International Food and Agribusiness Management Review
Volume 26, Issue 3, 2023; DOI: 10.22434/IFAMR2022.0086

Received: 31 July 2022 / Accepted: 6 December 2022

Special Issue: Agri-food systems transformation

Participation in a mutual fund covering losses due to pest infestation: analyzing key predictors of farmers' interest through machine learning

RESEARCH ARTICLE

Lisa Höschle[a], Samuele Trestini[b] and Elisa Giampietri[c]

[a]*PhD candidate at the Department of Agricultural Economics and Rural Development, University of Göttingen, Platz der Göttinger Sieben 5, 37073 Göttingen, Germany*

[b]*Associate Professor,* [c]*Research Associate, Department of Land, Environment, Agriculture and Forestry, University of Padova, Viale dell'Università 16, 35020 Legnaro, Italy*

Abstract

In the context of intensified *Halyomorpha halys* infestations in Italy, this paper provides a very first investigation of key factors that drive fruit growers' intention to participate in a mutual fund (MF) compensating production losses due to this invasive insect. Data were collected in Veneto Region in Italy, where many farmers suffered *H. halys* attacks, and interest in the development of innovative risk management tools is growing. The study investigates how behavioral (risk attitude, risk perception) and personality factors (self-efficacy, locus of control) explain farmers' intention to participate in the MF, additionally controlling for a large number of primary control data (e.g. farmers' perceptions and characteristics, farm characteristics). The study assumes approximate sparsity and applies the least absolute shrinkage and selection operator (LASSO), a machine learning technique which represents an original approach for research on risk management. Our empirical analysis reveals that farmers' intention to participate in the MF is driven by an interplay between the perceived risk of production loss, the benefits from participation in the fund, and the farm age, rather than by socio-economic characteristics of the farm. Results provide valuable insights for policymakers and local stakeholders to implement a mutual fund close to the farmers' needs.

Keywords: risk management, pest infestation, mutual funds, machine learning, LASSO
JEL-codes: D91, G32, Q10, Q18

Corresponding author: elisa.giampietri@unipd.it

1. Introduction

Over the last years, Italian farmers have suffered from increased severity and frequency of plant pests and diseases. This has put production at risk and compromised their economic stability and competitiveness (Möhring *et al.*, 2019; Zhang and Yu, 2021). The observed trend can be attributed to adverse effects of climate change and intensified international trade, whereas the latter is the main cause for the recent spread of *Halyomorpha halys* (*H. halys* hereafter) in the South of Europe. *H. halys* is an invasive insect – also known as brown marmorated stink bug – that originates from Asia (Haye *et al.*, 2015) and predominantly infects fruit trees and vegetable plants. The insect causes severe damages, ultimately resulting in substantial yield losses for farmers (Lee *et al.*, 2015). In case of infestation, the damage prevents both product consumption and processing. Recently, *H. halys* infestation has developed into a sanitary emergency for many farmers in Italy, particularly in the Northern regions[1]. Sanitary emergencies represent urgent situations in the context of plant health that significantly threaten production outcomes and economic profitability. The economic damage in the affected regions is estimated to be around 500 million € – only in 2019 (EC, 2020). Adequate solutions to foster farmers resilience against future attacks are still lacking (Zapponi *et al.*, 2021).

To cope with the biotic risk Italian farmers are facing, in 2019 the European Commission released a recommended course of action to encourage the use of targeted measures (EC, 2020). The intervention of the EC signals the recognition of the ongoing sanitary emergency, which caused substantial economic losses associated with *H. halys* infestation for Italian farmers.

Limited protection tools exist to control and limit the spread of *H. halys*. Farmers predominantly rely on phytosanitary measures (i.e. insecticides), biological control tools such as the 'Samurai' wasp (Scaccini *et al.*, 2020), or physical tools, namely anti-insect nets or traps. The effectiveness of mentioned tools is not proven. In addition to traditional self-copying strategies, farmers in the European Union can rely on innovative risk management tools to foster resilience against augmented pest infestation risks, such as mutual funds (MF) (Meuwissen *et al.*, 2018). MFs are collective instruments supported by the EU Common Agricultural Policy (CAP) that are based on solidarity and mutuality (Trestini *et al.*, 2017a,b). As regulated within the Art. 38 of the EU Regulation No. 1305/2013 (EC, 2013), MFs compensate farmers for production losses caused by sanitary risks, such as animal or plant diseases, pest infestations, adverse climate events or environmental incidents. Farmers that share similar production risks build a voluntary alliance to create a financial reserve – through annual contribution of the members – to indemnify farmers that face production losses. To participate in the MF, each farmer contributes with a membership fee, which is subsidized by 70% of the total amount by public resources (EU Regulation No. 2017/2393; EC, 2017). In the event of a production loss beyond the minimum damage threshold as set by the above-mentioned EU legislation, the fund provides compensation for affected farmers, within the limit of its financial availability (up to 100% of the total loss). Currently, the minimum damage threshold is calculated as 20% of the farmer's average annual production over the previous three-year period or as 20% of the three-year average of production based on the previous five-year period (excluding highest and lowest entry).

Between 2014 and 2020, only three Member States of the EU considered expenditures for MFs development in their budget of Rural Development Programs, namely Italy (€ 97 million), France (€ 60 million), and Romania (€ 200 million) (Cordier and Santeramo, 2019). On the one hand, low implementation rates can be attributed to different levels of risk exposure between EU member states, resulting in diverging policy implications. On the other hand, countries are un-constrained in the budget allocation of risk-specific CAP funds; therefore, financial resources compete with alternative rural development instruments (Meuwissen *et al.*, 2018).

[1] Namely, Veneto, Emilia-Romagna, Trentino Alto-Adige, Lombardia, Piemonte and Friuli-Venezia-Giulia region.

In Italy, MFs are complementary to traditional agricultural insurances[2], as they cover risks that are currently not insured otherwise (e.g. plant diseases). As a consequence, the tool reveals strong potential to manage production risks related to *H. halys* attacks. One main advantage of MFs is their high share of public financial support that makes them a comparatively low-cost instrument for risk diversification. However, the persistence of structural constraints and critical hindrances limits the development and the operativity of the tool (Trestini and Giampietri, 2018). Throughout Italy, MFs are geographically concentrated in the North and managed by the local Defence Consortia[3]. A key barrier for the further development and expansion of MFs in Italy is the required minimum number of 700 farmers for the set-up of the instrument, implemented by the national legislation. Another hurdle that limits the spread of the instrument is the lack of knowledge about farmers perception of MFs as a strategy to manage risk at the farm level as well as their intention to participate in such an instrument, both at national and EU level.

In the context of intensified infestation risk due to the growing prevalence of *H. halys* in Italy and simultaneously low adoption rates of subsidized MFs, this paper aims to fill in the unveiled research gap. The study is conducive to the research field of risk management in agriculture and determines key factors that impact upon farmers intention to participate in a MF that covers production losses owed to *H. halys* infestation. The insect represents a major threat for the investigated area. To the best of our knowledge, the underlying study is the first investigation of pest-specific MFs in the field of agricultural risk management. Data were collected in Veneto region in Italy, namely one of the regions mentioned in the EU recommendations that is most affected by the *H. halys* emergency. Consequently, interest in developing innovative risk management tools specific to the risk of *H. halys*, such as MFs, is growing.

Methodologically, the study employs the Least Absolute Shrinkage and Selection Operator (LASSO) algorithm, which is a machine learning technique for features selection (Maruejols *et al.*, 2022) which has recently gained increasing attention in the field of farmers´ behavior studies (Graskemper *et al.*, 2021, 2022; Wang *et al.* 2021). Robustness checks are realized through the application of cross-validated shrinkage penalties.

Derived conclusions are expected to provide valuable insights for both policymakers, local stakeholders (i.e. Defence Consortia or other MFs' managers, government officials, etc.), as well as agricultural advisors. Potential implications of the study entail improved implementation strategies of MFs, adjusted instrument design to manage production risks related to *H. halys*, increase adoption rates and efficient public spending of CAP funds.

The reminder of the paper is organized as follows. The next section provides insights into determinants of risk management tools' adoption at the farm level and – based on this – four hypotheses with respect to farmers' intention to participate in the MF are derived. In a next step, the employed data as well as the empirical framework and estimation strategy are developed. Lastly, obtained results are presented and discussed, followed by a conclusion of the insights.

1.1 Background

Past literature contributions (e.g. for more recent examples, see Hellerstein *et al.*, 2013; Menapace *et al.*, 2016; Meraner and Finger, 2019; Van Winsen *et al.*, 2016) investigated determinants of farmers' adoption of risk management tools, focusing especially on insurances, financial instruments, and self-copying strategies. Identified determinants are risk perception, risk attitude, farmers' preference towards uncertainty as well as time preferences (Coletta *et al.*, 2018). However, literature focusing on risk management instruments implemented by the CAP (e.g. MF) is rare. One exception to this is a recent paper by Giampietri *et al.*

[2] Agricultural crop, animal, and plant insurance (Art. 37 Reg. (EU) 1305/2013 and Reg. (EU) 2017/2393) provides support (up to 70%) for the insurance premium to activate coverage for production losses related to adverse weather conditions, animal and plant diseases, pest infestations, environmental emergencies, exceeding 20% of the farmer's average annual production in the previous three or five years. Although both the mutual fund and the insurance cover yield losses, according to the Italian legislation, the latter is currently mainly focused on weather related risks.

[3] These are: two mutual funds for plant diseases, one mutual fund for wine grape diseases, one mutual fund for arable crop diseases, and a fund for climatic and sanitary risks.

(2020), which provides evidence of farmers' willingness to participate in subsidized MFs and focuses on a sample of Italian farmers.

The authors focus on farmers' intention to participate in a general MF, covering losses due to a bundle of risks specified within the EU legislation (i.e. pests, plant diseases, adverse weather events and environmental emergencies). Giampietri *et al.* (2020) find a negative relationship between farmers' perceived barriers (i.e. scarce perception of benefits, low transparency, and difficult management at the farm level) and the intention to participate. Yet, a positive association between perceived risk frequency at farm level and farmers intention to participate is found. Similar results are obtained with respect to farmers awareness of MF´s new operating rules as provided by the EU legislation, the level of education and the adoption of investments (i.e. new farm structures and technologies) at the farm level. The authors do not find a significant effect for risk attitude. Overall, determinants of participation differ between specific risk management instruments (Giampietri *et al.*, 2020) and findings highlight the necessity of information initiatives to diffuse knowledge (benefits, operative rules, etc.) about specific risk management tools.

Following the literature, this study takes into consideration a broad variety of potential determinants of farmers' participation. The analysis considers the effect of past yield losses (Enjolras and Sentis, 2011) at the farm level on intended participation in an MF specific for *H. halys*. Precisely, losses occurred during the previous five years exceeding the threshold for indemnification (i.e. 20%) are accounted for. Inspired by Wąs and Kobus (2018), who studied the effect of farmers' expectations about both yields and weather conditions on farmers' insurance adoption, the study further accounts for the effect of farmers' predictions about future losses (i.e. related to the next year) due to *H. halys* on the intention to participate in the MF.

In line, additional determinants have been established in the literature to impact farmers' intention to adopt risk management strategies at the farm level. Among these and with a focus on the participation in MFs, the study considers the following: the use of self-copying strategies at the farm level (e.g. diversification, irrigation, off-farm income, etc.) (Enjolras and Sentis, 2011; Meraner and Finger, 2019), insurance adoption (Ye *et al.*, 2017), trust and perceived barriers linked to the participation in the MF (Giampietri *et al.*, 2020), as well as farm and farmer characteristics (see for instance Farrin *et al.*, 2016; Liesivaara and Myyrä, 2017; Menapace *et al.*, 2016; Ogurtsov *et al.*, 2009; Sun *et al.*, 2021; Wąs and Kobus, 2018).

We developed four main hypotheses on the effect of major determinants on farmers' intention to participate in MFs that are tested in the empirical analysis. The hypotheses focus on the effect of risk attitude, risk perception, locus of control and self-efficacy on farmers intention to participate in the MF. The development of each hypothesis is discussed hereafter.

1.2 Hypotheses development

Risk attitude is the position of individuals towards taking risk in the context of uncertainty (van Winsen *et al.*, 2016). Assuming farmers' rationality and following the literature (see for instance Cao *et al.*, 2019; Menapace *et al.*, 2016), we expect farmers who report high risk aversion levels to be more likely to participate in an MF to mitigate losses due to *H. halys* infestation. Risk perception, as opposed to risk attitude, is defined as the subjectively perceived likelihood of an uncertain but adverse event to happen and its resulting subjective impact, e.g. at the farm level (Slovic *et al.,* 1982). Therefore, we assume that high levels of risk perception are positively associated with the intention to participate in a MF to control for the potential risk of *H. halys* infestation (see Meraner and Finger, 2019). The literature on risk management decisions at farm level often analyzed the influence of risk attitude on risk perception (see for instance Nielsen *et al.*, 2013). In this regard, Van Winsen *et al.* (2016) found a non-significant effect between risk attitude and perceived production risk, in contrast to the significantly positive effect on perceived price risk. Risk attitude and risk perception were repeatedly singled out in the literature as main drivers of farmers' risk management decisions (Iyer *et al.*, 2020). We consequently assume that the intention to participate in a MF fund is driven by both subjective concepts. The following hypotheses are tested:

H1. Risk attitude affects the intention to participate in a mutual fund specific for *H. halys* infestation.

H2. Risk perception positively affects the intention to participate in a mutual fund specific for *H. halys* infestation.

In a recent paper, Knapp *et al.* (2021) consider the effect of personality traits, referred to as self-efficacy and locus of control, on choice behavior of Swiss fruit growers to adopt preventive measures against invasive pests (i.e. attacks from *Drosophila suzukii*). The authors find locus of control to be a strong predictor of farmer's behavior with respect to risk management decisions. Elkind (2008) described locus of control as 'the degree to which outcomes are attributed to one's own ability to alter a situation as opposed to external factors'. Locus of control refers to the conviction of farmers to achieve positive production outcomes owed to own ability and skills. Regarding self-efficacy, it is defined as an individual's belief to succeed in a specific domain (Wuepper and Sauer, 2016), hence, it refers to farmers' confidence in their own farming ability. Based on the similarity of our research with that by Knapp *et al.* (2021), this study considers locus of control and self-efficacy as potential determinants of farmers intention to participate in a MF specific for *H. halys* infestation. It is expected that farmers with a strong locus of control to achieve satisfactory production results are less likely to participate in the MF. Similarly, the less farmers believe in the success of their farming abilities, the more likely they are to participate in a MF to counterbalance the risk of infestation. The following hypotheses are tested:

H3: Locus of control affects the intention to participate in a mutual fund specific for *H. halys* infestation.

H4: Self-efficacy affects the intention to participate in a mutual fund specific for *H. halys* infestation.

Concluding, farmers' behavior with respect to risk management is driven by a variety of factors that were established within the context of risk management research, particularly focusing on agricultural insurances. Yet, the perception of CAP-financed MFs in light of an increased risk of *H. halys* infestation has so far not been investigated, consequently requiring an in-depth analysis. Based on developed hypotheses and given the background of risk management behavior at the farm level, the following section develops an empirical framework to determine key predictors of farmers intention to participate in a MF. First, the data employed is summarized.

2. Material and methods

2.1 Data description

The empirical analysis is based on primary farm data collected within a regional project focused on losses suffered by fruit growers as a consequence of *H. halys* infestation[4]. The aim of the survey is to investigate fruit growers' interest in the participation in a mutual fund specifically designed to compensate production losses owed to *H. halys*. Data were collected between November 2021 and May 2022 through an online survey among fruit growers, covering the north-eastern region of Veneto in Italy. As above mentioned, this is one of the regions which was most severely affected by this specific pest infestation since 2019. In this context, a structured questionnaire was distributed via social media to applicable farmers within the region. The full sample contains a total of 90 fruit growers.

The questionnaire was organized in the following way: in section one, after a brief description of the setup of the MF, information was collected on farmers' intention to participate in an MF with a specific design to compensate production losses due to *H. halys* (int). In section two, information on fruit growers' risk attitude (r_att), risk perception (r_perc), self-efficacy (self-eff), and locus of control (locus) was collected,

[4] One of the objectives of the project is to analyze the feasibility of an MF compensating production losses due to *H. halys* and to analyze fruit growers' preferences for the instrument and their intention to participate in it.

as well as perceived advantages of the instrument (adv), perceived barriers to participation (barr), and trust (trust). An overview of the investigated variables can be found in Table 1. As for r_att, self_eff, and locus, the Likert-type scales' items were derived from Knapp *et al.* (2021)[5], with some adjustments. The items for advantages, barriers and trust mainly derived from Giampietri *et al.* (2020) with adjustments. In the analysis, we used the average score of the items of each scale as risk attitude, self-efficacy, locus of control, perceived advantages, barriers, and trust. Risk perception was computed as a product of risk probability and risk impact (see Meraner and Finger, 2019). Section three collected information related to self-coping strategies already used by farmers (see variables whose code starts with 'sc_'), such as the use of less vulnerable varieties, diversification, irrigation, anti-hail nets or insect nets, higher workload, cooperation, off farm work, agricultural insurance, pesticide, and traps. The description of these variables and additional control variables can be found in Table 2 and 3. Moreover, the last part of the questionnaire was devoted to retrieving socio-demographic characteristics of the farmers.

On average, farmers of the sample are 51 years old and predominantly male (90%). 64% of all farmers in the sample completed high school of which 21% also hold a bachelor's degree. The majority of farmers in the sample are full-time farmers (81%). The predominant types of fruits grown in the sample are apples (31%) and pears (21%). Over the last five years, almost the entire sample has suffered production losses above the compensation threshold of the mutual fund (20%). During the next year, 47% of the sample expects production losses above the compensation threshold owed to *H. halys* attacks. The main self-coping strategies adopted by farmers are insurance (87%), irrigation (68%), and the use of pesticides (61%). 34% of all farmers attended two to three training courses in the previous year, whereby 49% of the sample took courses targeted for risk management. An extensive overview of the sample of interviewed fruit growers is reported in Supplementary Table S1.

Summary statistics reveal the following. Overall, summary statistics indicate a weak tendency of farmers towards participation in the MF with a mean value of 0.57. The risk attitude of farmers in the sample shows a mean value of 4.49 and a median of 4.5, hence farmers are neither highly risk averse nor risk seekers. Likewise, farmers evaluate their self-efficacy with a mean value of 5.1, that is neither a high nor low level of confidence in own farming abilities. A standard deviation of 1.53 confirms that most farmers estimate their farming ability to centre around the mean. Risk perception, however, is comparatively high with a mean value of 16.57. The value indicates that farmers perceive the probability of production losses due to *H. halys* and the associated impact (i.e. production loss) as a substantial threat to economic profitability. Farmers' capacity to determine positive production outcomes at the farm level shows a mean value of 6.67. Hence, we can conclude that farmer's expectation to influence on production outcomes is considerably higher than farmers' evaluation of own farming abilities.

The dimensionality of the data is high compared to the number of observations. Given the nature of the data and the context of the analysis, machine learning techniques are considered adequate in the context of the present study. A detailed explanation is given in the next section. Out of a broad variety of potential

[5] Knapp *et al.* (2021) derived the scales from: Dohmen *et al.* (2011) for risk attitude, Bandura and Wood (2006) for self-efficacy, and Rotter (1966) for locus of control.

Table 1. Summary statistics of main variables for hypotheses testing.[1]

Variable	No. Obs.	Mean	Median	Min	Max	SD
Intention to participate	90	0.57	1.0	0	1	0.50
Risk attitude	90	4.39	4.5	0	10	2.16
Risk perception	90	16.57	16.0	1	25	6.71
Self-efficacy	90	5.10	5.0	1	9.3	1.53
Locus of Control	90	6.67	6.7	1	10	1.67

[1]SD = standard deviation.

Table 2. Description of main variables for hypotheses testing.[1]

Variable (code)	Question	Description	Value	Type
Intention to participate (int)	In the coming years, should it be implemented, would you be willing to participate in a mutual fund against risks related to *H. halys*?		1 = yes, 0 = no	D
Risk attitude (r_att)	Are you willing to take risks, or do you try to mitigate risks? Indicate how much are you willing to take:	production risk	0 = not at all, 10 = very much	Ave
		commercialization risk		
		financial risk		
		agricultural risk in general		
Risk perception (r_perc)	*Risk probability*: Regardless of the level of damage, how do you assess the likelihood of production losses due to *H. halys*?		1 = completely unlikely, 5 = completely likely	Co
	Risk impact: Should you be attacked from *H. halys*, how would you assess their impact (production loss) in your farm?		1 = minimum impact, 5 = maximum impact	
Self-efficacy (self_eff)	When I encounter difficulties in agricultural production, I can usually think of a solution		0 = do not agree, 10 = fully agree	Ave
	I am confident that I can accomplish my production goals at the end of the harvest			
	I can solve most of the problems related to pests (such as *H. halys*) if I invest the necessary effort			
Locus of control (locus)	How successful my fruit production is depending mostly on my skills as a farmer		0 = do not agree, 10 = fully agree	Ave
	My production is more dependent on the weather than on what I do[2]			
	Success in the fruit production can only be slightly influenced by farmers[2]			

[1] Variable type: D = dummy, Ave = average score (continuous variable), Co = continuous. R_att = from 0 to 4 = risk averse, 5 = risk neutral, from 6 to 10 = risk seeker.

[2] Reversed.

predictors, a set of key factors that drive fruit grower's intention is selected through the application of the LASSO algorithm. By this means, relevant features in the context of risk behavior were determined to improve performance of CAP-funded management tools such as the MF in the future. Our analysis applies novel techniques in agricultural risk management, consequently contributing to innovative approaches to tackle adverse effects of pest infestation.

2.2 Empirical framework

The aim of the study is to investigate farmers willingness to participate in a *H. halys* specific MF within the framework of high dimensional data. In other words, the number of predictors is comparatively large in relation to the sample size, which is why the analysis assumes approximate sparsity and applies LASSO regularization method (Chernozhukov *et al.*, 2016). Research in the context of risk management strategies and adaptation behavior of MFs is scarce. Little is known about what drives participation in MFs. To exploit this new field of research, machine learning techniques are in advantage to learn from the data rather than

Table 3. Description of control variables.[1]

Variable (code)	**Question**		**Type**
Age in years (age)			Co
Sex (sex)	1=male, 2=female		Ca
Education (edu)	1=primary school, 2=secondary school, 3=high school, 4=university		Ca
Utilized agricultural area in hectares (uaa)			Co
Farm age (farm_age)	Since how many years do you manage the farm?		Co
Province (prov)	1=Belluno, 2=Padova, 3=Rovigo, 4=Treviso, 5=Venezia, 6=Vicenza, 7=Verona		Ca
Full-time farming (full)	1=yes, 0=no		D
Organic farming (organic)	1=yes, 0=no		D
Farm type (farm_type)	1=specialized farm, 0=otherwise		D
Fruit type (fruit_type)	1=cherries, 2=apricots, 3=peaches, 4=pears, 5=apples, 6=plums, 7=hazelnuts, 8=actinidia, 9=pomegranates, 10=walnuts, 11=chestnuts, 12=persimmons, 13=other		Ca
Farm future (farm_future)	1=rented, 2=sold, 3=inherited, 4=managed by partners/fam. members, 5= I will continue to manage it		Ca
Use of resistant varieties (sc_resvar)	1=yes, 0=no		D
Diversification (sc_nonagridivers)	1=yes, 0=no		D
Irrigation (sc_irrigat)	1=yes, 0=no		D
Anti-hail net (sc_hail_net)	1=yes, 0=no		D
Anti-insect net (sc_insect_net)	1=yes, 0=no		D
Higher workload (sc_work_more)	1=yes, 0=no		D
Cooperation with farmers (sc_coop)	1=yes, 0=no		D
Off farm work (sc_off_farm)	1=yes, 0=no		D
Insurance adoption (sc_insura)	1=yes, 0=no		D
Use of pesticides (sc_pesticid)	1=yes, 0=no		D
Use of traps (sc_trap)	1=yes, 0=no		D
Past losses (past_loss)	Have you suffered a production loss greater than 20% in the last 5 years?	1=yes, 0=no	D
Future losses from *H. halys* (fut_loss_hh)	Which will be the amount of production loss due to *H. halys* in your farm the next year?	1=0%, 2=1-10%, 3=11-20%, 4=more than 20%	O
Future losses from pests/ diseases (fut_loss_pest_ disease)	Which will be the amount of production loss due to other pests/diseases in your farm the next year?	1=0%, 2=1-10%, 3=11-20%, 4=more than 20%	O
Training courses (train_ courses)	How many training courses do you attend on average each year?	1=none, 2=1, 3=2-3, 4=4-5, 5=more than 5	O
Training courses on risk management (rm_courses)	Did you attend specific training course on risk management in the last year?	1=yes, 0=no	D
Advantages (adv)	MF membership promotes a secure and stable income for my farm		Ave
	MF membership enables the management of non-insurable risks		
	I believe that an instrument managed by farmers can work well		
Barriers (barr)	MF membership is difficult to manage at farm level		Ave
	Information on MF functioning is scarce		
	The benefits of joining the fund are not entirely clear to me		
	Technical support for joining the fund is poor		
	There is little transparency in the MF management mechanisms		
Trust (trust)	I have little trust in the fund managers		Ave
	I have little trust in other fund members (farmers)		
	I have a lot of uncertainty about the indemnification from the MF		
	I am very uncertain about the amount of compensation I will receive		

[1] Variable type: Co = continuous, Ca = categorical, O = ordinal, D = dummy, Ave = average score (continuous variable). The scale for adv, barr and trust was a 5-pt Likert type scale from 1 (fully disagree) to 5 (fully agree).

build on *a priori* established theories or assumptions (Baylis *et al.*, 2021; Storm *et al.*, 2020). As opposed to traditional econometric analysis, machine learning methods further count with higher flexibility with respect to the functional form of the model. These methods allow for multidimensional non-linear relationships (Storm *et al.*, 2020). Determinants of MF participation are assumed to count with complex properties, that are adequately captured by the LASSO algorithm. Participation in a *H. halys* related mutual fund is positively linked to economic profitability through risk reduction and diversification. The intention of the instrument is to foster farmers' resilience with respect to pest-induced yield losses. Yet, risk aversion and perception as well as locus of control and self-efficacy are expected to impact upon farmers' intention to participate in the fund. The prediction of specific determinants, which condition farmers' intention, can contribute to an improved instrument design, extending the scope and effectiveness of the tool. Insights into determinants of participation can further contribute to reduce access barriers. Consequently, results count with policy relevance.

To select key determinants, the regularization technique LASSO is implemented. The algorithm minimizes the log likelihood as well as a shrinkage penalty. The shrinkage penalty consists of the tuning parameter lambda (λ) and the sum of corresponding normalized beta coefficients. Compared to ordinary least squares (OLS), LASSO reduces variance but induces some bias to the coefficients (Friedman *et al.*, 2010; Lesmeister, 2016; Maruejols *et al.*, 2022). Recently, Knapp *et al.* (2021) have used a similar approach in their study on Swiss farmers but focusing on farmers' choices and insurance uptake.

Formally, LASSO is expressed as:

$$\min(log - likelihood + \lambda \textstyle\sum_{i=1}^{N} |\beta_i|) \tag{1}$$

The function of intended participation is defined as follows:

$$P(y = 1, x) = \frac{e^{X\beta}}{1+e^{X\beta}} \tag{2}$$

where y=1 denotes the willingness to participate in a MF for *H. halys*, and y=0 non-participation, respectively. X represents a vector of variables, that determine participation, and β the corresponding vector of coefficients. The binomial normal log-likelihood function for Equation 2 is:

$$Max.\, l(x, \beta) = Max. \textstyle\sum_{i=1}^{N} y_i * (X_i\beta) - \sum_{i=1}^{N} \log(1 + e^{X_i\beta}) \tag{3}$$

where N designates the sample size. LASSO was employed to improve prediction accuracy of feature selection. The algorithm represents a powerful tool to avoid multicollinearity of independent covariates during the model building process via the penalization of the negative log-likelihood function (Equation 3) (Friedman *et al.,* 2010). The binomial normal LASSO log-likelihood function is defined as:

$$Min. \log\left(1 + e^{X_i\beta}\right) - \frac{1}{N} \textstyle\sum_{i=1}^{N} y_i * (X_i\beta) + \lambda \sum_{i=1}^{N} |\beta_i| \tag{4}$$

where λ is the regularization parameter which determines the strength of the shrinkage penalty. The choice of λ is fundamental for the model building process as the selection of key predictors by the LASSO algorithm strongly depends on the penalization of covariates. Covariates with little explanatory power are shrunk towards zero whereas the algorithm keeps relevant features in the regression (Lesmeister, 2015; Maruejols *et al.*, 2022). One main assumption is that in a high dimensional setting, the majority of coefficients are adjacent to zero (Friedman *et al.*, 2010). Therefore, only covariates with strong explanatory power for the outcome are selected. In case that the value of the tuning parameter λ is too large, key variables are omitted and coefficients are shrunk excessively, despite their explanatory power for the outcome. Thus, the predictive accuracy of the model deteriorates.

To account for mentioned limitations, the study implements a series of literature-based, diverging regularization parameters to confirm the robustness of results. Model selection is based on prediction performance. The baseline model employs a rigorous LASSO technique as suggested in Chernozhukov *et al.* (2016), where the choice of λ is theoretically grounded and data driven. To cross-validate results and to confirm the robustness of feature selection, the analysis subsequently alters the regularization parameter based on Friedman *et al.* (2010), who provided a distinct approach to define λ. Contrarily to a data-driven penalty as in rigorous LASSO (Belloni *et al.*, 2013; Chernozhukov *et al.*, 2016), Friedman *et al.* (2010) use k-fold cross validation and apply a series of different penalty factors. The prediction performance of each model hinges on the level of λ. For logistic regressions, the authors determine the optimal value of the tuning parameter (λ) as the value that minimizes the prediction error of the model through cross-validation. The prediction error is measured in binomial deviance. In other words, the optimal value of lambda minimizes binomial deviance. Alternatively, Friedman *et al.* (2010) apply values of lambda within one standard error of the prediction error to compare model performance. That is the maximum value of lambda at which binomial deviance is within one standard error of the smallest binomial deviance. If different penalty factors are applied in the same regression, that is the application of individual penalty factors for each covariate to keep certain variables in the model, weighted LASSO is applied. The analysis compares model performance of rigorous LASSO with penalized regressions that employ a cross-validated value of lambda that either minimizes the prediction error or is within one standard error of the prediction error. Consequently, the estimation strategy is subdivided into baseline model (rigorous LASSO) and robustness check (cross-validation).

3. Results

3.1 Baseline results

The baseline approach of the empirical analysis entails three subsequent stages (Column 1 to 3), whereby the final baseline model is reported in Column 2, Table 4. In a first stage, a generalized logistic model is estimated (Column 1, Table 4), which reports predicted log-odds of the intention to participate in the MF for *H. halys* infestation. The model takes into consideration the full set of covariates included in the dataset. Overall, positive values indicate a positive association between covariates and the outcome variable, whereas negative values show a negative link (MacKenzie *et al.*, 2018). Accounting for the full set of independent covariates in the first model specification (Column 1), none of the variables yields statistically significant results, potentially indicating multicollinearity of predictors or inadequate choice of covariates. Also, log-odds are difficult to interpret.

To circumvent concerns with respect to model specification and variable selection, the second stage of the baseline analysis entails a rigorous LASSO logit estimation (Table 5). Rigorous LASSO computes a data-driven penalty to the regression term (Belloni *et al.*, 2013), as opposed to a cross-validated choice of penalty (Friedman *et al.*, 2010). Specific coefficients of selected covariates by the algorithm are reported Table 4. Out of many potential covariates, a set of four non-zero coefficients are determined as key predictors of farmers' intention to participate in a *H. halys* specific MF. Key predictors are farmers risk perception with respect to yield losses due to *H. halys* (r_perc), the age of the farm (farm_age), the perceived assessment of future yield losses due to *H. halys* (fut_loss_hh) as well as perceived advantages of participation in the fund (mean_adv). It is mainly covariates of subjective perception of risks and benefits as well as the maturity of the farm that appear to drive farmers' decision to participate in the MF. Interestingly, socio-economic covariates, such as farmer age, gender, and the type of cultivation, i.e. fruit type, are not selected by the algorithm. Similarly, perceived barriers to access the instrument or a lack of trust in the fund do not appear to predict fruit farmers' intention to participate in the mutual fund. It is rather perceived benefits of participation that are linked with farmers intention.

Based on the selection of influential predictors through LASSO (Table 5), the third stage computes a reduced logistic regression (Colum 2, Table 4). This model is also referred to as baseline model from now on. Due to the limited interpretability of odd ratios and to account for absolute changes in the outcome, estimates of marginal effects at the mean are reported in Column 3 of Table 4. Apart from expected production losses due to *H. halys* in the following year, all selected covariates yield statistically significant results in the reduced

Table 4. Baseline results (Logistic Regression, reduced Logit, Marginal Effects at the mean).[1]

	Logistic regression (log-odds units)		Reduced Logit (log-odds units)		Marginal effects (percent points)	
(intercept)	-1,764.173	(1,269,761.105)	-8.893***	(2.042)		
r_perc	22.225	(127,17.224)	0.146**	(0.052)	0.021***	(0.006)
age	13.150	(6,617.772)				
sex	-150.089	(319,047.606)				
edu	45.374	(45,727.730)				
uaa	0.276	(774.564)				
farm_age	-6.311	(6,665.450)	0.050 *	(0.024)	0.007 *	(0.003)
prov	2.576	(16,924.628)				
full	-85.937	(113,089.329)				
organic	-153.961	(189,335.286)				
farm_type	168.683	(143,055.153)				
fruit_type	0.177	(16,125.620)				
train_courses	59.071	(23,206.873)				
rm_courses	87.743	(170,782.431)				
sc_resvar	69.076	(233,445.724)				
sc_nonagridivers	243.314	(298,355.981)				
sc_irrigat	31.892	(120,564.099)				
sc_hail_net	6.937	(152,941.751)				
sc_insect_net	140.184	(233,364.181)				
sc_work_more	-84.406	(143,182.502)				
sc_coop	37.735	(117,480.488)				
sc_off_farm	150.304	(116,486.058)				
sc_insura	-200.958	(393,965.902)				
sc_pesticid	-21.419	(216,704.830)				
sc_trap	-106.922	(256,995.892)				
farm_future	-12.551	(69,262.047)				
past_loss	-109.916	(299,481.878)				
fut_loss_hh	44.422	(45,758.442)	0.313	(0.340)	0.045	(0.048)
fut_loss_pest_disease	21.326	(32,927.175)				
mean_r_att	-6.753	(18,059.527)				
mean_self_eff	57.468	(33,106.758)				
mean_locus	-15.387	(59,828.324)				
mean_adv	116.989	(94,573.149)	1.250**	(0.445)	0.178***	(0.053)
mean_barr	90.999	(158849.944)				
mean_trust	-45.075	(150528.455)				
AIC	70.000		89.805		89.805	
BIC	157.493		102.304		102.304	
Log likelihood	-0.000		-39.902		-39.902	
Deviance	0.000		79.805		79.805	
N.obs.	90		90			
Deviance (Null)					123.162	
df.null					89	
DF Resid.					85	
N.obs					90	

[1] Standard errors are reported in parenthesis. Significance at * *P*<0.05, ** *P*<0.01, *** *P*<0.001.

Table 5. Post LASSO coefficient selection.

Outcome variable	Variables chosen	Coefficient
Intention to participate in a MF for *H. halys* (int)	Intercept	8.893
	risk_perc	0.146
	farm_age	0.050
	fut_loss_hh	0.313
	mean_adv	1.250

logistic regression model. In terms of magnitude, it is the mean of expected advantages from participation in the fund that is most strongly associated with farmers' intention to participate. Likewise, despite of lower magnitude, the perception of *H. halys* associated risks appear to drive farmers intention to participate in the MF. Since the marginal effect of mean_adv is remarkably larger than the coefficient for r_perc, advantages associated with the participation in the fund appear to outweigh the impact of *H. halys* related risk perception. Besides, it is the farm age that positively affects farmers intention to participate in the fund, highlighting the importance of farm maturity and farming experience for farmers to adopt innovative risk management tools at the farm level.

In sum, participation in the MF appears to be distinctly influenced by farmers' expectations of future benefits. Furthermore, it is farm maturity rather than socio-economic factors that appear to have an effect on farmers intention to participate.

3.2 Robustness check through cross-validation

The selection of covariates through LASSO strongly depends on the penalization of covariates. To confirm robustness of results, the analysis subsequently cross-validates the selection of key predictors according to Friedman *et al.* (2010). The optimal value of lambda that minimizes the prediction error of the model is determined through k-fold cross-validation and defined either as lambda that minimized binomial deviance (Lambda min) (Models 2 and 4) or the maximum value of lambda at which binomial deviance is within one standard error of the smallest binomial deviance (Lambda 1 SE) (Models 1 and 3). In Models 1 and 3, the optimal lambda is 0.142 and 0.099, respectively. The value of lambda is 0.056 in Model 2 and 0.047 in Model 4.

Figure 1 plots the binomial normal log-likelihood against the log of lambda for Models 1 and 2. The left line represents the value of lambda that minimizes binomial deviance whereas the right line shows the value

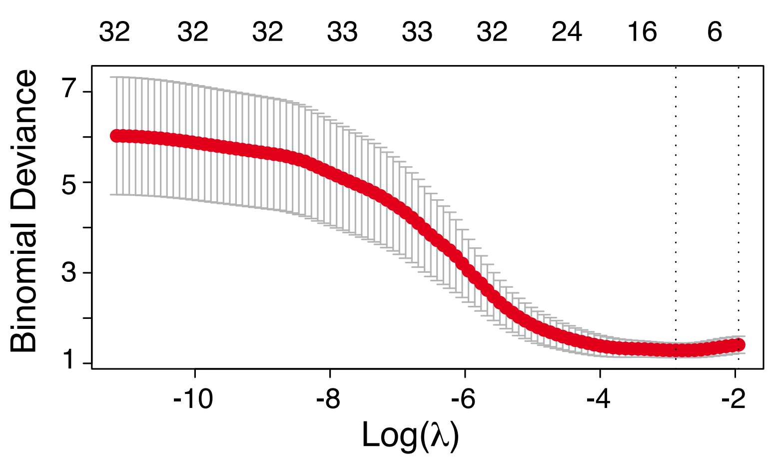

Figure 1. The relation between log(lambda) and binomial deviance.

of lambda within one standard error of the minimum binomial deviance. Overall, the red line defines the mean cross-validated error curve plus one standard deviation bandwidth (Friedman *et al.*, 2010). Binomial deviance is minimized when only a small number of covariates (top legend of Figure 1) is considered and the value of lambda low (bottom legend of Figure 1).

Furthermore, Models 1 and 2 are weighted LASSO regressions that apply different penalty factors to each coefficient. Default penalty factors of one are applied to all variables except for risk perception of *H. halys* infestation (r_perc) as well as the mean of risk attitude (mean_r_att), locus of control (r_locus), and self-efficacy (mean_self_eff). A penalty factor of zero is applied to mentioned covariates to keep them in the model. Coefficients are penalized with equal strength (same value of lambda) if the default penalty factor of one is applied. A penalization factor of zero indicates no penalization of coefficients; that is equivalent to the OLS estimator (Hastie *et al.*, 2021). Table 6 summarizes cross-validated, weighted LASSO regression results of Model 1 and 2. Comparing Model 1 and 2, the number of selected covariates decreases with an increasing value of lambda. In Model 1, only non-penalized covariates are selected; namely risk perception, the mean of risk attitude, self-efficacy and locus of control. Covariates selected in Table 5 (farm age, future losses, mean advantages) are not selected in Model 1. Similarly, socio-economic covariates are shrunk towards zero and not included in Model 1. Considering Model 2, which applies the value of lambda that minimizes binomial deviance, the selection of covariates is in line with the rigorous baseline model. Additional covariates are selected, such as farm type or insurance adoption. Again, socio-economic characteristics are not selected in Model 2.

In Model 3 (Lambda 1 SE) and 4 (Lambda Min), we conduct k-fold cross validation without defined penalty factors. Results are displayed in Table 7 comparing non-weighted LASSO regressions results for Lambda min and Lambda 1 SE (Supplementary Figure S1 shows the relationship between lambda and binomial deviance without penalty factor). Model 3 selects identical predictors as the rigorous baseline model of farmers intention to participate in the MF. Other covariates shrink to zero. In Model 4, all baseline predictors are included in the model. The predictive strength of baseline predictors with respect to farmers' intention to participate in the instrument is consequently confirmed. Besides, self-copying strategies such as insurance adoption or a higher workload at the farm as well as participation in training courses are selected in Model 4.

Model specifications with either data-driven or cross-validated penalization parameters differ in variable selection. Divergences between model specifications are considerably subtle and hinge on small difference in lambda. Although results are sensitive to the adequate choice of the shrinkage parameter lambda, the

Table 6. Cross-validated, weighted LASSO.

Model 1: Lambda 1 SE		**Model 2: Lambda min**	
Penalization Factor		**Penalization Factor**	
Variables	**Estimate**	**Variables**	**Estimate**
(intercept)	-2.535	(intercept)	-6.936
r_perc	0.177	r_perc	0.167
mean_r_att	-0.089	mean_r_att	-0.043
mean_self_eff	0.093	mean_self_eff	0.160
mean_locus	-0.023	mean_locus	-0.002
Lambda	0.142	farm_age	0.021
		farm_type	0.070
		train_courses	0.134
		sc_insura	-0.128
		fut_loss_hh	0.096
		mean_adv	0.726
		Lambda	0.056

Table 7. Cross-validation without penalization.

Model 3: Lambda 1SE		Model 4: Lambda min	
No penalization		**No penalization**	
Variables	**Estimate**	**Variables**	**Estimate**
(intercept)	-3.049	(intercept)	-6.515
r_perc	0.059	r_perc	0.096
farm_age	0.014	age	0.005
fut_loss_hh	0.201	sex	-0.0004
mean_adv	0.381	edu	-0.002
Lambda	0.099	farm_age	0.022
		full	0.006
		farm_type	0.230
		train_courses	0.214
		sc_work_more	0.176
		sc_insura	-0.297
		farm_future	0.055
		fut_loss_hh	0.316
		mean_adv	0.715
		Lambda	0.047

robustness of results is confirmed. Models 2, 3 and 4 select identical covariates as the baseline model. Both data-driven and cross-validated shrinkage penalties confirm the selection of covariates, strengthening the validity of determined predictors of farmers' intention to participate in a *H. halys* specific MF. Overall, farmers' intention to participate in a *H. halys* specific MF is driven by the interplay between perceived risks and benefits. Although expected advantages of participation appear to outweigh the impact of risk on the decision in terms of magnitude, it is risk perception that continuously is selected by the algorithm to predict Italian fruit growers' intention to participate in the MF. Independent of the choice of lambda or penalty factor, risk perception is included in all models. The age of the farm and thus its maturity and cumulated experience further appear to impact the balance between risks and benefits. We therefore conclude that perceived risk of future yield losses linked to *H. halys* decisively impacts farmers' intention to participate in the MF, coupled with fund-related advantages and the farm's maturity.

4. Discussion

To boost farmers' resilience against pest-related yield losses, a set of key predictors is singled out. The analysis provides empirical evidence for a link between farmers' intention to participate in a MF designed to cover production losses due to *H. halys* and fruit growers' risk perception and perceived benefits. Throughout the empirical analysis, different model specifications were defined to obtain robust results. Specifications vary with respect to shrinkage penalty and penalty factor. Based on obtained results, established hypotheses are discussed hereafter.

Previous literature has shown the link between farmers' risk attitude and adoption behavior of on-farm risk management tools (Iyer *et al.*, 2020). Menapace *et al.* (2013) expect farmers to preserve an adverse attitude towards risk (H1). Most research has focused on agricultural insurances as predominant risk management tool, showing mixed results regarding the link between risk attitude and the adoption of the instrument (see for instance Hellerstein *et al.*, 2013; Santeramo, 2019; Van Winsen *et al.*, 2016). Our study shows that despite its inclusion in the dataset, risk attitude is not determined as key predictor for participation in the fund, consistent with results of Giampietri *et al.* (2020). The hypothesis of an impact of risk attitude on MF participation (H1) is consequently rejected.

Risk perception is selected as a key predictor (H2), thus it appears to impact farmers' intention to participate in the fund. Obtained findings are in line with Giampietri *et al.* (2020), who establish a positive link between risk perception and the intention to participate in MFs. Alongside a strong perception of the insect-related risk among farmers and thus the likely interest in the instrument, the decision to create a mutual fund to tackle *H. halys*-related risks appears sound.

Locus of control summarizes the farmer's capacity to positively impact upon production outcomes at the farm level. The assessment of its effect on farmers intention to participate in the fund (H3) does not show a clear relationship. Locus of control is selected as key predictor in Model 1 and 2 of the cross-validation, however, the covariate is not included in the baseline model. Consequently, the hypothesis of locus of control affecting farmers intention to participate cannot be confirmed within the context of this research, contrasting with Knapp *et al.* (2021).

Self-efficacy represents farmers´ self-assessment of farming ability in terms of efficiency and success. The covariate is selected in Model 1 and 2 of the cross validation, yet, not included in the baseline model. Singular predictor related to self-efficacy (such as additional workload) are however selected in some model specifications. Results with respect to self-efficacy are rather ambiguous. The assumption of farmers self-efficacy on the intention to participate in the MF is therefore neither confirmed nor rejected (H4). In their recent paper, Knapp *et al.* (2021) find a negative association between self-efficacy and the decision to adopt hail insurance. Their results highlight the ambiguity between self-efficacy and farmers intention to participate in risk management tools.

In general, additional key determinants are selected by the algorithm that were not previously specified as hypotheses. Expected future losses due to *H. halys* infestation are included in the baseline model, indicating an effect of future expectations on present actions and adaptation behavior. Farmers might fear that future events, such as unforeseen pest infestations, jeopardize present or past efforts. Unforeseen events are out of farmers' active control, hence, positively impacting upon their intention to participate in the MF. Advantages associated with the participation in the MF are assumed to balance out adverse effects of risk perception and potential future losses.

Apart from future production losses, it is Model 4 of the cross-validation that selects a variety of farmer and farm characteristics, e.g. farmer age, gender, educational level, and farm type. Generally, a strong impact of individual farm characteristics is assumed to influence farmers' intention to participate. However, it is only farm age that is included in the baseline model. Farm age is a proxy for maturity and cumulated experience also with risk, representing farmers' past efforts and ability to meet production goals. In line, Enjolras and Sentis (2011) testify a positive link between the intention to apply risk management instruments and the use of self-copying strategies. Also, Meraner and Finger (2019) found more risk literate farmers being more likely to use off-farm measures as insurance. Being innovative and applying farm-level solutions to cope with risk and adverse effects of pest infestation also represents one component of self-efficacy. On that account, both Model 2 (insurance adoption) and Model 4 (insurance adoption and additional workload) select self-copying strategies as influential predictors for fund participation.

5. Conclusions

H. halys infestation represents a sanitary emergency in many regions of Northern Italy, causing high production losses for fruit growers and compromises economic stability. With the intention to counteract this risk, the EU CAP promotes the adoption of innovative risk management tools such as mutual funds. Uptake rates are so far low and the success of the instrument uncertain. Consequently, an improved understanding of determinants that drive farmers´ intention to participate in the tool is required. Due to the multidimensional nature of determinants that impact upon farmers intention, the study applies LASSO algorithms to predict key determinants.

The empirical analysis reveals that fruit growers' intention to participate in a specified MF which covers production losses due to *H. halys* is driven by the interplay between perceived risks and benefits. Our rigorous baseline model selects risk perception with respect to *H. halys* related yield losses, the age of the farm, the assessment of future production losses due to *H. halys* infestation, as well as perceived advantage of participation in the MF as key predictors for the intention to participate in the risk management instrument. This selection is confirmed through cross-validation of the shrinkage penalty lambda. Except for the covariate of expected future production losses due to *H. halys* infestation, all covariates yield statistically significant results in the post LASSO logistic regression.

We conclude that farmers intention to participate in a *H. halys* specific MF is strongly linked to perceived benefits of the instrument. Furthermore, it is farmers' perception of risk concerning future yield losses due to *H. halys* infestation that decisively impacts upon the intention to participate in the MF, coupled with MF-related advantages as well as farm maturity. As opposite, socio-economic covariates are not selected, indicating a weak effect on farmers' intention to participate in the MF. Also, risk attitude, perceived barriers to access the instrument or a lack of trust in the fund do not predict farmers' intention of participation.

Policy implications, which can be derived from gained insights, comprise the importance of an adequate instrument design for MFs. For the tool to be successful, accessibility and knowledgeability of farmers is required. Indeed, among farmers, information about MFs is scarce. Promoting benefits associated with the participation in the fund is therefore likely to increase uptake rates of MFs since perceived benefits are key drivers for farmers' intention to participate in the fund. Hence, our finding indicates that better communication on this innovative risk management tool might be needed to allow farmers to foster their resilience also diversifying risk management strategies outside of on-farm. Ultimately, it is risk perception that is strongly linked with intended participation. However, to date, legislative burdens hamper farmers' ability to access the instrument according to their needs. It follows that the decision to create a mutual fund for *H. halys* represents a sound idea if we consider the high perceived risk of losses related to insect attacks among farmers and its role as a key predictor of intention to participate in the fund.

Generally speaking, it would be very useful for the managers of MFs to have at their disposal information related to the individual risk perception and perceived future loss of farmers to be combined with information on loss distribution, in order to preliminary assess and eventually prevent problems of adverse selection.

The contribution of the study is twofold. To the best of the authors knowledge, this study is the first to analyze farmers' intention to participate in a CAP-subsidized mutual fund that covers production losses due to a pest infestation, namely *H. halys*. Moreover, the study assesses risk management strategies at the farm level, taking into consideration a large set of potential predictors. Methodologically, the study applies the machine learning technique LASSO, which is a novel approach in risk management studies.

Nevertheless, the study is not without limitations. First, the analyzed sample is not representative for Italian fruit growers on a national level and the sample size is small. Future research may amplify the sample size to analyze the effect of behavioral and personality traits on MF adaptation rates. Besides, the extension of the study to other countries of the European Union would enable cross-country comparison in farmers' behavior. Second, LASSO estimates are inherently skewed as they induce bias to the coefficients. Effect sizes need to be interpreted cautiously. In general, fostering farmers' resilience against adverse agricultural risks – exacerbated through climate change – is not only a concern of fruit growers in Northern Italy, but evolves more and more into a global problem. This study consequently contributes to an improved understanding of farmers' interest towards more innovative risk management tools, such as MFs. To conclude, similar analyses, which can be applied to other countries or other pest infestations, are expected to provide valuable insights for policy makers, MFs' related stakeholders, and agricultural advisors, to facilitate the creation of mutual funds which are closer to farmers needs and to increase participation rates among them.

Supplementary material

Supplementary material can be found online at https://doi.org/10.22434/IFAMR2022.0086

Table S1. Descriptive statistics of the sample.
Figure S1. Binominal deviance (k-fold cross validation with no penalty factor).

Acknowledgements

We thank Xiaohua Yu for his assistance and precious comments on an earlier version of the paper. Financial support from the research project: 'Regional plan to tackle the spread of allochthonous insects harmful to fruit production – DGR Veneto Region No. 611 (19.05.2020)'. Action 5: 'Feasibility study of a mutual fund to cover the risk associated with pest infestations in fruit growing' (CUP H19E20000150002).

Conflict of interest

The authors declare no conflict of interest.

References

Bandura, A. and R. Wood. 1989. Effect of perceived controllability and performance standards on self-regulation of complex decision making. *Journal of Personality and Social Psychology* 56(5): 805. https://doi.org/10.1037//0022-3514.56.5.805

Baylis, K., Heckelei, T. and H. Storm. 2021. Machine learning in agricultural economics. In *Handbook of Agricultural Economics* 5: 4551-4612. Elsevier. https://doi.org/10.1016/bs.hesagr.2021.10.007

Belloni, A., V. Chernozhukov, I. Fernández-Val and C. Hansen. 2013. Program evaluation and causal inference with high-dimensional data. *Econometrica* 85(1): 233-298. https://doi.org/10.3982/ECTA12723

Cao, R., A. Carpentier and E. Zurich. 2011. Measuring farmers' risk aversion: the unknown properties of the value function. Paper prepared for presentation at the EAAE 2011 Congress: Change and Uncertainty, Challenges for Agriculture, Food and Natural Resources, ETH Zurich, Zurich, Switzerland, 30 August-2 September 2011.

Chernozhukov, V., C. Hansen and M. Spindler. 2016. High-dimensional metrics in R. Available at: http://R-Forge.R-project.org. https://doi.org/10.48550/arXiv.1603.01700

Coletta, A., E. Giampietri, F.G. Santeramo, S. Severini and S. Trestini. 2018. A preliminary test on risk and ambiguity attitudes, and time preferences in decisions under uncertainty: towards a better explanation of participation in crop insurance schemes. *Bio-based and Applied Economics* 7(3): 265-277. https://doi.org/10.13128/bae-7679

Cordier, J. and F.G. Santeramo. 2019. Mutual funds and the income stabilisation tool in the EU: retrospect and prospects. *EuroChoices* 19(1): 53-58. https://doi.org/10.1111/1746-692X.12210

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp and G.G. Wagner. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* 9(3): 522-550. https://doi.org/10.1111/j.1542-4774.2011.01015.x

Elkind, P.D. 2008. Perceptions of risk, stressors, and locus of control influence intentions to practice safety behaviors in agriculture. *Journal of Agromedicine* 12: 7-25. https://doi.org/10.1080/10599240801985167

Enjolras, G. and P. Sentis. 2011. Crop insurance policies and purchases in France. *Agricultural Economics* 42(4): 475-486. https://doi.org/10.1111/j.1574-0862.2011.00535.x

European Commission (EC). 2013. Regulation (EU) No 1305/2013 of the European Parliament and of the Council of 17 December 2013 on support for rural development by the European Agricultural Fund for Rural Development (EAFRD) and repealing Council Regulation (EC) No 1698/2005. *Official Journal of the European Union* L 347, 20.12.2013: 487-548.

European Commission (EC). 2017. Regulation (EU) 2017/2393 of the European Parliament and of the Council of 13 December 2017 amending Regulations (EU) No 1305/2013 on support for rural development by the European Agricultural Fund for Rural Development (EAFRD), (EU) No 1306/2013 on the financing, management and monitoring of the common agricultural policy, (EU) No 1307/2013 establishing rules for direct payments to farmers under support schemes within the framework of the common agricultural policy, (EU) No 1308/2013 establishing a common organisation of the markets in agricultural products and (EU) No 652/2014 laying down provisions for the management of expenditure relating to the food chain, animal health and animal welfare, and relating to plant health and plant reproductive material. *Official Journal of the European Union* L 350, 29.12.2017: 15-49.

European Commission (EC). 2020. Commission Implementing Regulation (EU) 2020/465 of 30 March 2020 on emergency measures in support of fruit and vegetables producer organisations in the Italian regions of Emilia Romagna, Veneto, Trentino Alto-Adige, Lombardia, Piemonte and Friuli Venezia Giulia in view of the damage caused to their production by the Asian brown marmorated stink bug (*Halyomorpha halys*). *Official Journal of the European Union* L 98, 31.3.2020: 26-29.

Farrin, K., M.J. Miranda and E. O'Donoghue. 2016. How do time and money affect agricultural insurance uptake? A new approach to farm risk management analysis. USDA Economic Research Report No. 40. Available at: https://www.ers.usda.gov/webdocs

Friedman, J., T. Hastie and R. Tibshirani. 2010. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software* 33(1), 1.

Giampietri, E., X. Yu and S. Trestini. 2020. The role of trust and perceived barriers on farmer's intention to adopt risk management tools. *Bio-based and Applied Economics* 9(1): 1-24. https://doi.org/10.13128/bae-8416

Graskemper, V., X. Yu and J.H. Feil. 2021. Farmer typology and implications for policy design – an unsupervised machine learning approach. *Land Use Policy* 103: 105328. https://doi.org/10.1016/j.landusepol.2021.105328

Graskemper, V., X. Yu and J.H. Feil. 2022. Values of farmers-evidence from Germany. *Journal of Rural Studies* 89:13-24. https://doi.org/10.1016/j.jrurstud.2021.11.005

Hastie, T., J. Qian and K. Tay. 2021. An introduction to glmnet. Available at: https://glmnet.stanford.edu/articles/glmnet.html

Haye, T., T. Gariepy, K. Hoelmer, J.P. Rossi, J.C. Streito, X. Tassus and N. Desneux. 2015. Range expansion of the invasive brown marmorated stinkbug, *Halyomorpha halys*: an increasing threat to field, fruit, and vegetable crops worldwide. *Journal of Pest Science* 88(4): 665-673. https://doi.org/10.1007/s10340-015-0670-2

Hellerstein, D., N. Higgins and J. Horowitz. 2013. The predictive power of risk preference measures for farming decisions. *European Review of Agricultural Economics* 40: 807-833. https://doi.org/10.1093/erae/jbs043

Iyer, P., M. Bozzola, S. Hirsch, M. Meraner and R. Finger. 2020. Measuring farmer risk preferences in Europe: a systematic review. *Journal of Agricultural Economics* 71(1): 3-26. https://doi.org/10.1111/1477-9552.12325

Knapp, L., Wuepper, D. and R. Finger. 2021. Preferences, personality, aspirations, and farmer behavior. *Agricultural Economics* 52(6): 901-913. https://doi.org/10.1111/agec.12669

Lee, J.C., A.J. Dreves, A.M. Cave, S. Kawai, R. Isaacs, J.C. Miller, S. van Timmeren, S. and D.J. Bruck. 2015. Infestation of wild and ornamental noncrop fruits by *Drosophila suzukii* (Diptera: Drosophilidae). *Annals of the Entomological Society of America* 108(2): 117-129. https://doi.org/10.1093/aesa/sau014

Lesmeister, C. 2015. *Mastering machine learning with R: master machine learning techniques with R to deliver insights for complex projects* (1st ed.). Packt Publishing Ltd.

Liesivaara, P. and S. Myyrä. 2017. The demand for public-private crop insurance and government disaster relief. *Journal of Policy Modeling* 39(1): 19-34. https://doi.org/10.1016/j.jpolmod.2016.12.001

MacKenzie, D.I., J.D. Nichols, J.A. Royle, K.H. Pollock, L.L. Bailey and J.E. Hines. 2018. Fundamental principals of statistical inference. In: D.I. MacKenzie, J.D. Nichols, J.A. Royle, K.H. Pollock, L.L. Bailey and J.E. Hines (eds.). Occupancy estimation and modeling (2nd ed.). Academic Press, London, UK, pp. 71-111. https://doi.org/10.1016/B978-0-12-407197-1.00004-1

Maruejols, L.L., L. Höschle and X. Yu. 2022. Vietnam between economic growth and ethnic divergence: a LASSO examination of income-mediated energy consumption. *Energy Economics* 114: 106222.

Menapace, L., G. Colson and R. Raffaelli. 2013. Risk aversion, subjective beliefs, and farmer risk management strategies. *American Journal of Agricultural Economics* 95(2): 384-389. https://doi.org/10.1093/ajae/aas107

Menapace, L., G. Colson and R. Raffaelli. 2016. A comparison of hypothetical risk attitude elicitation instruments for explaining farmer crop insurance purchases. *European Review of Agricultural Economics* 43(1): 113-135. https://doi.org/10.1093/erae/jbv013

Meraner, M. and R. Finger. 2019. Risk perceptions, preferences, and management strategies: evidence from a case study using German livestock farmers. *Journal of Risk Research* 22(1): 110-135. https://doi.org/10.1080/13669877.2017.1351476

Meuwissen, M.P., Y.D. Mey and M. van Asseldonk. 2018. Prospects for agricultural insurance in Europe. *Agricultural Finance Review* 78: 174-182. https://doi.org/10.1108/AFR-04-2018-093

Möhring N., S. Gaba and R. Finger. 2019. Quantity based indicators fail to identify extreme pesticide risks. *Science of the Total Environment* 646: 503-523. https://doi.org/10.1016/j.scitotenv.2018.07.287

Nielsen, T., A. Keil and M. Zeller. 2013. Assessing Farmers' risk preferences and their determinants in a marginal upland area of Vietnam: a comparison of multiple elicitation techniques. *Agricultural Economics* 44: 255-273. https://doi.org/10.1111/agec.12009

Ogurtsov, V.A., M.A. Van Asseldonk and R.B. Huirne. 2009. Purchase of catastrophe insurance by Dutch dairy and arable farmers. *Review of Agricultural Economics* 31(1): 143-162. https://doi.org/10.1111/j.1467-9353.2008.01430.x

Rotter, J.B. 1966. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied* 80(1): 1-28. https://doi.org/10.1037/h0092976

Santeramo, F.G. 2019. I learn, you learn, we gain. Experience in crop insurance markets. *Applied Economic Perspectives and Policy* 41(2): 284-304. https://doi.org/10.1093/aepp/ppy012

Scaccini, D., M. Falagiarda, F. Tortorici, I. Martinez-Sañudo, P. Tirello, Y. Reyes-Domínguez, A. Gallmetzer, L. Tavella, P. Zandigiacomo, C. Duso and A. Pozzebon. 2020. An insight into the role of *Trissolcus mitsukurii* as biological control agent of *Halyomorpha halys* in Northeastern Italy. *Insects* 11(5): 306. https://doi.org/10.3390/insects11050306

Slovic, P., B. Fischhoff and S. Lichtenstein. 1982. Why study risk perception? *Risk Analysis* 2: 83-93. https://doi.org/10.1111/j.1539-6924.1982.tb01369.x

Storm, H., K. Baylis and T. Heckelei. 2020. Machine learning in agricultural and applied economics. *European Review of Agricultural Economics* 47(3): 849-892. https://doi.org/10.1093/erae/jbz033

Sun F., D. Abler and X. Yu. 2021. Crop allocation and increasing returns to fertilizer use in China. *Land Economics* 97(2): 491-508. https://doi.org/10.3368/le.97.2.491

Trestini, S., E. Pomarici and E. Giampietri. 2017a. Around the economic sustainability of Italian viticulture: do farm strategies tackle income risks? *Quality – Access to Success* 18: 461-467.

Trestini, S., E. Giampietri and V. Boatto. 2017b. Toward the implementation of the income stabilization tool: an analysis of factors affecting the probability of farm income losses in Italy. *New Medit* 16(4): 24-30.

Trestini, S. and E. Giampietri. 2018. Re-adjusting risk management within the CAP: evidence on implementation of the income stabilisation tool in Italy. In: M. Wigier and A. Kowalski (2018) The common agricultural policy of the European union-the present and the future, EU member states point of view, no 73.1, IAFE-NRI, Warsaw, Poland.

Van Winsen, F., Y. de Mey, L. Lauwers, S. Van Passel, M. Vancauteren and E. Wauters. 2016. Determinants of risk behaviour: effects of perceived risks and risk attitude on farmer's adoption of risk management strategies. *Journal of Risk Research* 19(1): 56-78. https://doi.org/10.1080/13669877.2014.940597

Wang H., L. Maruejols and X. Yu. 2021. Predicting energy poverty with combinations of remote-sensing and socioeconomic survey data in India: evidence from machine learning. *Energy Economics* 102: 105510. https://doi.org/10.1016/j.eneco.2021.105510

Wąs, A. and P. Kobus. 2018. Factors differentiating the level of crop insurance at Polish farms. *Agricultural Finance Review* 78(2): 209-222. https://doi.org/10.1108/AFR-06-2017-0054

Wuepper, D. and L. Sauer. 2016. Explaining the performance of contract farming in Ghana: the role of self-efficacy and social capital. *Food Policy* 62: 11-27. https://doi.org/10.1016/j.foodpol.2016.05.003

Ye, T., Y. Liu, J. Wang, M. Wang, and P. Shi. 2017. Farmers' crop insurance perception and participation decisions: empirical evidence from Hunan, China. *Journal of Risk Research* 20(5): 664-677. https://doi.org/10.1080/13669877.2015.1118151

Zapponi, L., F. Tortorici, G. Anfora, S. Bardella, M. Bariselli, L. Benvenuto, I. Bernardinelli, A. Butturini, S. Caruso, R. Colla, E. Costi, P. Culatti, E. Di Bella, M. Falagiarda, L. Giovannini, T. Haye, L. Maistrello, G. Malossini, C. Marazzi, L. Marianelli, A. Mele, L. Michelon, S.T. Moraglio, A. Pozzebon, M. Preti, M. Salvetti, D. Scaccini, S. Schmidt, D. Szalatnay, P.F. Roversi, L. Tavella, M.G. Tommasini, G. Vaccari, P. Zandigiacomo and G. Sabbatini-Peverieri. 2021. Assessing the distribution of exotic egg parasitoids of *Halyomorpha halys* in Europe with a large-scale monitoring program. *Insects* 12(4): 316. https://doi.org/10.3390/insects12040316

Zhang, X and X. Yu. 2021. Short supply chain participation, and agrochemicals' use intensity and efficiency: evidence from vegetable farms in China. *China Agricultural Economic Review* 13(4): 721-735. https://doi.org/10.1108/CAER-05-2020-0108