prob > F =	0.0000
R-squared =	0.8609
Adj R-squared =	0.8588
Root MSE =	20363

FNI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
T	-2773.724	2853.444	-0.97	0.331	-8373.016	2825.569
OUT	1.053151	.018749	56.17	0.000	1.01636	1.089942
SPC	-.709009	.0960284	-7.38	0.000	-.8974449	-.5205731
OVER	-1.626404	.0740393	-21.97	0.000	-1.771691	-1.481117
UL	2888.91	1421.737	2.03	0.042	99.04616	5678.774
PL	12813.36	1420.667	9.02	0.000	10025.59	15601.12
UAA	299.7262	33.43773	8.96	0.000	234.1116	365.3408
LIA	-.0267325	.0115895	-2.31	0.021	-.0494745	-.0039905
EXT	-1.177048	.0692049	-17.01	0.000	-1.312848	-1.041248
LFA	-1366.502	1744.575	-0.78	0.434	-4789.868	2056.863
MA	2619.695	1628.21	1.61	0.108	-575.3294	5814.719
ORG	913.402	1894.057	0.48	0.630	-2803.291	4630.095
MACH	-.0530447	.0235605	-2.25	0.025	-.0992772	-.0068122
GDPC	-385.6279	169.0751	-2.28	0.023	-717.4028	-53.85305
FHC	-1.52988	.7542044	-2.03	0.043	-3.009849	-.0499104
_cons	4331.405	3664.313	1.18	0.237	-2859.048	11521.86

(2) PaRVI model for olives 2014; ATT= -11899.45 + 9146.39= -2753.06

```
. reg FNI T OUT OUT_T SPC SPC_T OVER OVER_T UL UL_T PL PL_T UAA UAA_T LIA LIA_T EXT EXT_T LFA LFA_T MA
> MA_T ORG ORG_T MACH MACH_T GDPC GDPC_T FHC FHC_T
```

Source	SS	df	MS
Model	2.6242e+12	29	9.0488e+10
Residual	4.1598e+11	1006	413494265
Total	3.0401e+12	1035	2.9373e+09

Number of obs =	1036
F(29, 1006) =	218.84
Prob > F =	0.0000
R-squared =	0.8632
Adj R-squared =	0.8592
Root MSE =	20335

FNI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
T	-11899.45	27172.24	-0.44	0.662	-65220.21	41421.31
OUT	1.047319	.0193439	54.14	0.000	1.00936	1.085278
OUT_T	.1138724	.1321325	0.86	0.389	-.1454145	.3731593
SPC	-.6977267	.0969915	-7.19	0.000	-.8880556	-.5073978
SPC_T	.0179402	1.08143	0.02	0.987	-2.104177	2.140057
OVER	-1.646505	.074332	-22.15	0.000	-1.792368	-1.500641
OVER_T	.8926233	1.003665	0.89	0.374	-1.076894	2.862141
UL	2990.05	1463.166	2.04	0.041	118.8431	5861.257
UL_T	-1292.967	8020.323	-0.16	0.872	-17031.45	14445.51
PL	12769.08	1430.045	8.93	0.000	9962.865	15575.29
PL_T	-30161.04	18196.85	-1.66	0.098	-65869.17	5547.086
UAA	319.4817	34.31024	9.31	0.000	252.1539	386.8096
UAA_T	-331.6224	168.6757	-1.97	0.050	-662.619	-.6258239

LIA	-.0255729	.011651	-2.19	0.028	-.0484359	-.0027099
LIA_T	-.1209253	.3749353	-0.32	0.747	-.8566702	.6148197
EXT	-1.169413	.0698148	-16.75	0.000	-1.306412	-1.032413
EXT_T	1.076367	1.0814	1.00	0.320	-1.045691	3.198425
LFA	-1361.971	1814.99	-0.75	0.453	-4923.571	2199.63
LFA_T	878.0346	8495.109	0.10	0.918	-15792.13	17548.2
MA	2710.491	1658.6	1.63	0.103	-544.222	5965.203
MA_T	-720.1669	9441.353	-0.08	0.939	-19247.17	17806.83
ORG	896.5287	1973.387	0.45	0.650	-2975.898	4768.955
ORG_T	3397.879	10215.94	0.33	0.740	-16649.12	23444.88
MACH	-.055291	.0237695	-2.33	0.020	-.1019344	-.0086475
MACH_T	-.0914924	.201172	-0.45	0.649	-.4862572	.3032725
GDPC	-431.0066	171.4976	-2.51	0.012	-767.5405	-94.47264
GDPC_T	836.7625	1449.278	0.58	0.564	-2007.191	3680.716
FHC	-1.448104	.7586245	-1.91	0.057	-2.936772	.0405638
FHC_T	-1.360658	7.188048	-0.19	0.850	-15.46594	12.74463
_cons	4847.735	3718.094	1.30	0.193	-2448.373	12143.84

(3) PSM (NN(1) matching) for olives 2014; ATT= 12169.77

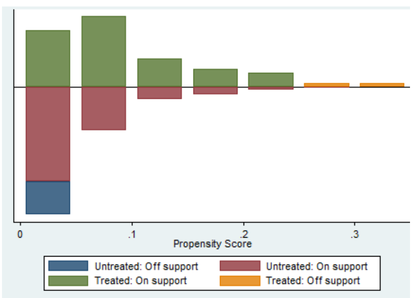
```
. psmatch2 T OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC MKM MKM2, outcome(FNI) neighbor (1) ate common

Probit regression                               Number of obs   =       1036
                                                LR chi2(16)     =      39.34
                                                Prob > chi2     =     0.0010
Log likelihood = -195.31057                     Pseudo R2       =     0.0915
```

T	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
OUT	5.48e-06	2.30e-06	2.38	0.017	9.71e-07	9.98e-06
SPC	-.0000176	.0000142	-1.24	0.217	-.0000454	.0000103
OVER	-8.93e-06	.0000126	-0.71	0.480	-.0000337	.0000158
UL	.2008846	.1467503	1.37	0.171	-.0867407	.4885099
PL	.0322036	.2281213	0.14	0.888	-.414906	.4793132
UAA	.0026774	.00412	0.65	0.516	-.0053976	.0107524
LIA	7.41e-07	1.86e-06	0.40	0.691	-2.91e-06	4.39e-06
EXT	-9.52e-06	.000013	-0.73	0.463	-.0000349	.0000159
LFA	.3489431	.1806091	1.93	0.053	-.0050443	.7029305
MA	-.1374657	.1975047	-0.70	0.486	-.5245678	.2496364
ORG	-.0300747	.1996781	-0.15	0.880	-.4214366	.3612872
MACH	-4.69e-06	3.47e-06	-1.35	0.176	-.0000115	2.10e-06
GDPC	-.0072222	.0259543	-0.28	0.781	-.0580918	.0436473
FHC	-.0003953	.0001443	-2.74	0.006	-.0006781	-.0001124
MKM	-.0381467	.0117806	-3.24	0.001	-.0612362	-.0150573
MKM2	.0003162	.0001418	2.23	0.026	.0000383	.000594
_cons	-.8976359	.5045608	-1.78	0.075	-1.886557	.0912852

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	27563.6773	23377.0882	4186.58909	7512.51231	0.56
	ATT	28302.8576	16133.0897	12169.7678	7161.49988	1.70
	ATU	22400.4457	19497.7119	-2902.73375	.	.
	ATE			-1978.14743	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	170	811	981
Treated	2	53	55
Total	172	864	1,036



```
. pstest OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC MKM MKM2, both treatment (T)
```

Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
OUT	U	39750	39081	0.9		0.05	0.957	0.30*
	M	40990	23176	24.2	-2563.0	2.31	0.023	4.12*
SPC	U	5252.2	7007.3	-12.3		-0.66	0.508	0.07*
	M	5437.4	3027	16.8	-37.3	2.79	0.006	1.76*
OVER	U	6430.6	7941.7	-9.5		-0.52	0.603	0.09*
	M	6609.2	4936.6	10.6	-10.7	1.59	0.114	2.21*
UL	U	.95908	.87251	19.3		1.37	0.172	0.91
	M	.95206	.93481	3.8	80.1	0.20	0.845	0.92
PL	U	.48653	.58124	-9.9		-0.56	0.577	0.17*
	M	.50489	.27197	24.3	-145.9	2.68	0.009	2.06*
UAA	U	21.04	22.354	-4.8		-0.31	0.760	0.49*
	M	21.566	16.128	20.0	-313.7	1.26	0.212	0.98
LIA	U	2372.9	5652.9	-3.5		-0.18	0.853	0.01*
	M	2142.4	169.98	2.1	39.9	1.44	0.153	436.98*
EXT	U	7516.5	10362	-13.9		-0.75	0.452	0.07*
	M	7784.3	4355	16.8	-20.5	2.64	0.010	1.45
LFA	U	.50909	.33843	34.9		2.59	0.010	.
	M	.49057	.5283	-7.7	77.9	-0.39	0.701	.
MA	U	.29091	.39348	-21.6		-1.52	0.129	.
	M	.30189	.28302	4.0	81.6	0.21	0.833	.
ORG	U	.14545	.16208	-4.6		-0.33	0.745	.
	M	.13208	.13208	0.0	100.0	-0.00	1.000	.
MACH	U	14312	21422	-29.3		-1.75	0.080	0.32*
	M	14852	11211	15.0	48.8	1.09	0.278	0.95
GDPC	U	17.915	18.465	-15.6		-1.03	0.304	0.62
	M	18.006	18.966	-27.3	-74.5	-1.32	0.190	0.52*
FHC	U	396.07	658.69	-36.8		-2.14	0.033	0.25*
	M	399.89	302.74	13.6	63.0	1.13	0.261	1.15
MKM	U	19.255	22.419	-24.2		-1.87	0.062	1.33
	M	19.981	21.698	-13.1	45.7	-0.61	0.542	0.82
MKM2	U	562.27	649.29	-10.3		-0.60	0.551	0.22*
	M	583.49	696.42	-13.4	-29.8	-0.57	0.569	0.14*

* if variance ratio outside [0.58; 1.71] for U and [0.58; 1.73] for M

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.092	39.34	0.001	15.7	13.1	86.3*	0.57	77
Matched	0.148	21.75	0.152	13.3	13.5	48.6*	25.35*	54

* if B>25%, R outside [0.5; 2]

Many farms are excluded from the estimation because no matches were found. Even the estimated propensity scores of adopters are rather low. The pseudo R^2 is low and even decreased after matching, which indicates that differences between treated and control farms are even larger after matching. The balancing test also reports that the two groups significantly differ with respect to OUT, SPC, PL and EXT, although the differences were not significant before matching. The results suggest that PSM is not a good choice to estimate income effects of GIs for olives specialists.

(4) PSM (NN(5) matching) for olives 2014; ATT= 2304.47

```
. psmatch2 T OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC MKM MKM2, outcome(FNI) neighbor (5) ate common
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	27563.6773	23377.0882	4186.58909	7512.51231	0.56
	ATT	28302.8576	25998.3859	2304.47163	7322.05672	0.31
	ATU	22400.4457	20844.3037	-1556.14198	.	.
	ATE			-1319.32194	.	.

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.092	39.34	0.001	15.7	13.1	86.3*	0.57	77
Matched	0.043	6.32	0.984	8.3	6.3	46.2*	1.32	46

* if B>25%, R outside [0.5; 2]

Results of the probit model and the numbers of treated and untreated farms off/on support is the same as for PSM NN(1). However, the balancing test indicates that matches are better in terms of reduced differences between treated and control group. There are no significant differences with respect to the observed variables after matching (results of the balancing test/*pstest* command are not fully reported here). The pseudo R^2 has also decreased from 0.092 to 0.043. However, the estimated propensity scores are very low for both adopters and non-adopters.

(5) PSM (kernel matching) for olives 2014; ATT= 4895.76

```
. psmatch2 T OUT SPC OVER UL PL UAA LIA EXT LFA MA ORG MACH GDPC FHC MKM MKM2, outcome(FNI) kernel ate common
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	27563.6773	23377.0882	4186.58909	7512.51231	0.56
	ATT	28302.8576	23407.099	4895.75853	6672.12745	0.73
	ATU	22400.4457	24422.1285	2021.68284	.	.
	ATE			2197.9861	.	.

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.092	39.34	0.001	15.7	13.1	86.3*	0.57	77
Matched	0.017	2.44	1.000	4.3	3.9	30.9*	0.60	31

* if B>25%, R outside [0.5; 2]

Results of the probit model and the numbers of treated and untreated farms off/on support is the same as for PSM NN(1). Similar to the NN(5) matching techniques, differences between treated and control group are no longer significant after matching (results of the balancing test/*pstest* command are not fully reported here). The pseudo R^2 decreased to 0.017. Again, many farms are excluded from the estimation because of missing matches. Propensity scores remain low for adopters and non-adopters.

Quality wine specialists 2015

(1) PaRCI model for wine 2015; ATE= 1238.21

```
. reg FNI T OUT SPC OVER UL PL UAA LIA EXT ORG MACH GDPC FHC
```

Source	SS	df	MS
Model	4.0173e+13	13	3.0902e+12
Residual	9.4756e+11	1206	785704078
Total	4.1121e+13	1219	3.3733e+10

Number of obs =	1220
F(13, 1206) =	3933.09
Prob > F =	0.0000
R-squared =	0.9770
Adj R-squared =	0.9767
Root MSE =	28030

FNI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
T	1238.205	1932.453	0.64	0.522	-2553.137	5029.548
OUT	1.027491	.0057665	178.18	0.000	1.016177	1.038804
SPC	-1.126902	.0155851	-72.31	0.000	-1.157479	-1.096325
OVER	-1.210258	.057918	-20.90	0.000	-1.32389	-1.096627
UL	-2435.633	1345.827	-1.81	0.071	-5076.055	204.7886
PL	-625.6613	301.4285	-2.08	0.038	-1217.044	-34.27878
UAA	97.68873	30.82022	3.17	0.002	37.22152	158.1559
LIA	-.0145225	.0062091	-2.34	0.020	-.0267043	-.0023406
EXT	-1.221819	.0394512	-30.97	0.000	-1.29922	-1.144418
ORG	2593.802	5678.72	0.46	0.648	-8547.465	13735.07
MACH	-.099028	.0151352	-6.54	0.000	-.1287222	-.0693337
GDPC	-202.7701	142.5251	-1.42	0.155	-482.3948	76.85459
FHC	4.473843	1.337263	3.35	0.001	1.850223	7.097463
_cons	8814.238	3699.404	2.38	0.017	1556.255	16072.22

(2) PaRVI model for wine 2015; ATT= 2612.15 + 425.90= 3038.05

```
. reg FNI T OUT OUT_T SPC SPC_T OVER OVER_T UL UL_T PL PL_T UAA UAA_T LIA LIA_T EXT EXT_T ORG ORG_T
> MACH MACH_T GDPC GDPC_T FHC FHC_T
```

Source	SS	df	MS
Model	4.0234e+13	25	1.6093e+12
Residual	8.8707e+11	1194	742936453
Total	4.1121e+13	1219	3.3733e+10

Number of obs =	1220
F(25, 1194) =	2166.20
Prob > F =	0.0000
R-squared =	0.9784
Adj R-squared =	0.9780
Root MSE =	27257

FNI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
T	2612.147	7594.778	0.34	0.731	-12288.45	17512.74
OUT	1.032977	.0062373	165.61	0.000	1.02074	1.045214
OUT_T	-.1631748	.0451866	-3.61	0.000	-.2518287	-.0745209
SPC	-1.135552	.0164311	-69.11	0.000	-1.167789	-1.103315
SPC_T	.2307152	.0757271	3.05	0.002	.0821421	.3792882
OVER	-1.054928	.0638422	-16.52	0.000	-1.180183	-.9296723
OVER_T	-.3623273	.1703611	-2.13	0.034	-.6965678	-.0280869
UL	-2824.582	1544.633	-1.83	0.068	-5855.08	205.9152
UL_T	2860.857	2978.874	0.96	0.337	-2983.552	8705.267
PL	569.8623	453.8097	1.26	0.209	-320.4909	1460.216
PL_T	-2238.341	652.4124	-3.43	0.001	-3518.343	-958.3385
UAA	203.0452	36.65867	5.54	0.000	131.1227	274.9678
UAA_T	-231.7278	77.20827	-3.00	0.003	-383.2068	-80.24883
LIA	-.0079871	.0066792	-1.20	0.232	-.0210913	.0051172
LIA_T	-.024307	.0172734	-1.41	0.160	-.0581966	.0095826
EXT	-1.389794	.0452481	-30.71	0.000	-1.478568	-1.301019
EXT_T	.5722483	.0929851	6.15	0.000	.3898159	.7546806
ORG	1743.402	6712.772	0.26	0.795	-11426.74	14913.54
ORG_T	694.6263	11957.34	0.06	0.954	-22765.12	24154.37
MACH	-.1818059	.0207539	-8.76	0.000	-.2225241	-.1410878
MACH_T	.1912218	.0325552	5.87	0.000	.12735	.2550936
GDPC	3.061913	166.6112	0.02	0.985	-323.8213	329.9451
GDPC_T	210.5009	342.7059	0.61	0.539	-461.8719	882.8736
FHC	6.524794	1.541948	4.23	0.000	3.499565	9.550023
FHC_T	-7.14564	3.013481	-2.37	0.018	-13.05795	-1.233333
_cons	2216.496	4413.539	0.50	0.616	-6442.659	10875.65

(3) PSM (NN(1) matching) for wine 2015; ATT= 3878.33

```
. psmatch2 T OUT SPC OVER UL PL UAA LIA EXT ORG MACH GDPC FHC MKM MKM2, outcome(FNI) neighbor (1) ate
> common
```

Probit regression	Number of obs	=	1220
	LR chi2(14)	=	187.71
	Prob > chi2	=	0.0000
Log likelihood = -595.47214	Pseudo R2	=	0.1362

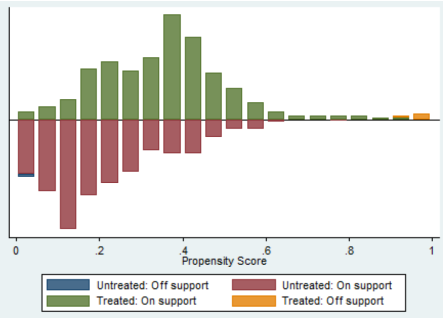
T	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
OUT	-3.09e-06	8.97e-07	-3.44	0.001	-4.85e-06	-1.33e-06
SPC	-1.87e-06	9.15e-07	-2.04	0.041	-3.67e-06	-7.66e-08
OVER	2.36e-06	3.29e-06	0.72	0.473	-4.09e-06	8.81e-06
UL	.1709615	.0739547	2.31	0.021	.0260128	.3159101
PL	.0353836	.0164037	2.16	0.031	.0032329	.0675342
UAA	-.0114719	.0019316	-5.94	0.000	-.0152578	-.0076859
LIA	2.23e-07	3.50e-07	0.64	0.525	-4.64e-07	9.10e-07
EXT	.0000112	3.06e-06	3.67	0.000	5.24e-06	.0000172
ORG	.0647144	.2760306	0.23	0.815	-.4762956	.6057243
MACH	1.72e-06	8.80e-07	1.95	0.051	-4.37e-09	3.45e-06
GDPC	-.072334	.0086337	-8.38	0.000	-.0892558	-.0554122
FHC	-.0000347	.0000726	-0.48	0.633	-.000177	.0001077
MKM	.0034435	.0080909	0.43	0.670	-.0124145	.0193014
MKM2	.0002709	.0001038	2.61	0.009	.0000674	.0004743
_cons	.5974977	.2073066	2.88	0.004	.1911842	1.003811

Note: 2 failures and 0 successes completely determined.

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	24748.8613	48186.7093	-23437.848	12090.5161	-1.94
	ATT	25769.3661	21891.0385	3878.32761	6450.27239	0.60
	ATU	37923.8256	47363.8095	9439.98396	.	.
	ATE			8043.8291	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	5	907	912
Treated	4	304	308
Total	9	1,211	1,220




```
. pstest OUT SPC OVER UL PL UAA LIA EXT ORG MACH GDPC FHC MKM MKM2, both treatment (t)
```

Variable	Unmatched Matched	Mean Treated	Control	%bias	%reduct |bias|	t-test t	p>|t|	V(T)/ V(C)
OUT	U	76196	96493	-9.0		-1.19	0.233	0.22*
	M	68080	68224	-0.1	99.3	-0.02	0.987	0.66*
SPC	U	12936	16268	-4.7		-0.61	0.542	0.14*
	M	11265	11138	0.2	96.2	0.06	0.954	1.05
OVER	U	13252	12830	1.2		0.17	0.862	0.57*
	M	11351	9707	4.8	-289.7	1.12	0.261	1.45*
UL	U	1.0564	1.0819	-3.8		-0.59	0.554	1.15
	M	1.0681	1.0181	7.6	-97.1	0.93	0.353	1.11
PL	U	1.9474	.86793	20.2		3.63	0.000	3.81*
	M	1.4808	1.5377	-1.1	94.7	-0.13	0.900	0.75*

UAA	U	18.357	23.367	-13.7		-2.03	0.042	0.81
	M	17.397	23.962	-18.0	-31.0	-2.38	0.018	0.96
LIA	U	31137	19791	7.8		1.14	0.255	0.70*
	M	28981	26983	1.4	82.4	0.16	0.874	0.54*
EXT	U	19388	15433	6.8		1.00	0.316	0.78*
	M	14979	14904	0.1	98.1	0.03	0.979	0.67*
ORG	U	.02597	.01864	5.0		0.79	0.433	.
	M	.02632	.04934	-15.6	-214.0	-1.49	0.137	.
MACH	U	37502	20746	19.2		3.45	0.001	3.68*
	M	31289	40627	-10.7	44.3	-1.08	0.280	0.44*
GDPC	U	20.838	23.787	-48.3		-7.38	0.000	1.06
	M	20.946	20.426	8.5	82.3	1.13	0.259	1.42*
FHC	U	230.45	213.31	2.8		0.43	0.668	1.17
	M	233.48	179.51	8.7	-214.8	1.12	0.263	1.45*
MKM	U	30.195	26.425	23.0		3.76	0.000	1.78*
	M	30.474	32.569	-12.8	44.4	-1.44	0.151	1.12
MKM2	U	1254	891.52	28.3		5.01	0.000	3.35*
	M	1269.1	1363.9	-7.4	73.9	-0.83	0.404	1.86*

* if variance ratio outside [0.80; 1.25] for U and [0.80; 1.25] for M

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.136	187.71	0.000	13.8	8.4	79.4*	0.72	69
Matched	0.042	35.35	0.001	6.9	7.5	48.6*	0.63	69

* if B>25%, R outside [0.5; 2]

Differences between the treated and the control group with respect to PL, MACH, GDPC, MKM and MKM2 are no longer significant after matching. The difference in UAA remains significant. The pseudo R^2 decreased from 0.136 to 0.042.

(4) PSM (NN(5) matching) for wine 2015; ATT=1293.24

```
. psmatch2 T OUT SPC OVER UL PL UAA LIA EXT ORG MACH GDPC FHC MKM MKM2, outcome(FNI) neighbor (5) ate
> common
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	24748.8613	48186.7093	-23437.848	12090.5161	-1.94
	ATT	25769.3661	24476.1311	1293.235	8552.42454	0.15
	ATU	37923.8256	44170.2264	6246.40081	.	.
	ATE			5002.99668	.	.

Results of the probit model and the number of untreated and treated farms off/on support are the same as for NN(1). All differences between treated and control group are no longer significant after matching (results of the balancing test/*pstest* command are not fully reported here). Pseudo R^2 has decreased to 0.026.

(5) PSM (kernel matching) for wine 2015; ATT= 2702.60

```
. psmatch2 T OUT SPC OVER UL PL UAA LIA EXT ORG MACH GDPC FHC MKM MKM2, outcome(FNI) kernel ate common
```

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	24748.8613	48186.7093	-23437.848	12090.5161	-1.94
	ATT	25769.3661	23066.7676	2702.59857	6415.21555	0.42
	ATU	37923.8256	40852.7162	2928.89065	.	.
	ATE			2872.08405	.	.

Note: S.E. does not take into account that the propensity score is estimated.

Results of the probit model and the number of untreated and treated farms off/on support are the same as for NN(1). Again, all differences between treated and control group are no longer significant after matching (results of the *pstest* command are not fully reported here). Pseudo R^2 has decreased to 0.019.

Olives specialists 2015

(1) PaRCI model for olives 2015; ATE= -3495.41

```
. reg FNI T OUT SPC OVER UL PL UAA LIA EXT ORG MACH GDPC FHC
```

Source	SS	df	MS
Model	4.1200e+12	13	3.1692e+11
Residual	4.2524e+11	978	434809100
Total	4.5452e+12	991	4.5865e+09

Number of obs =	992
F(13, 978) =	728.87
Prob > F =	0.0000
R-squared =	0.9064
Adj R-squared =	0.9052
Root MSE =	20852

FNI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
T	-3495.41	2844.116	-1.23	0.219	-9076.682	2085.862
OUT	1.055251	.0162711	64.85	0.000	1.023321	1.087181
SPC	-1.656224	.0951296	-17.41	0.000	-1.842906	-1.469543
OVER	-.892892	.0401095	-22.26	0.000	-.9716026	-.8141814
UL	1700.089	1471.897	1.16	0.248	-1188.35	4588.528
PL	1967.948	1906.177	1.03	0.302	-1772.72	5708.616
UAA	284.2531	31.49237	9.03	0.000	222.4527	346.0535
LIA	-.0590768	.0093667	-6.31	0.000	-.0774579	-.0406957
EXT	-.8200094	.1038258	-7.90	0.000	-1.023756	-.6162625
ORG	-1662.42	2156.554	-0.77	0.441	-5894.426	2569.585
MACH	-.064152	.0249788	-2.57	0.010	-.1131701	-.0151338
GDPC	-878.3093	214.9697	-4.09	0.000	-1300.164	-456.4544
FHC	-1.196096	.943902	-1.27	0.205	-3.048402	.6562102
_cons	17032.38	4077.543	4.18	0.000	9030.636	25034.12

(2) PaRVI model for olives 2015; ATT= -42172.75 + 38696.07= -3476.68

```
. reg FNI T OUT OUT_T SPC SPC_T OVER OVER_T UL UL_T PL PL_T UAA UAA_T LIA LIA_T EXT EXT_T ORG ORG_T
> MACH MACH_T GDPC GDPC_T FHC FHC_T
```

Source	SS	df	MS
Model	4.1241e+12	25	1.6496e+11
Residual	4.2110e+11	966	435919214
Total	4.5452e+12	991	4.5865e+09

Number of obs =	992
F(25, 966) =	378.43
Prob > F =	0.0000
R-squared =	0.9074
Adj R-squared =	0.9050
Root MSE =	20879

FNI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
T	-42172.75	27692.01	-1.52	0.128	-96516.19	12170.68
OUT	1.048948	.0171992	60.99	0.000	1.015196	1.0827
OUT_T	.1051681	.0840659	1.25	0.211	-.0598048	.270141
SPC	-1.657474	.0960712	-17.25	0.000	-1.846007	-1.468942
SPC_T	.8569209	1.648373	0.52	0.603	-2.377883	4.091725
OVER	-.8906705	.0406944	-21.89	0.000	-.9705302	-.8108109
OVER_T	-.4000975	.97598	-0.41	0.682	-2.315383	1.515188
UL	1944.888	1502.147	1.29	0.196	-1002.961	4892.736
UL_T	1731.897	9295.198	0.19	0.852	-16509.21	19973.01
PL	2196.178	1923.203	1.14	0.254	-1577.959	5970.315
PL_T	-20978.13	21356.78	-0.98	0.326	-62889.16	20932.89
UAA	286.6127	31.8041	9.01	0.000	224.1996	349.0258
UAA_T	-214.0309	293.1412	-0.73	0.465	-789.2979	361.2362
LIA	-.0570886	.0094392	-6.05	0.000	-.0756122	-.038565
LIA_T	-.0747457	.4658658	-0.16	0.873	-.9889715	.83948
EXT	-.8176852	.1048039	-7.80	0.000	-1.023355	-.6120156
EXT_T	.9202236	1.433244	0.64	0.521	-1.892406	3.732853
ORG	-1738.626	2194.131	-0.79	0.428	-6044.438	2567.185
ORG_T	166.8369	17279.5	0.01	0.992	-33742.84	34076.51
MACH	-.0645329	.0254756	-2.53	0.011	-.1145268	-.014539
MACH_T	-.1440318	.1616848	-0.89	0.373	-.4613258	.1732622
GDPC	-963.1028	221.0075	-4.36	0.000	-1396.813	-529.3927
GDPC_T	2293.524	1566.896	1.46	0.144	-781.3889	5368.437
FHC	-1.098043	.9742487	-1.13	0.260	-3.009931	.8138451
FHC_T	1.421922	4.578089	0.31	0.756	-7.562224	10.40607
_cons	18362.55	4175.365	4.40	0.000	10168.72	26556.38

(3) PSM (NN(1) matching) for olives 2015; ATT= 10810.92

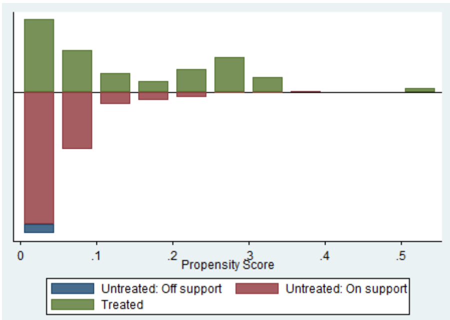
```
. psmatch2 T OUT SPC OVER UL PL UAA LIA EXT ORG MACH GDPC FHC MKM MKM2, outcome(FNI) neighbor (1) a
> te common
```

Probit regression			
	Number of obs	=	992
	LR chi2(14)	=	57.63
	Prob > chi2	=	0.0000
Log likelihood = -192.13375	Pseudo R2	=	0.1304

T	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
OUT	6.24e-06	1.69e-06	3.70	0.000	2.94e-06	9.54e-06
SPC	-.0000716	.0000244	-2.94	0.003	-.0001194	-.0000238
OVER	-3.61e-06	4.74e-06	-0.76	0.446	-.0000129	5.68e-06
UL	.0287842	.1625054	0.18	0.859	-.2897206	.347289
PL	.0023529	.2894692	0.01	0.994	-.5649964	.5697022
UAA	.0062824	.0039503	1.59	0.112	-.0014599	.0140248
LIA	3.67e-06	3.56e-06	1.03	0.302	-3.30e-06	.0000106
EXT	-.0000153	.0000156	-0.98	0.326	-.0000458	.0000152
ORG	-.3287798	.2746688	-1.20	0.231	-.8671207	.2095611
MACH	-3.00e-06	3.30e-06	-0.91	0.362	-9.46e-06	3.46e-06
GDPC	.0240674	.0276315	0.87	0.384	-.0300893	.078224
FHC	-.000219	.0001218	-1.80	0.072	-.0004576	.0000197
MKM	-.0581447	.0123217	-4.72	0.000	-.0822947	-.0339947
MKM2	.0004922	.0001498	3.28	0.001	.0001985	.0007859
_cons	-.8827109	.512222	-1.72	0.085	-1.886648	.1212258

Note: 9 failures and 0 successes completely determined.

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	35255.3268	34428.1552	827.171593	9169.09278	0.09
	ATT	35255.3268	24444.4111	10810.9157	10294.981	1.05
	ATU	28021.7255	34621.002	6599.27654	.	.
	ATE			6855.868	.	.



psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	40	894	934
Treated	0	58	58
Total	40	952	992

```
. pstest OUT SPC OVER UL PL UAA LIA EXT ORG MACH GDPC FHC MKM MKM2, both treatment (T)
```

Variable	Unmatched Matched	Mean Treated	Control	%bias	%reduct |bias|	t-test t	p>|t|	V(T)/ V(C)
OUT	U	46282	55698	-9.3		-0.56	0.575	0.28*
	M	46282	38870	7.3	21.3	0.65	0.518	1.40
SPC	U	3302.8	7738.7	-31.4		-1.71	0.087	0.03*
	M	3302.8	3670.7	-2.6	91.7	-0.53	0.595	0.55*
OVER	U	7370.1	10828	-12.2		-0.67	0.506	0.03*
	M	7370.1	8786.3	-5.0	59.0	-0.35	0.728	0.06*
UL	U	.92571	.91116	3.5		0.23	0.818	0.57*
	M	.92571	.94587	-4.8	-38.6	-0.31	0.757	1.08

PL	U	.40527	.77303	-25.1		-1.41	0.160	0.09*
	M	.40527	.34759	3.9	84.3	0.55	0.580	1.23
UAA	U	18.774	25.9	-18.9		-1.07	0.284	0.12*
	M	18.774	18.58	0.5	97.3	0.06	0.956	0.76
LIA	U	1411.2	6370.4	-5.2		-0.28	0.781	0.00*
	M	1411.2	658.29	0.8	84.8	0.66	0.513	4.32*
EXT	U	6731.5	13223	-23.9		-1.32	0.186	0.07*
	M	6731.5	5244.9	5.5	77.1	0.92	0.357	1.52
ORG	U	.05172	.13276	-28.2		-1.79	0.073	.
	M	.05172	.03448	6.0	78.7	0.45	0.651	.
MACH	U	16002	24604	-31.2		-1.95	0.051	0.39*
	M	16002	15699	1.1	96.5	0.09	0.928	1.81*
GDPC	U	17.59	17.953	-11.3		-0.82	0.414	0.89
	M	17.59	18.84	-39.0	-244.4	-1.88	0.063	0.61
FHC	U	512.59	451.51	8.4		0.63	0.531	1.04
	M	512.59	597.02	-11.6	-38.2	-0.65	0.515	1.26
MKM	U	15.879	22.089	-47.5		-4.08	0.000	1.83*
	M	15.879	16.638	-5.8	87.8	-0.24	0.811	0.62
MKM2	U	469.84	608.54	-18.7		-1.15	0.249	0.34*
	M	469.84	627.16	-21.2	-13.4	-0.66	0.514	0.09*

* if variance ratio outside [0.59; 1.69] for U and [0.59; 1.69] for M

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.130	57.63	0.000	19.6	18.8	74.7*	0.33*	85
Matched	0.078	12.48	0.567	8.2	5.2	67.0*	0.72	38

* if B>25%, R outside [0.5; 2]

Similar to the sample of olives specialists in 2014, the estimated propensity scores are very low, even for adopters. 40 untreated farms are off support. The pseudo R^2 decreased from 0.130 to 0.078 after matching. Before, differences between treated and control group were significant (at 10% level or lower) for SPC, ORG, MACH and MKM. The balancing test shows that the differences are no longer significant after matching.

(4) PSM (NN(5) matching) for olives 2015; ATT= 8723.52

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	35255.3268	34428.1552	827.171593	9169.09278	0.09
	ATT	35255.3268	26531.8064	8723.52042	11591.4079	0.75
	ATU	28021.7255	36286.7641	8265.03862	.	.
	ATE			8292.97133	.	.

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.130	57.63	0.000	19.6	18.8	74.7*	0.33*	85
Matched	0.041	6.62	0.948	4.1	2.2	48.0*	0.59	54

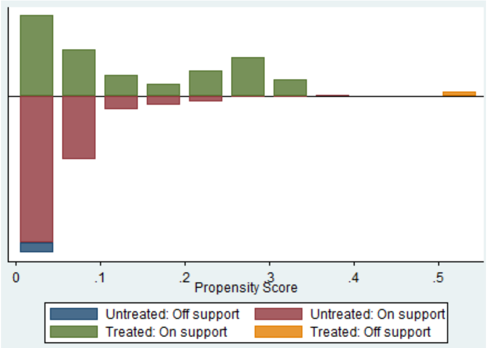
Results of the probit estimation and the numbers of treated and untreated farms off/on support is the same as before. Again, differences between treated and control group are no longer significant after matching. The pseudo R^2 has further decreased to 0.041. Overall, the estimated propensity scores stay low.

(6) PSM (kernel matching) for olives 2015; ATT= -5499.85

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
FNI	Unmatched	35255.3268	34428.1552	827.171593	9169.09278	0.09
	ATT	28494.7212	33994.5703	-5499.84906	6640.48911	-0.83
	ATU	28021.7255	32397.4247	4375.6992	.	.
	ATE			3783.78936	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	40	894	934
Treated	1	57	58
Total	41	951	992



Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.130	57.63	0.000	19.6	18.8	74.7*	0.33*	85
Matched	0.034	5.41	0.979	7.1	6.4	43.6*	0.49*	62

* if B>25%, R outside [0.5; 2]

One adopter is excluded from the estimation when using the kernel matching technique. The propensity scores remain low for both treated and untreated farms. However, the pseudo R^2 further decreased to 0.034, which indicates that matching successfully reduced differences between the control group and the treated group.