Don't bet the Farm on Crop Insurance Subsidies

Marc Yeterian
marc.yeterian@dauphine.psl.eu

Céline Grislain-Letrémy
celine.grislain-letremy@banque-france.fr

Bertrand Villeneuve
bertrand.villeneuve@dauphine.psl.eu

February 14, 2024
Working paper

Abstract

Crop insurance is one of the most important tools that farmers have to protect themselves against climate-related risks. Yet and despite being heavily subsidized, insurance uptake in France remains extremely low. The goal of this paper is twofold ; first, we explain this paradox by analyzing the heterogeneous benefits and adverse effects of taking up crop insurance, and second, we provide concrete policy recommendations to increase insurance uptake in a welfare-maximizing way. Using an original micro-level panel of 17 000 French farmers over 20 years, we first use a moments-based regression to identify the local average effects (LATE) of insurance on expected revenues and variance, before investigating the factors that might cause heterogeneity in these effects, both observables through interaction terms and unobservables through a marginal treatment effect design. We conclude that insurance subsidies have very little impact on crop insurance demand, especially for those who would benefit the most, and suggest other less costly and more efficient ways to increase insurance uptake such as information campaigns.

Keywords: *Insurance; Agriculture; Marginal treatment effects; Instrumental variable.*
JEL Classification: *G22; Q10; Q12.*

1 Introduction

Agriculture is an important sector for the French economy, contributing 100 Bn EUR (net of insurance subsidies) to the national GDP in 2022 (about 3%) and employing 1.5% of the workforce (GÉRY, HECQUET, and LUCAS, 2023). France is the first agricultural producer in Europe and the fifth agricultural exporter in the world (VIE PUBLIQUE, 2022). However, the sector is declining ; the number of farmers has fallen rapidly in the past 40 years with their share in overall employment being almost divided by 5 (CHARDON, JAUNEAU, and VIDALENC, 2020).

Because climate shocks and risks have been steadily increasing for farmers around the world and in France (SHUKLA et al., 2019), risk management strategies have become a major focus of the agronomic literature, as evidenced by a meta-analysis by ANDERSON, BAYER, and EDWARDS (2020). While the literature mostly focuses on production technique, economic tools are often considered a second-best choice. We argue that multi-risk crop insurance, despite its low subscription rate in France, can be an important boon for farmers suffering from high risk exposure. We show the benefits of insurance subscription both through the average impact and the heterogeneous effects, and explain why, despite being heavily subsidized, crop insurance struggles to find its way into farmer's portfolios.

The French agricultural sector has been largely affected by climate change over the past decades. There have been numerous studies trying to quantify the impact of climate change on agricultural yields (both variability i.e. yield risk, and mean) with highly varying results. On a macro level for the EU, the cumulated effects of climate change (including human behavior and all ecological impacts) on productivity loss seems to lie in an order of magnitude between 10% and 15% (BOSELLO and ZHANG, 2005). These averages hide heterogeneous realities and massive variability. While globally, macroeconomic results show that wheat and corn production has lost 3.8% to 5.5% between 1990 and 2008 (MATHIAS et al., 2022), a more granular study in the Brittany region of France shows that climate change has increased corn yields by 0.12 tons of dry matter per Ha per year (+1.2% on average) via a shortening of the growth period (LIGNEAU et al., 2020). This has also led to a decrease in crop resistance to floods and droughts as a result of the more intensive harvest rate, which also accentuated revenue variability. In France specifically, GAMMANS, MEREL, and ORTIZ-BOBEA (2017) use a fixed effects model combined with the RCP8.5 (high emissions) scenario to show that with constant technology, overall productivity of wheat and barley will decline by 17%-33% by 2100. To our knowledge there have not been studies assessing the past overall impacts of climate change on agricultural yields in France,[1] but it is clear that the agricultural sector is declining and at least part of the trend can be explained by climate change. One of the factors contributing to this decline is the arduous and strenuous nature of the work (55 hours of work per week on average in France) which is going to get worse with climate change ; BRISSON (2010) identifies the increase in the number of hot days as one of the main factors for productivity losses in the agricultural sector, with workers being able to work less hours at a reduced productivity rate.

These impacts of climate change on agriculture have led the French Government to subsidize crop insurance schemes from 2005 onwards (KOENIG et al., 2022). These insurance subsidies have become an important part of both the French and European aid to agriculture, and have been integrated with the CAP in 2016 (MAA, 2022). Still, the general penetration rate of climate insurance is relatively low in France ; only 13.3% of farms were insured in 2020 (LOI DU 2 MARS 2022, 2022), despite insurance

[1] The Climator Project (BRISSON, 2010) assesses negative and positive channels on specific crops between 2007 and 2010 but not a global effect.

being highly subsidized and generally considered a net positive (DI FALCO et al., 2014). Indeed, the claim/premium ratio was around 200% for cereals and 156% for all crops in 2020 (MAA, 2022).

Providing insight into this paradox of low adoption is one of the main motivations behind this paper. We first analyze the impact of crop insurance on farmer's revenues to determine whether subscription is indeed a generally optimal choice, before explaining the low uptake through an analysis of the determinants of crop insurance and the heterogeneous nature of the benefits. Finally, we perform a counterfactual policy analysis to analyze the impact of potential additional insurance subsidies and other types of campaigns such as information or non-financial incentives.

This study is made possible by the use of granular pseudo-panel survey data on over 17 000 farmers in France for the period 2002-2020 (RICA, 2022). The Registre d'Information Comptable Agricole (RICA) is part of a European program (Farm Accountancy Data Network, FADN) that "monitors farms' income and business activities. It is also an important informative source for understanding the impact of the measures taken under the common agricultural policy" (*FADN* 2023). To our knowledge, this is the first time that this dataset has been used for such a study in France. This dataset is then matched to equally granular weather data over the full period (see section 6 for more details). For the first step, we use a parametric moments-based approach inspired by ANTLE (1983) and reused in DI FALCO et al. (2014) and WANG, REJESUS, and AGLASAN (2021) specifically for the purpose of analyzing the LATE of insurance on the revenue distribution. We refine the methodology used in these previous works by employing agronomic indicators for our weather variables, as well as proposing a new instrument to combat endogeneity, which would make our estimates less biased. For the second step, we perform a Probit regression on the probability to be insured, taking into account the dynamic nature of the market trough entry and exit variables in addition to a static insurance dummy. Third, we analyze the heteregeneous nature of the benefits first through an interacted regression, and then through a Marginal Treatment Effect framework. Finally, we use the marginal treatment effect framework to provide a counterfactual analysis of two policies : a 20% increase in insurance subsidies and an information campaign. The marginal treatment effect framework (HECKMAN and VYTLACIL, 2007) has never been used in this literature before, despite its benefits in providing a much finer analysis of the impacts of insurance than the classic average effect method.

We therefore contribute to the literature in four different ways :

- We go further than the methodologies used in the past by including a heterogeneity analysis on both observable and unobservable characteristics, which allows for precise targeting recommendation for policy. The MTE approach specifically has, to our knowledge, never been used in the context of crop insurance and provides insights on policy discussions regarding the right policies to maximize insurance uptake and social welfare.

- We develop a method to assess climate shocks impacts on agriculture more in-line with the agronomic literature. Using the concept of Growing Degree Days (LUO, 2011), we refine the methodology used in WANG, REJESUS, and AGLASAN (2021) to create aggregated indicators that capture the heterogeneous and non-linear effects of temperature on various types of crops. We specifically survey the agronomic literature to create three categories of crops according to their sensitivities. Using these into our regressions allows us to get more accurate and coherent results than the extremum method still currently used in most of the agricultural economics literature.

- We use a novel and granular dataset composed from individual data on farmers, including agronomic and financial variables, weather data at a 0.1° resolution and administrative data for climate disasters. Mixing these datasets, along with their precision, allows us to perform an analysis on a large scale (Mainland France for 20 years) while still remaining on the micro level. We appear to be the first to perform this kind of study on the French agricultural sector. As seen in the literature review, most other similar papers focus either on the Italian, US or developing markets. This allows us to not only confirm what has been found in previous research, but also to create an original and quantifiable result for policy purposes, especially with regards to the 2023 reform.

- We use an original continuous instrument on the micro level : the national subsidy rate by crop. This allows us both to circumvent the issues with instruments previously used in the literature and to use the previously mentioned MTE framework (which requires a continuous instrument). The continuous nature of the instrument also allows for more accurate results on the average impacts of insurance, as opposed to the limited external validity we would get by using specific reform years.

The paper is structured as follows : Section 2 provides context on the crop insurance market in France and a literature review of previous works on the topic, section 3 provides the empircal framework for the paper, section 4 discusses the estimation strategy, section 5 presents the data and summary statistics, section 6 provides and discusses the results, section 7 provides the counterfactual policy analysis, and section 8 concludes the paper and offers a policy discussion.

2 Literature review

2.1 Institutional context

Farmer's insurance options against climate risks in France are divided into two distinct regimes (MAA, 2022).[2] First, there is a fully public insurance scheme called the "Dispositif des Calamités Agricoles" (DCA) that insures farmers against losses caused by "exceptional" events, which include natural disasters. This pillar is maintained by a fund called the FNGRA co-financed by companies (one third) and the State (two thirds) specifically designed to compensate environmental and health risks that have been recognized on the national level. The payouts can be financed directly by the national fund is losses are small (less than 30% of annual production) or by the FEADER (European fund) is they are larger. According to BABUSIAUX (2000), this scheme is very limited and needs to be complemented with private insurance. First, the payouts are usually low (rate of compensation below 45%). Furthermore, there are threshold effects that discourage diversification and higher yields ; to be covered, a crop needs to constitute more than 13% of potential earnings (so farms with a high number of different crops are not covered), and the rate of compensation is calculated via the district average productivity, which means that the farmers with the highest productivity are be proportionally less compensated.

Second, the private subsidized insurance regime offers more customizable and diversified products (FOLUS et al., 2020). These include crop insurance on both the quantity and quality of crop loss, insurance against non-publicly covered climate risks (e.g. frost), or insurance against a loss in turnover below a guaranteed threshold. The crop insurance is publicly subsidized between 45% and 65% depending on the level of guarantee it offers (the higher the level, the lower the subsidy). Because of the high insurance subsidies, these contracts are also regulated and therefore relatively homogeneous. For example, they typically include a 20% deductible and cover all climate-related shocks that are not covered by the first pillar. Higher, less subsidized tiers, might include market insurance (i.e. protection against price drops or demand losses) or compensation for supply chain issues.

The French insurance system is similar to the ones found in other developed countries, which typically have a hybrid public-private system with varying degrees of public intervention. In the US the public insurance scheme takes the form of reinsurance scheme as a last resort (USDA, 2022), whereas in Italy the system is almost identical to France with public ex-post payments and a subsidized private sector (CAPITANIO et al., 2011). Finally, the market for agricultural insurance is closely linked to the Eu-

[2] For the studied period (2005-2021),which does not include the changes introduced by the 2022 reform.

ropean Common Agricultural Policy (CAP). Almost every farmer in France has received direct aids from the CAP scheme in the past year and these correspond on average to 88% of the revenues of farmers in France (CHATELLIER and GUYOMARD, 2020). These aids interact with insurance and farmer behaviors since they provide semi-stable revenues to farmers regardless of their yields (aids are mostly based on inputs rather than outputs) which limits the need for insurance (CHATELLIER, 2020).

2.2 Previous literature on the topic

This paper fits in the body of literature analyzing the empirical links between climate change, agricultural yields and risk management strategies (WALTHALL et al., 2013; VELANDIA et al., 2009). These studies establish a model of rational choice based on the impact of climate change on yield mean and variability and econometrically assess the impact of climate-related variables (temperatures, rainfall) and farm's characteristics (crop diversity, land) on the probability to opt into an insurance contract through probit regressions. Unsurprisingly, they find that farmers that have been hit with extreme weather events in the past tend to insure more. Furthermore, farms with the highest risks (and therefore highest potential claims) seem to be more insured, which would suggest that farmers indeed make a rational decision when choosing insurance.

On the other hand, the impacts of crop insurance have been scarcely studied, especially in Europe, but a few core papers have contributed to the research. The main one, and the base inspiration for this paper, is DI FALCO et al. (2014) which looks at the impacts of crop insurance on Italian farmer's revenues at the micro level through an instrumental variable approach, as well as the determinants of insurance adoption through a probit regression (which also constitutes the first step of the IV approach). On the effects of insurance, the paper finds that is has a positive impact on mean revenue, a negative impact on variance and a positive impact on skewness. In other words, this means that subsidized crop insurance both increases revenues and reduces risks for farmers, which makes it a very attractive option. On the determinants of insurance, the size of the farm and the value of inputs in the production function seem to increase the probability to be insured. Climate variables yield less coherent results with minimum temperature decreasing the probability to be insured, and maximum temperature increasing it. Since DI FALCO et al. (2014), other papers have refined the methods (ROLL, 2019; SANTERAMO et al., 2016; BLANC and SCHLENKER, 2017), the last of which is WANG, REJESUS, and AGLASAN (2021). The authors use interaction terms to determine whether crop insurance magnifies the effect of high temperatures on revenues, which would imply a moral hazard effect . While a moral hazard effect is found, the model is specifically applied on corn in the US at a county level, which distinguishes it from DI FALCO et al. (2014) and limits its external validity.

Finally, several other studies have used similar methods to assess the benefits of crop insurance, mainly in developing countries (BIRTHAL et al., 2022; ADDEY, JATOE, and KWADZO, 2021; FANG et al., 2021). The results from these studies unveil several interesting mechanisms on the indirect impacts of insurance. FANG et al. (2021) shows that, in China, crop insurance tends to increase total factor productivity, even when controlling for scale. This would suggest that insurance might encourage farmers to invest in more productive or intensive growing methods, creating a net positive impact even without taking into account the claims and premiums paid. BIRTHAL et al. (2022) performs a similar study in India and shows the heterogeneous nature of the benefits depend on farm characteristics including scale and exposure to climate shocks.

While these studies inform us on potential mechanisms and provide a baseline for our expected results, to our knowledge nothing of the sort has been done in France, which makes our study the first one to provide an assessment of the efficiency of crop insurance especially for French farmers. Furthermore, as outlined in the next subsection, previous studies appeared to have biases that we aim to correct.

3 Model

3.1 Baseline framework

We assume that a farmer's i utility function depends on both their average revenues (positively) and the variance of these revenues (negatively, i.e. farmers are risk-averse). In the traditional insurance economic literature and assuming perfect competition, on average buying insurance means balancing a trade-off between the mean and the variance of revenues. In other words, paying to decrease risk. Insurance is, by nature, a "losing" bet in perfect competition in the sense that average revenues are generally lower with insurance than without. This is a basic equilibrium condition to allow insurance companies to at least reach a 0-profit threshold. Frictions (such as administrative costs) are the main reason why insurance uptake is not a beneficial decision as far as average revenues are concerned. Because utility functions are generally comprised of both mean and variance of revenues, some farmers might be willing to sacrifice a bit of their mean to decrease their variance, hence the decision to insure. We model farmer's average revenues as (indices i are implied for simplicity)

$$E(R) = F(\boldsymbol{X}, \boldsymbol{B}) - L(N, \boldsymbol{B}, \boldsymbol{X}) + D[L(N, \boldsymbol{B}, \boldsymbol{X})] - D[q(N, \boldsymbol{B}, \boldsymbol{X}, \lambda)] \tag{1}$$

And the variance as

$$V(R) = (1 - D)[\sum_{t=1}^{n} E(L(N, \boldsymbol{B}, \boldsymbol{X})) - L_t(N, \boldsymbol{B}, \boldsymbol{X})]^2 \tag{2}$$

With $D \in [0:1]$ the insurance decision, F a production function depending positively on a vector of inputs $\boldsymbol{X}$ and negatively on a vector of protection behaviors $\boldsymbol{B}$, L a loss shock function based on a random variable N for the probability of occurrence of a climate chock, negatively on the protection behavior $\boldsymbol{B}$ and positively on the input of vectors $\boldsymbol{X}$, and n the number of periods. Finally, the farmer can decide to pay a price q to purchase a crop insurance contract with a level of protection D between 0 and 1. The price of the contract depends on the characteristics that the insurance company observes and an administrative cost λ. As stated above, excluding λ, in perfect competition, the function q should be equivalent to the function L, which means that the insurance company perfectly observe and predict risk. In that case, the impact of purchasing insurance on average revenues should be negative, as shown in the derivative of Equation (1) wrt. D :

$$\frac{\partial E(R)}{\partial D} = [L(N, \boldsymbol{B}, \boldsymbol{X})] - [q(N, \boldsymbol{B}, \boldsymbol{X}, 0)] \tag{3}$$

With $L(N, \boldsymbol{B}, \boldsymbol{X}) \leq q(N, \boldsymbol{B}, \boldsymbol{X}, \lambda)$ due to $\frac{\partial q(N, \boldsymbol{B}, \boldsymbol{X}, \lambda)}{\partial \lambda} \geq 0$ and $L(N, \boldsymbol{B}, \boldsymbol{X}) = q(N, \boldsymbol{B}, \boldsymbol{X}, 0)$ when $\lambda = 0$. Of course, the impact on variance is negative as well and the decision to purchase insurance depends on the form of the farmer's utility function, assuming it includes both variance and mean of revenues.

When insurance subsidies come into play though, they may change the parameters of the decision. Because insurance companies can now offer a lower price than what payouts actually cost them, farmers increase their average revenues while still decreasing variance. Formally, this means that we now have :

$$E^S(R) = F(\boldsymbol{K}, \boldsymbol{B}) - L(N, \boldsymbol{B}, \boldsymbol{X}) + D[L(N, \boldsymbol{B}, \boldsymbol{X})] - D[q(N, \boldsymbol{B}, \boldsymbol{X}, \lambda)] + DS[q(N, \boldsymbol{B}, \boldsymbol{X}, \lambda)] \tag{4}$$

$$\frac{\partial E^S(R)}{\partial D} = [L(N, \boldsymbol{B}, \boldsymbol{X})] - [q(N, \boldsymbol{B}, \boldsymbol{X}, \lambda)][1 - S] \tag{5}$$

With S the level of insurance subsidies. This means that if S is high enough, unless a farmer exhibits a highly downwards risk-taking behavior (unlikely), insurance becomes an optimal choice, increasing their utility regardless of their behavior. We can also allow for imperfect information, with the price function not perfectly predicting the loss function. In that case, farmers may have an incentive to engage in moral hazard behaviors by reducing $\boldsymbol{B}$ when they are insured, which would increase L but not q. In order to estimate these behaviors, we can simply substract insurance subsidies on both sides to get

$$E^S(R) - DS[q(N, \boldsymbol{B}, \boldsymbol{X}, \lambda)] = F(\boldsymbol{K}, \boldsymbol{B}) - L(N, \boldsymbol{B}, \boldsymbol{X}) + D[L(N, \boldsymbol{B}, \boldsymbol{X})] - D[q(N, \boldsymbol{B}, \boldsymbol{X}, \lambda)] \tag{6}$$

$$\frac{\partial E^S(R) - DS[q(N, \boldsymbol{B}, \boldsymbol{X}, \lambda)]}{\partial D} = [L(N, \boldsymbol{B}, \boldsymbol{X})] - [q(N, \boldsymbol{B}, \boldsymbol{X}, \lambda)] \tag{7}$$

We get the same Equation as in (5), except that the left-hand side is now net of insurance subsidies. This means that if Equation (7) is still positive, then farmers are increasing their revenues in the non-subsidized frameworks, and there are behavioral implications. Importantly, we cannot distinguish whether these behavioral changes are due to the insurance companies under-estimating risks (with q being lower than L), or if insurance companies somehow impact the production function of farmer by, for instance, encouraging them to increase inputs in a way that both increases their revenues and reduces their risks. Distinguishing these is not a goal of this paper, as we are mainly interested in the overall impact of insurance on welfare, rather than the channels.

In other words, this mechanism means that we expect to find an average positive benefit of insurance for mean and a negative effect for variance. However, the reduction in variance would only happen if farmers maintain the same production techniques whether they are insured or not. Farmers might engage in moral hazard behaviors by increasing their risk if they know that they will be compensated in terms of loss. Additionally, they might increase their production while maintaining their risk, creating increased yields not directly linked to insurance benefits. By looking at the impact of insurance on revenues net of insurance subsidies, we might be able to capture this effect (at least on mean revenue).

3.2 Heterogeneous analysis

Equations (3) and (7) show us that the benefits of insurance uptake are heterogeneous based on farmer's characteristics. Taking the cross-derivative of (6) wrt. D and $\boldsymbol{X}$, we get :

$$\frac{\partial^2 E^S(R) - DS[q(N,\boldsymbol{B},\boldsymbol{X},\lambda)]}{\partial D \partial \boldsymbol{X}} = [\frac{\partial L(N,\boldsymbol{B},\boldsymbol{X})}{\partial \boldsymbol{X}}] - [\frac{\partial q(N,\boldsymbol{B},\boldsymbol{X},\lambda)}{\partial \boldsymbol{X}}] \tag{8}$$

The sign of Equation (8) is not determined in the general case, meaning that each farmer may experience different benefits of insurance uptake, with some even being negative, depending on the form of their loss function and the efficiency of the prediction of insurance companies. Even in the presence of insurance subsidies, this means that insurance uptake could actually have a negative impact on average revenues (and variance) in some cases, and this is something that we need to observe in order to inform policy. Specifically, we need to identify which variables in $\boldsymbol{X}$ may be responsible for those differences. Assuming that farmers are rational, we can identify those variables by looking at the determinants of insurance subscription. This would mean that farmers select into the treatment (insurance subscription) based on their potential gains. We then get

$$D^* = \beta \boldsymbol{X} + \boldsymbol{U} \tag{9}$$

With $\boldsymbol{U}$ the vector of unobserved characteristics that also impact insurance subscription and benefits. Estimating $\boldsymbol{X}$ through a Probit regression is easy given enough data, and once we know the variables contained in it, we can check their impact on insurance benefits through Equation (8). In addition to providing a verification, it informs us on whether farmers do indeed select into treatment based on observables characteristics.

3.3 MTE framework

Estimating $\boldsymbol{U}$ is also extremely important for policy and for this we borrow from the MTE literature. Following Heckman and Vytlacil (2007), Andresen (2018), Kamhöfer, Schmitz, and Westphal (2019) and others, we use a marginal treatment effect estimation framework to analyze the heterogeneous impacts of crop insurance subscription over non-observable characteristics. This approach complements the last two subsections by providing insights on selection into treatment and showing that, even controlling for variables in $\boldsymbol{X}$, insurance can have highly differentiated impacts on farmers. Formally, we start with a Roy model such that

$$R^1 = \boldsymbol{X}\beta_1 + \boldsymbol{U_1}$$

$$R^0 = \boldsymbol{X}\beta_0 + \boldsymbol{U_0}$$

$$D^* = Z'\partial - V, D = 1[D^* \geq 0 <=> Z'\partial \geq V] \tag{10}$$

With R the outcome (revenue) with and without insurance (1/0), $\boldsymbol{X}$ the same vector of observable characteristics, $\boldsymbol{U_1}, \boldsymbol{U_0}$ the unobservable characteristics , D^* the latent desire to take up insurance that depends on $\boldsymbol{Z}$ ($\boldsymbol{X}$ plus an instrument) and an unobserved V. We use the average subsidy rate over crops and year as an instrument, and discuss this choice in the next section.

The intuition of this model is that as the instrument part of $\boldsymbol{Z}$ increases, more farmers are going to take up insurance. At the marginal point and keeping $\boldsymbol{X}$ constant, these farmers needed the extra incentive to insure. In other words, when insurance subsidies are low, only the farmers that had a very low "resistance" to insurance are getting insured, and as insurance subsidies increase, more and more reluctant farmers start insuring as well. The MTE allows us to estimate the equivalent of a local average treatment effect (LATE) differentiated over that subsidy scale. Hence, from Equation (10), we apply a monotonous transformation to get $p(Z)$ (Propensity score as quantiles of Z) and U_D (Quantiles of V). Rearranging, we can define the MTE as

$$MTE(X, U_D) \equiv \frac{\partial E(R|\boldsymbol{X}, p)}{\partial p} \tag{11}$$

That is, the variation of the expected outcome of insurance given the observables, as the propensity score increases. Simply put, this is the marginal effect of the treatment (insurance participation) for every quantile of the subsidy rate. To understand this properly, it helps to think about selection into treatment ; assuming that the marginal effects of insurance were decreasing along U_D (negative slope of the MTE), this would mean that the farmers who need the most insurance subsidies to insure are the ones getting the least benefit out of it, which, in a rational choice model, would make sense. If the curve was increasing, this would mean that farmers have previous biases which prevent them from seeing the benefits.

Finally, we can estimate the MTE following the procedure laid out in KAMHÖFER, SCHMITZ, and WESTPHAL (2019), starting with a basic LATE framework and taking the derivative wrt. p.

$$E(R|X, p) = \boldsymbol{X}\beta_0 + \boldsymbol{X}(\beta_1 - \beta_0)p + K(p)$$

$$\frac{\partial E(R|X,p)}{\partial p} = \boldsymbol{X}(\beta_1 - \beta_0) + \frac{\partial K(p)}{\partial p} \tag{12}$$

Fixing $\boldsymbol{X}$ at mean value (meaning that $\beta_1 = \beta_0 = \beta$) we get a derivative in p that is constant in X. This allows for different levels of potential outcomes while still conditioning for X. We still allow for different effects depending on the initial levels of p

$$E(R|\boldsymbol{X}, p) = \boldsymbol{X}\beta + (\alpha_1 - \alpha_0)p + K(p) \tag{13}$$

With α_1, α_0 the intercepts. This estimation strategy does not allow us say anything about the variations of $\boldsymbol{X}$ over p, but this is not what we are interested in. The only potential bias that might arise if the effects of $\boldsymbol{X}$ vary over different subpopulations is a difference in the level of the MTE, but not their structure. Since the goal of this section is specifically to look at the structure of the MTE (as the levels have already been estimated with a different strategy), this sacrifice seems reasonable.

4 Estimation strategy

4.1 Average impact of insurance subscription on revenues

4.1.1 Specification

Following our framework and Di Falco et al. (2014), we first estimate an OLS regression of Equations (2), (4) and (6) by using a fixed effect model estimate revenues with and without insurance subsidies from insurance status (D), inputs and individual characteristics $\boldsymbol{X}$ and climate shocks $\boldsymbol{L}$, before using the error term to the 2nd power (moments-based approach) to estimate the effects of these same variables on variance. [3]

$$\begin{aligned} R_{it} &= \alpha + \beta_1 D_{it} + \beta_2 \boldsymbol{X_{it}} + \beta_3 \boldsymbol{\Lambda_{it}} + \beta_4 \boldsymbol{\Lambda_{it-1}} + \theta_i + \theta_t + \epsilon_{it} \\ \epsilon_{it}^2 &= \alpha' + \beta_{12} D_{it} + \beta_{22} \boldsymbol{X_{it}} + \beta_{32} \boldsymbol{\Lambda_{it}} + \beta_{42} \boldsymbol{\Lambda_{it-1}} + \theta_i' + \theta_t' + \epsilon_{it}' \end{aligned} \tag{14}$$

With R_{it} the revenue variable in log (EBITDA with or without insurance subsidies), D_{it} the decision to insure (binary), $\boldsymbol{X_{it}}$ the vector of input variables (including subsidies and dummies for cattle/greenhouses, see section 3 for a complete list), $\boldsymbol{\Lambda_{it}}$ the vector of climate variables (sum of out-of-

[3] For a demonstration of the moments-based approach, see Antle (1983).

bound hot and cold GDDs for the specific crops grown by the farm, floods and droughts), θ_i the farm fixed effects θ_t the time fixed effects, and ϵ_{it} and ϵ'_{it} the unconditional error terms. All variables except dummies are expressed in log.

As mentioned in more details in section 3.3., the inclusion of GDDs stems from the agronomic literature and allows us to identify more accurately the effect of climate on revenue. The production variables include costs for gas/oil, costs for crop protection products, water used for irrigation, total work hours, total surface of the farm and production subsidies. The choice to put all the variables in log[4] stems from the high heterogeneity of the sample (see descriptive statistics) for both the dependant and independant variables, along with the assumption of negative returns to scale on the Cobb-Douglas production function. Furthermore, this choice is in line with the rest of the literature (DI FALCO et al., 2014; WANG, REJESUS, and AGLASAN, 2021). However, to reassure the readers on the solidity of the results, we include in the Appendix two alternative specifications where the dependant variable is expressed in direct form or with the Inverse Hyperobolic Sine (IHS) function, which would eliminate any bias arising from the negative values. Finally, the inclusion of the lagged weather variable comes from the argument we were making in section 2.3. against the use of these same variables as instruments ; lagged weather variables are a determinant of both present revenue through the adaptation behavior of farmers and insurance status due to past experience influencing present choices, and should therefore be included as a control.

This OLS specification presents high risks of endogeneity, mainly due to the omitted variable bias ; it is entirely possible that farmers with a better business sense (therefore improving their revenue) subscribe more to crop insurance (which should also increase their revenues), which would cause an overestimation of the insurance coefficient. Beliefs and social environment, while partially taken into account through fixed effects, might also play a role in both revenue and insurance subscription.

Endogeneity remains one of the main threats that can occur when studying the behavior of farmers. The decision to insure might indeed be linked to unobservable variables such as the farmer's beliefs. If these variables are linked to revenue (belief in climate change might for instance inform both the farmer's insurance and growing strategies), biases can occur. To correct this issue, previous authors have used an instrumental variable approach inspired by ANGRIST and KRUEGER (2001). The choice of instruments is crucial, especially when it comes to the exclusion restriction (WOOLDRIDGE, 2010) which states that the instrument must be exogenous, i.e. not causally linked with the dependant variable other than through

[4]For negative values, we add to the entire sample the minimum value + 1 before taking the natural log.

the instrumented variable.

DI FALCO et al. (2014) use lagged weather variables as their instruments. The argument is that lagged weather variables do not affect revenue in the present year. However, this appears to be a strong hypothesis for two reasons. First, past weather events might drive present adaptation strategies due to persistent effects ; a farmer hit by a flood in year t-1 might have built a tarp to protect its crop. This tarp remains in year t and affects revenues as well as the insurance strategy. Secondly, weather is highly autocorrelated and weather effects are persistent beyond a year ; past weather events have an influence on current weather events, which themselves can effect revenue through other channels than insurance. These arguments are further explored in MELLON (2022).

WANG, REJESUS, and AGLASAN (2021) corrects these issues by using two sets of instruments ; policy changes and national subsidy rates. Taken together, both of these are perfect instruments in the sense that they do not affect farmer's revenues through any other channel than insurance. Discreet reforms on their own might be weak instruments for two reasons ; first, they only exploit a limited source of variation in the sample, and second, they are incompatible with year fixed effects due to colinearity, which can create other sources of biases.

We employ an instrumental variable strategy through an institutional source of variation. Following WANG, REJESUS, and AGLASAN (2021), CONNOR, REJESUS, and YASAR (2022), DELAY (2019) and GOODWIN, VANDEVEER, and DEAL (2004), among others, we use the evolution of the average subsidy rate for insurance for each type of crops as an instrument. Insurance subsidy rates are decided at the EU level every year (MASA, 2022) since 2015 (and at the French level before that) and are differentiated between each type of culture. Since we do not have access to the official documents for insurance subsidies before 2015, we compute the average subsidy rate from our sample using the share of insurance subsidies of total premiums paid for the insured farmers growing a specific culture every year, then applying that share to the uninsured farmers in the sample.[5] Our final specification for the average impact of insurance on revenues therefore takes the following form for expected revenues and variance with and without insurance subsidies[6]

[5]See variable construction for more details.

[6]We also provide a robustness test in the Appendix using only the official documents on a sub-sample post-2015.

$$
\begin{aligned}
D^*_{it} &= \alpha + \beta_1 E(S|t,c)_{ct} + \beta_2 \boldsymbol{X_{it}} + \beta_3 \boldsymbol{\Lambda_{it}} + \beta_4 \boldsymbol{\Lambda_{it-1}} + \theta_i + \theta_t + \epsilon_{it} \\
R_{it} &= \alpha^{'} + \beta_{12} D^*_{it} + \beta_{22} \boldsymbol{X_{it}} + \beta_{32} \boldsymbol{\Lambda_{it}} + \beta_{42} \boldsymbol{\Lambda_{it-1}} + \theta^{'}_i + \theta^{'}_t + \epsilon^{'}_{it} \\
\epsilon'^2_{it} &= \alpha^{''} + \beta_{13} D^*_{it} + \beta_{23} \boldsymbol{X_{it}} + \beta_{33} \boldsymbol{\Lambda_{it}} + \beta_{43} \boldsymbol{\Lambda_{it-1}} + \theta^{''}_i + \theta^{''}_t + \epsilon^{''}_{it}
\end{aligned} \tag{15}
$$

4.1.2 Instrument validity

We make the argument that average national subsidy rates over crops and year are a valid instrument, meaning they respect both the strong instrument clause and the exclusion restriction. For the first part, we include the first-stage estimates along with the F-test, which is widely regarded as a valid way of testing weak identification. The F-stat is extremely high (over 160), meaning that even the worries laid in ANDREWS, STOCK, and SUN (2018) should not matter in our case.

In addition, while there is no proper mathematical way of testing for the exclusion restriction. We argue that because it is a policy based instrument decided at the national level (EU after 2015) before the beginning of contract season, it is highly unlikely that it would affect revenue in any other way than through insurance uptake. One worry is that we do not know exactly how subsidy rates are determined, and if some farmers had a way of influencing the decision process, they might increase insurance subsidies for their specific crops, while having the most power due to their ability to influence rates. If that were the case, we could be capturing the impact of influence rather than insurance and incur an upward bias. However, because we are also using fixed effects (both farmer and year), this would only matter if the influence of specific crop farmers changed over the course of our sample period. If some farmers had always had high influence, this would not matter because we only capture the changes in influence within the period. Furthermore, assuming that influence rises and decreases randomly across the period, the biases incurred by those changes would cancel out. The possibility of a specific crop rising to power in the past ten years is still not completely out of the question, but we have no reason to believe it happened.

4.2 Heterogeneous analysis

4.2.1 Identifying sources of heterogeneity in expected utility

While opting into insurance is, on average, an optimal choice, this does not mean that it is true for every farmer in the distribution. It might be the case that benefits are highly concentrated on a subset of farmers, which would explain the low insurance subscription figures despite the high average effect. Formally, as explained in the empirical framework through Equation (9), this would mean that Equation (15) might yield different results according to a function of a vector of variables $\boldsymbol{X}$. One such variable might for example be the farm size.

The aim this subsection is to identify these variables. Assuming farmers are rational, we can use a deterministic approach to deduce that variables in $\boldsymbol{X}$ influence their choice of insurance. Looking into the determinants of insurance subscription therefore allows us to identify the main variables of interest. In other words, we can check whether the criteria that influence insurance decisions are the same criteria that influence insurance benefits, which can inform us on hidden costs or other unobserved criteria.

4.2.2 Specification

To estimate the variables of interest in $\boldsymbol{X}$, we employ a Probit regression with fixed effects, using the same production and weather variable than in the base framework, this time to assess their effect on the probability to subscribe to insurance. Taking into account the dynamic nature of the market, we also run the Probit regression on the probability to enter or exit the market. This gives us three different specifications

$$P(D_{it}=1) = \alpha + \beta_1 \boldsymbol{X_{it}} + \beta_2 \boldsymbol{\Lambda_{it}} + \beta_3 \boldsymbol{\Lambda_{it-n}} + \beta_4 R_{it} + \beta_6 E(q|c,t)_{ct} + \theta_t + \epsilon_{it} \tag{16}$$

$$P(D_{it}=1|D_{it-1}=0) = \alpha^{'} + \beta_{12} \boldsymbol{X_{it}} + \beta_{22} \boldsymbol{\Lambda_{it}} + \beta_{32} \boldsymbol{\Lambda_{it-n}} + \beta_{42} R_{it} + \beta_{52} E(q|c,t)_{ct} + \theta_t^{'} + \epsilon_{it}^{'} \tag{17}$$

$$P(D_{it}=0|D_{it-1}=1) = \alpha^{''} + \beta_{13} \boldsymbol{X_{it}} + \beta_{23} \boldsymbol{\Lambda_{it}} + \beta_{33} \boldsymbol{\Lambda_{it-n}} + \beta_{43} R_{it} + \beta_{53} E(q|c,t)_{ct} + \theta_t^{''} + \epsilon_{it}^{''} \tag{18}$$

With D_{it} the dummy for crop insurance. The production and climate variable are the same as in Equation (15). We additionally include R_{it} as the revenue variables (EBITDA and yields) and $E(q|c,t)_{ct}$ as the price of insurance in the region by crop.

We are less worried about endogeneity in the Probit regression, since we are specifically searching for observable determinants of insurance (i.e. this is a predictive model, not a causal inference). For example, it might be the case that size is a proxy for another unobserved variable, and therefore is not the true determinant of insurance. This would not change our assessment that size is an observable criteria which allows us to predict insurance subscription. The causal or proxy nature of these determinants is not relevant to the research question and does not necessitate the use of an IV.

4.2.3 Heterogeneous benefits of insurance

Using the main variables of $\boldsymbol{X}$ identified in the previous subsection, we can then run the same regression as in Equation (15) (using the IV) but this time with an interaction term between our variable of interest and insurance uptake. This gives us a regression that we can interpret through the cross-derivative given in Equation (8). If the coefficients of the interaction are positive, this means that increasing this variable (for example, farm size) also increases the benefit of being subscribed to insurance. Formally

$$\begin{aligned} D^*_{it} &= \alpha + \beta_1 E(S|t,c)_{ct} + \beta_2 \boldsymbol{X_{it}} + \beta_3 \boldsymbol{\Lambda_{it}} + \beta_4 \boldsymbol{\Lambda_{it-1}} + \theta_i + \theta_t + \epsilon_{it} \\ (D^*_{it} * \boldsymbol{X_{it}})^* &= \alpha + \beta_{12} E(S|t,c)_{ct} + \beta_{22} E(S|t,c)_{ct} * \boldsymbol{X_{it}} + \beta_{32} \boldsymbol{\Lambda_{it}} + \beta_{42} \boldsymbol{\Lambda_{it-1}} + \theta_i^{'} + \theta_t^{'} + \epsilon_{it}^{'} \\ R_{it} &= \alpha^{''} + \beta_{13} D^*_{it} + \beta_{23} (D^*_{it} * \boldsymbol{X_{it}})^* + \beta_{33} \boldsymbol{X_{it}} + \beta_{43} \boldsymbol{\Lambda_{it}} + \beta_{53} \boldsymbol{\Lambda_{it-1}} + \theta_i^{''} + \theta_t^{''} + \epsilon_{it}^{''} \\ \epsilon_{it}^{''2} &= \alpha^{''} + \beta_{14} D^*_{it} + \beta_{24} (D^*_{it} * \boldsymbol{X_{it}})^* + \beta_{34} \boldsymbol{X_{it}} + \beta_{44} \boldsymbol{\Lambda_{it}} + \beta_{54} \boldsymbol{\Lambda_{it-1}} + \theta_i^{'''} + \theta_t^{'''} + \epsilon_{it}^{'''} \end{aligned} \tag{19}$$

With $\boldsymbol{X_{it}}$ the variables of interest and β_{23} and β_{24} the coefficients we want to interpret. Notice that when $\boldsymbol{X}$ increases, the benefit from going to non-insured to insured also increases. Formally, we can derive the third line of Equation (19) wrt. D to get

$$\frac{\partial R_{it}}{\partial D} = \beta_{13} + \beta_{23} \boldsymbol{V_{it}} \tag{20}$$

Notice how the benefit of D depends on the sign of β_{23}. A negative β_{23} would mean that increasing the characteristics in actually decreases the benefits of insurance, whereas a positive sign would mean the opposite. This is the exact setup of the cross-derivative from Equation (8).

4.3 MTE framework

4.3.1 Estimation

We perform the regression from Equation (13) using first a parametric approach for $K(p)$ (quadratic form, as is standard in the literature), and second using a semi-parametric approach (LIV) laid out in

ANDRESEN (2018) and first explained in HECKMAN and VYTLACIL (2007) (note that, in order to respect the MTE exclusion restriction, the revenue variable is net of insurance subsidies). The included controls are the same as in subsection 4.1.

We interpret mainly the semi-parametric regression in the results but do not bootstrap the confidence interval due to computational limitations. However, there is no reason to believe that confidence intervals should vary greatly between parametric and semi-parametric methods, and we therefore interpret the intervals from the parametric graphics.

Finally, note from Figure 7 (appendix) that the common support is cut at the ends due to lack of data. As we discuss in section 5.3.2., this does not bias the results.

4.3.2 Hypotheses

In this section, we provide a discussion on the MTE hypotheses that need to be satisfied in order for the results to be unbiased. We argue that despite the treatment (insurance uptake) being highly endogenous, our instrument is strong enough to yield unbiased estimates.

The first hypothesis concerns the exclusion restriction, which is slightly more constrained than a regular IV. Specifically, as stated in BRAVE and WALSTRUM (2014), we need to satisfy the following conditions for the general Roy model

$$Cov(Z, U_0) = 0$$

$$Cov(Z, U_1) = 0$$

Where Z is the instrument, and U_0, U_1 are the treatment effects. This means that not only do we need the instrument to only affect our outcome through the instrumented variable (standard exclusion restriction), but the instrument should also be completely uncorrelated with the treatment conditional to treatment adoption. This is why we estimate the MTE on revenue net of insurance subsidies, because otherwise our instrument (the average insurance subsidy rate over year and crop) would not satisfy this condition.

We therefore argue that average insurance subsidy rate is not correlated to revenue net of insurance subsidies either through the treatment effect or any other channel. One objection might be that these insurance subsidies could be invested by the farmer, which would indirectly increase revenue even net of insurance subsidies. However, we note that these insurance subsidies are typically given to farmers at the end of the growing season (MASA, 2022), whereas insurance is paid at the beginning or on a month-by-month basis. This means that, unless some strong anticipation behaviors are occurring, farmers could not invest the insurance subsidies to increase their production over year t. Furthermore, the subsidy bases are recalculated every year at the national level (before 2015) and European level (2016 and onwards), which means that farmers could hardly anticipate or influence this choice. This assumption also allows us to satisfy the "as good as random" assumption, i.e. $U_0, U_1, V \perp Z | \boldsymbol{X}$.

The second assumption is, as argued in HECKMAN and VYTLACIL (2007), that the propensity score needs to be as continuous as possible in order to estimate the whole range of treatment effects. In other words, we need the average subsidy rate over year and crop to be sufficiently heterogeneous to allow proper identification. We provide common supports for our MTE estimations in the Appendix that show that $P(Z)$ is indeed estimated on the (almost) full spectrum with enough variability. The analysis of takers on the common support also shows that monotonicity is achieved as there is no reversal of the order.

Furthermore, our instrument takes over 100 different values, varying from 0 for some crops (namely vines) up to 45% for other depending on crop and year. While true continuity is therefore not achieved, HECKMAN and VYTLACIL (2007) shows that the MTE curve can still be estimated over the common support. The results we provide in the next section therefore "fill in the gaps". The only assumption we need to make here is that there are no jumps (violation of monotonicity) between our point estimates, which we argue is fairly weak. It would not make sense for the propensity score to suddenly jump between two point estimates of close subsidy rates.

5 Data

5.1 Data sources

The data used in this paper comes from several distinct sources which have been matched together using the geolocalisation of the farms. The next subsection describes each source and gives an overview of the steps taken (if any) to transform the data for econometric analysis.

5.1.1 Farm level data

The financial, input and insurance data of farms comes from the French "Réseau d'Information Comptable Agricole" (RICA, 2022). This is a survey-based pseudo-panel dataset containing 17 743 individual firms observed over the 2002-2021 period. The RICA database is produced and managed directly by the French Government and constitutes the primary source for this paper. While it is a national dataset, it is part of the FADN network at the European level, and similar data sets, like the one used by DI FALCO et al. (2014), exist in other countries such as Italy and Germany. [7]

5.1.2 Climate data

Climate data comes from two distinct sources :

First, we use meteorological data provided by the National Meteorological and Hydrological Services from EU countries and aggregated by Copernicus, a European Union program dedicated to observing the Earth's climate (BOOGAARD et al., 2022). The dataset contains observations of temperatures and precipitations every six hours in France with a precision of 0.1° latitude/longitude for the 1950-2022 period. The data comes from 82 institutions and 22 600 weather stations (represented in Figure 2). Considering the station density in France is not large enough to collect all the data at such a granular level, a prediction model (reanalysis) is employed by Copernicus to fill the missing data (the detailed process can be found at BOOGAARD et al. (2022) in the dataset documentation).

The large nature of the dataset means that we had to extract the specific longitude/latitude and time ranges using Python before being able to fit it for econometric analysis. The GDD and precipitation indicators have first been computed directly for each coordinates before being extracted. Once the extraction was done, we matched these indicators to the RICA database using the "Base des Codes Postaux" (postal codes data) from the French Government (*Communes de france - Base des codes postaux* 2020) which provides the latitude/longitude coordinates for the center of each French city. Using a least squares method, we therefore matched each weather observation to a city, then matched the cities to the postal codes of each firm provided in the RICA database.

Second, the data on droughts and floods comes from the Caisse Centrale de Réassurance (CCR, Accessed in 2023), which is a public reinsurance company (100% owned by the French Government)

[7] Access to the RICA is restricted and confidential, and only summary-level data can be extracted and presented in this paper.

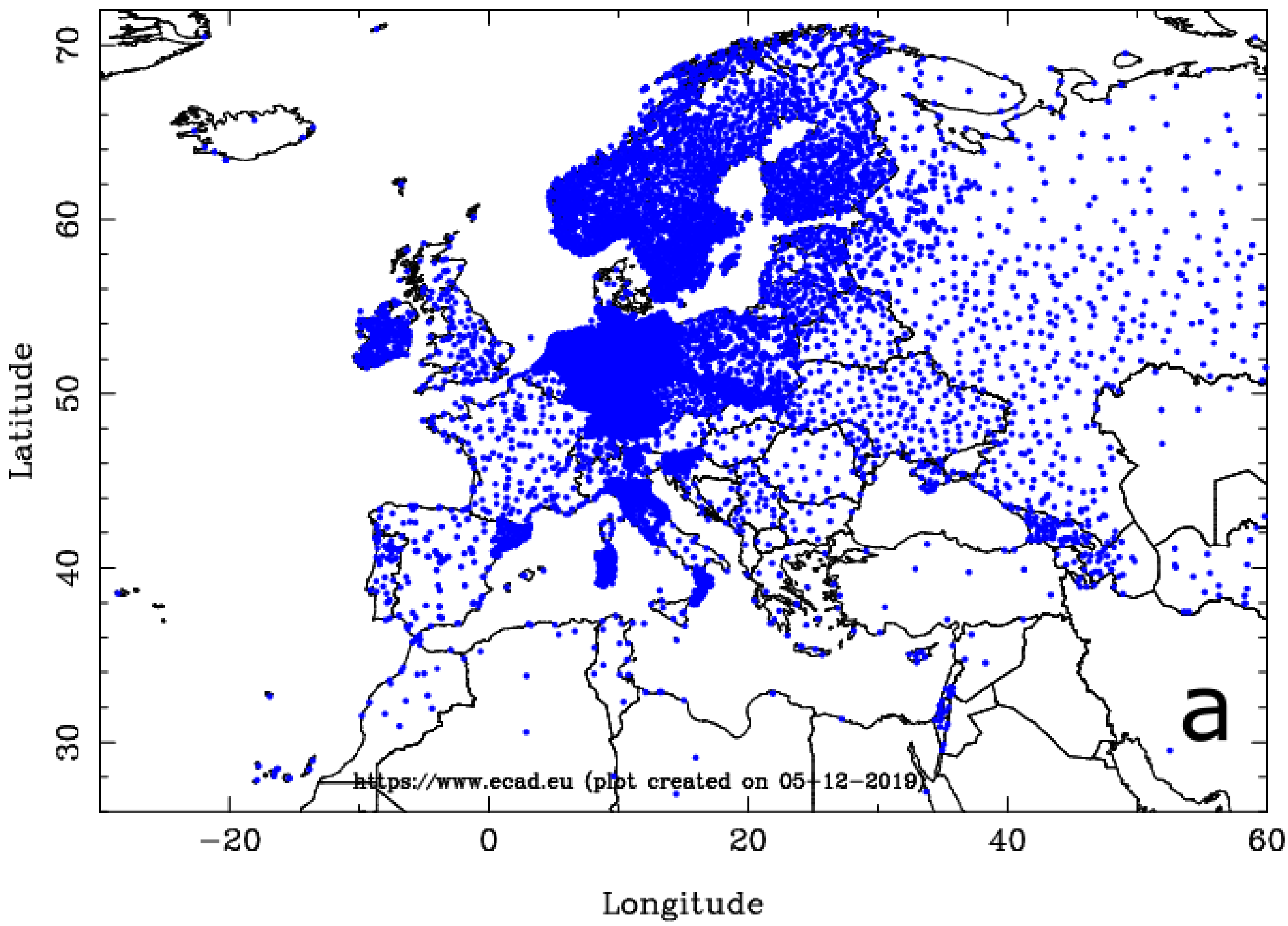

Figure 1: Weather station density across Europe for the Copernicus dataset. Source: Copernicus

that produces its own data on droughts and floods claims. These claims are filed by the mayor or prefect of the affected region, then collected and aggregated by the CCR in a database. Again, using the postal codes database, we match these claims to the RICA database.

5.1.3 Calamities data

The "Calamités Agricoles" (agricultural calamities) data comes directly from the claims filed by the mayors or prefects to the agricultural calamities regime (public insurance scheme). This data is aggregated but not treated by the Government, which means we only have access to the textual city-specific claims (see Appendix for an example). We search these claims for keywords (like "Apples") and match them to our list of crops available in the RICA. From there, we use the postal codes data to assign calamities claims to farmers. [8].

[8]This data is only used as a sub-sample robustness check in our regression, since we only have access to the claims from 2011 to 2020

5.2 Variable construction

5.2.1 Insurance variables

Consistent with DI FALCO et al. (2014), we use a dummy for crop insurance (1 if the farm is insured in a given year, 0 otherwise). We determine the status of insurance by using the premiums paid during the year. If high enough premiums were paid by the farm, we consider it insured.[9]

5.2.2 Weather variables

Because agriculture and weather are very finely linked, the choice of weather variables needs to be precise and has to rely on the agronomic literature. For example, the use of common statistical methods to identify temperature extremes (e.g. top 5% of temperatures or deviation from previous means) does not work in this context because the relationship between plant growth and temperature is not linear (SPINONI, VOGT, and BARBOSA, 2015; ANNAN and SCHLENKER, 2015). Furthermore, there are several types of crops which react very differently to temperature changes, e.g. maize resists to and even thrives in extreme hot temperatures, which is not the case of wheat (LUO, 2011). Not recognizing and including these effects in the regressions can lead to unobserved heterogeneity and an underidentification of the weather effects, such as in DI FALCO et al. (2014) which does not find a negative impact of cold temperatures on revenue.

Weather variables need to be accurate representations of the actual links between climate and agricultural yields. Consistent with the agronomic literature (LUO (2011), BLANC and SCHLENKER (2017), KORRES et al. (2016), HORTON (2018)), we therefore use the index of "Growing Degree Days" (GDDs) to compute extreme temperatures over a year. This index represents heat accumulation and captures th e non-linear effects of temperatures over the growing season. Plants do not grow if the mean temperature over a day (Tm is below a certain threshold (Tb) and slow their growth above certain upper limits (Tu). GDDs are an index of heat accumulation, and the plant needs a certain amount of accumulated heat to grow (GDD_{opt}). These thresholds and limits are different depending on the plant and can be found in the literature (cf. table 1).

[9] We ensure that there are no measurement error by only considering as insured the farmers paying more than 20 euros/Ha in insurance. This allows us to more accurately weed out those that might have either wrongfully answered the survey or that only insure an extremely small part of their production

Type of crops	Tb	Tu	GDDopt	Growing period in France	Sources
C3 Winter crops (Wheat, barley, oats, rye + Lettuce)	5.5 (0 in September to March)	30	1725	September to September	SPINONI, VOGT, and BARBOSA (2015) LUO (2011) RÖTTER and VAN DE GEIJN (1999) GRIGORIEVA, MATZARAKIS, and DE FREITAS (2010)
Potatoes and roots	8	26	1000	September to September	WORTHINGTON and HUTCHINSON (2005) LUO (2011)
C4 crops, fruits and vegetables (maize, rice, tomatoes)	10	32	1400	March to September	RÖTTER and VAN DE GEIJN (1999) LUO (2011) GRIGORIEVA, MATZARAKIS, and DE FREITAS (2010)

We first compute a GDD index over the year using the classic methodology, then use this index to derive our extreme temperatures. The base GDD index is calculated as follows

$$GDD_t = \frac{\sum GDD_i}{GDD_{opt}} \tag{21}$$

$$\text{If } Tm_i < Tb_i : GDD_i = 0$$

$$\text{If } Tb_i < Tm_i < Tu_i : GDD_i = Tm - Tb$$

$$\text{If } Tm_i > Tu_i : GDD_i = 2Tu - Tm$$

With t being the index for the year and i the index for the day.

From this index, we can compute our variable of interest for temperature : The sum of out-of-bound GDDs (OOB) for cold and hot temperatures, as defined in SCHLENKER, HANEMANN, and FISHER (2007). That is the sum of sum of the differences between the temperatures below (above) Tb (Tu), which represents the sum of cold (hot) GDDs received throughout the year by the crops. The index is then matched to each farm using the share of agricultural surface allocated to each crop. E.g. a farm growing 50% tomatoes and 50% wheat would receive 50% of the OOB for C4 crops and 50% of the OOB for C3 crops.

Finally, we also include two variables for floods and droughts. These are discreet variables that count the number of floods or droughts which have hit the farm in a given year. This data comes directly from the Caisse Centrale de Réassurance (CCR), which is the French public company for crop reinsurance (see section data sources for more information). There are no national definitions for floods and droughts, instead local officials (Mayors and prefets) can declare a state of flood or drought, which is then examined and confirmed by the State. We then match this data with our micro-level dataset. Note that this creates a limitation ; it might be possible that some regions declare floods more easily than others. However, we think that using local assessments still produces less biased results than a global approach that might not accurately capture the specificity of each region, especially in France where the climate can be highly heterogeneous.

5.2.3 Revenues

Our main variable of interest for revenue is annual EBITDA, which is a classic accounting variable that takes into account revenues from sales, subsidies, claims and stock variations, subtracting costs and insurance premiums. The advantage of using this instead of a more simple variable like operating profit is that EBIDTA tracks the farm's performance before any policy change (taxes), which yields a more accurate estimate of the impact of climate change and insurance on revenues (there is no reason to think that an extreme weather event would affect taxes or interests).

Additionally, we perform our regressions on the EBIDTA net of all insurance subsidies. By comparing the two indicators, we can estimate the raw welfare effect of insurance before any transfer from the State. This allows us to differentiate the overall benefit to the farmer (EBIDTA) and the actual insurance/production benefit (EBIDTA net of insurance subsidies). Furthermore, as we discuss in section 5, the MTE framework demands that we use the second indicator (net of insurance subsidies) to satisfy the exclusion restriction.

5.2.4 Instrumental variable

We use the national subsidy rate per crop and per year as an instrumental variable. While the official rate (45-65%) does not vary, the value of the insured premiums does (for instance potatoes might only be insured up to 40 euros/t one year, and 35 the next). This subsidy rate is decided at the European level, and therefore constitutes a good exogenous IV, as farmers have more interest to insure if the insurance subsidies are high, and insurance subsidies only affect farmer's revenues through their impact on insurance subscription. There is no reason to think that farmers could influence this rate at the European level. We can therefore safely assume that this national subsidy rate does satisfy the exclusion restriction.

The main difficulty lies in the fact that the official documents were only published since 2015. We can however estimate those rates for previous years by using the actual insurance subsidies paid to farmers as a ratio of their total insurance costs, and then apply this ratio to every farmer with a given crop for a given year. Formally

$$Srate_{ijt} = \frac{\sum_{i=1}^{n} sub_{ijt}}{\sum_{i=1}^{n} Pr_{ijt}} \tag{22}$$

With j the index for the crop, i the index for the farm, t the index for the year, $Srate$ the subsidy rate applied to farmer i, sub the actual subsidy received and Pr the premium paid.

To ensure that our measure is consistent with the official data, we perform a robustness test in the Appendix using only the official data on a reduced sample and find similar results (albeit less significant) than in our base framework.[10]

5.3 Summary statistics

5.3.1 Main statistics

Tables 2 and 3 show the summary statistics for all the variables used in the regressions[11]. As expected with firm data, the sample is highly heterogeneous, with some firms earning negative revenues in given years and others earning millions. The low EBITDAs can be explained by the cyclical nature of some agricultural productions (fallow for instance) and the heterogeneity in inputs (Std larger than or equal to the mean for all inputs) can also be attributed to the vastly different needs of the various crops represented in the sample ; organic agriculture uses very little phytosanitary products, while wheat in wet climates (North of France) might require little to no irrigation. On the other hand, tomatoes grown in greenhouses require a lot more inputs to grow. For gross production, stock variations are the main cause of negative values.

The insurance subscription rate is 26%, double what would be expected considering the national average of 13%. It is important to note that this 26% figure does not capture the share of insured farmers for a specific year, like the 13% does, but rather the cross-sectional mean over the whole sample. Still, this high figure can be explained by the fact that very small farms (who are typically under-insured) are under-represented in the sample. According to the French national statistics bureau (*Tableaux de*

[10] Note that a direct comparison between the subsidy rates from our data and the official data does not make sense since the way it is officially measured is through the base insured value per ton sold, whereas we measure overall subsidy rates.

[11] The hot and cold GDDs also exhibit a lot of variations, which is normal considering the various needs of the plants and the highly heterogeneous climate in France. Cold GDDs seem to be a lot more numerous than hot GDDs, which makes sense considering the sample is representative of agriculture in France, with the majority of farms being located in the North. Furthermore, most crops produced in French agriculture are more sensitive to colder temperature than hotter ones (e.g. winter wheat). As an example, winter wheat's upper bound is 30°C on average over a day, a temperature that is almost never reached in the north (BOOGAARD et al., 2022). Droughts and floods on the other hand appear to be fairly rare, but a mean of 0.06-0.08 signifies that, on average, every farm in our sample has experienced at least one flood/drought over the 18-year period.

	Mean	SD	Q1	Q2	Q3	Min	Max	Count
Dummy for crop insurance status (1=insured)	0.27	0.44	0.00	0.00	1.00	0.00	1.00	123700
Insurance spending per Ha (EUR/Ha)	24.22	55.91	0.00	2.32	22.81	0.00	450.00	123700
EBIDTA with insurance subsidies (KEUR)	85.70	87.45	35.93	64.18	110.31	-504.04	3755.93	123700
EBIDTA net of insurance subsidies (KEUR)	85.70	86.94	36.08	64.29	110.32	-504.04	3755.93	122039
Subsidy rate (year, culture)	8.40	9.38	0.00	6.34	15.51	0.00	46.58	123575
Sum of cold GDDs across the year (°C)	49.50	50.78	15.20	33.38	65.54	1.00	582.41	119940
Sum of hot GDDs across the year (°C)	1.06	0.46	1.00	1.00	1.00	1.00	63.79	119940
Number of floods/year	0.07	0.29	0.00	0.00	0.00	0.00	6.00	123700
Number of droughts/year	0.09	0.32	0.00	0.00	0.00	0.00	4.00	123700

Table 1: Summary statistics for the main variables

	Mean	SD	Q1	Q2	Q3	Min	Max	Count
Number of workers (work hours)	3922.07	4262.48	1600.00	3200.00	4600.00	45.00	216158.00	123700
Used agricultural surface (Ha)	104.21	81.40	46.20	85.42	141.50	0.32	795.49	123700
Diversification index (1=Not diversified)	0.48	0.28	0.25	0.46	0.67	0.00	1.00	123700
Subsidies received (EUR)	36949.61	30564.74	15750.69	30834.21	50784.93	0.00	1106312.00	123700
Cattle dummy	0.39	0.49	0.00	0.00	1.00	0.00	1.00	123353
Greenhouse dummy	0.02	0.15	0.00	0.00	0.00	0.00	1.00	123700
Organic agriculture dummy (1= at least partial)	0.52	0.82	0.00	0.00	1.00	0.00	5.00	123700
Real costs for gas/oil (EUR)	6744.66	6592.25	2519.90	4890.05	8835.00	0.00	172891.27	123700
Real costs for pesticides/Fertilizers (EUR)	12312.19	14809.97	2693.80	7426.96	16614.63	0.00	311599.00	123700
Agrotourism revenues (EUR)	77.58	1292.50	0.00	0.00	0.00	0.00	147940.00	123700

Table 2: Summary statistics for the control variables

l'économie française 2020), 31% of farms had yields lower than 25 000 EUR/year in 2016; whereas the first quartile of our sample is 89 000 EUR/year. Note that the mean of both revenues and yields is representative of the French landscape (average yields of 202 000 EUR/year according to INSEE vs. 221 000 EUR in our sample), but the distribution in our sample is highly concentrated towards the middle compared to the real distribution of farms. While this might seem like a selection bias at first glance, we argue that since our estimation only focuses on the within estimator (fixed effects), the bias would only exist if the movement dynamics for small farms and the benefits of insurance subscription were significantly different than for the rest of the sample. Furthermore, this bias towards the center can also be explained by the fact that we only observe farmers in continental France and not the overseas territories (whereas the official figure of 13% does)[12].

[12] The mean of the cattle dummy (0.39) might also seem high, but it is important to keep in mind that this dummy is equal to 1 if even a small fraction of the production is dedicated to cattle farming. Most farms have at least a few animals for auto-consumption, which does not mean that their main activity lies in cattle. The dummy is mostly here to control for the stability that cattle can bring to a production in times of crisis, and as we will see, remains highly significant.

Status of insurance compared with the previous year	Frequency	Percent
Kept insurance	43794	47.88
Canceled insurance	3256	3.56
Opted into insurance	4050	4.43
Stayed uninsured	40361	44.13

Table 3: Distribution of movements within the full sample

5.3.2 Geographic distribution of insurance uptake

Figure 2 provides maps of insurance subscription rates (left) and probability to get it by a flood or drought in a given year over the entire time sample (right, 2002-2021). The lack of correlation between uptake and exposure is very apparent, with the largest uptake being by far the Ile-de-France (Paris) region (72%), despite having a low risk exposure (7%). The most exposed region (PACA, 26%) also has a below average uptake (18%). These maps show that the insurance decision is a far more complex task than just applying an optimization problem, and requires further investigation.

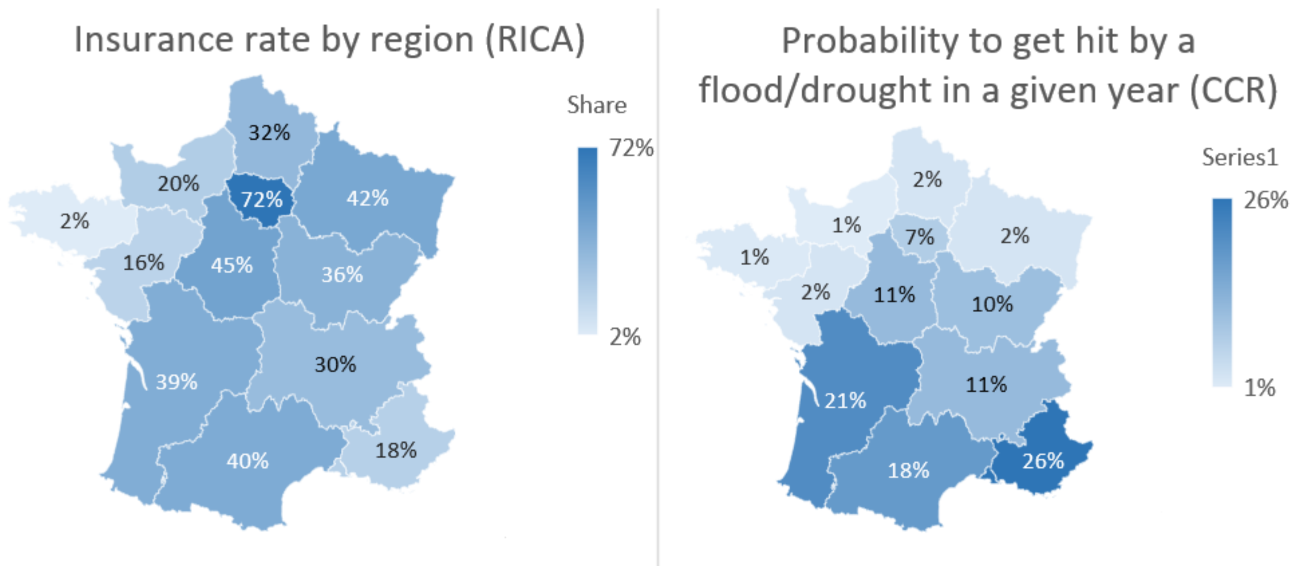

Figure 2: Map of insurance rate and risk exposure by region. Data source : RICA, Caisse Centrale de Reassurance ; Authors' Production

6 Results

6.1 LATE : Insurance increases revenues on average

Table 4 shows the results for the IV estimation (equation (15)),[13] that is the impact of crop insurance and weather variables on the EBITDA distribution with and without insurance subsidies. Columns 1 and 3 present the results for mean, while columns 2 and 4 show the results for variance.

The results confirm that crop insurance is indeed an optimal choice on average. Subscribing to crop insurance increases revenues on average by about 22% with insurance subsidies, and 19% without insurance subsidies. This means that we expect payments from crop insurance to be much higher than premiums, and therefore to increase mean revenues.

The comparable increases in revenue with and without insurance subsidies means that there are likely behavioral impacts to being insured. Assuming no behavioral change, we would expect revenues net of insurance insurance subsidies to remain the same with and without insurance, since on average premiums and claims need to be balanced, as outlined in the theoretical discussion. This means that insured farmers actually produce more than non-insured ones, probably because they have more incentive to invest in their fields when they know that they will be compensated should climate shocks arise.

This increased production also explains why the impact of insurance on variance appears to be non-significant. While insurance is supposed to reduce variance, if farmers cultivate riskier, higher-value crops when subscribing to insurance, mean revenues might increase while compensating the positive effect that insurance has on variance. We therefore have a double effect of insurance ; first, for the effect on revenues net of insurance subsidies, it creates an incentive for farmers to increase their (unobserved) inputs and effort, since the increased risk is covered by insurance. Second, it encourages farmers to adopt better protection behaviors to not pay the deductibles and might increase their general skill[14].

The size of the coefficients is high ; a 20% increase in revenues just for insurance might appear large, but it is consistent with both past literature (DI FALCO et al., 2014; WANG, REJESUS, and AGLASAN, 2021) and every specification of the model we have tried and presented in the robustness tests. Agricultural revenues are highly volatile and climate shocks in particular can easily destroy 60-100% of the produc-

[13]OLS and first-stage estimates are available in the Appendix.

[14]The effect of the weather variables on both revenue variables is also in line with the agronomic literature and our expected results. Cold temperatures appear to increase revenues, but only on the short-term, as an increase in past GDDs decreases revenues (-0.7% per GDD for the third lag). Indeed, in the short-term, cold GDDs are an indication for colder years, which in general feature less climate shocks. However, consistent colder years hurt crop growth. The effect of hot temperatures is also significant, at a much higher level ; -2% for the current year and about 0.9% for the second lag. Floods have a significant negative impact on EBITDA, while droughts are non-significant, which also makes sense since droughts can be compensated with irrigation, while floods can cause much larger damage.

tion, which would lead to a complete loss of revenues (save for insurance subsidies and the Calamités Agricoles scheme) for a farmer hit by such a shock without being insured. Because regressions only show average numbers, these very high losses that would primarily hit smaller farms factor into the coefficient in a high proportion, hence the large numbers (20%).

Furthermore, a comparison with the estimates obtains with OLS in the Appendix shows that the coefficient in the IV estimation are orders of magnitude larger (both specifications showcase strongly significant positive coefficients, but the OLS are around 0.4%). This confirms our intuition that the IV estimation corrects the average effect by eliminating the bias caused by variables in $\boldsymbol{U}$. The MTE estimation makes this very apparent and shows the high heterogeneity in insurance benefits based on the levels of $\boldsymbol{U}$ as defined in equation (9), i.e. the unobservable characteristics.

The comparison with the OLS estimates, the high coefficients and the limited level of actual insurance subscription all point to a high heterogeneity in the treatment effect of insurance. This is why stopping the analysis at the ATE (OLS)/LATE (IV) is not possible if we aim to understand how insurance actually impacts revenues. Investigating the heterogeneous impacts of insurance requires a more detailed look into both the determinants and the effects of insurance, which is the object of the next sections. Finally, this detailed analysis allows for actual policy recommendations based on who should insure.

	EBIDTA with insurance subsidies		EBIDTA without insurance subsidies	
	Mean	Variance	Mean	Variance
Dummy for crop insurance status (1=insured)	0.221***	-0.002	0.187***	0.002
	(0.028)	(0.008)	(0.026)	(0.007)
Cold GDDs (log)	0.006***	0.000	0.006***	0.000
	(0.001)	(0.000)	(0.001)	(0.000)
L.Cold GDDs (log)	-0.002	-0.001**	-0.002*	-0.001**
	(0.001)	(0.000)	(0.001)	(0.000)
L2.Cold GDDs (log)	0.003**	-0.001**	0.002**	-0.001**
	(0.001)	(0.001)	(0.001)	(0.000)
L3.Cold GDDs (log)	-0.007***	0.001	-0.006***	0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Hot GDDs (log)	-0.020***	-0.001	-0.020***	-0.000
	(0.004)	(0.001)	(0.004)	(0.001)
L.Hot GDDs (log)	-0.007	0.000	-0.006	0.000
	(0.004)	(0.001)	(0.004)	(0.001)
L2.Hot GDDs (log)	0.009***	-0.001	0.009***	-0.001
	(0.003)	(0.001)	(0.003)	(0.001)
L3.Hot GDDs (log)	0.005	-0.002**	0.002	-0.002**
	(0.004)	(0.001)	(0.004)	(0.001)
Number of floods (log)	-0.010***	-0.001	-0.010***	-0.000
	(0.002)	(0.001)	(0.002)	(0.001)
L.Number of floods (log)	-0.006**	0.001	-0.006**	0.001
	(0.002)	(0.001)	(0.002)	(0.001)
L2.Number of floods (log)	-0.002	0.000	-0.002	0.000
	(0.002)	(0.001)	(0.002)	(0.001)
L3.Number of floods (log)	-0.007***	-0.000	-0.007***	-0.000
	(0.002)	(0.001)	(0.002)	(0.001)
Number of droughts (log)	0.001	-0.001	0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.001)
L.Number of droughts (log)	0.004*	0.000	0.004*	-0.000
	(0.002)	(0.001)	(0.002)	(0.001)
L2.Number of droughts (log)	0.002	0.001	0.003	0.000
	(0.002)	(0.001)	(0.002)	(0.001)
L3.Number of droughts (log)	0.004	-0.000	0.004*	-0.000
	(0.002)	(0.001)	(0.002)	(0.001)
Subsidy rate (1st stage)	0.004***		0.004***	
	(0.000)		(0.000)	
Observations	69790	69790	69006	69006
Weak Ident.	168.984	168.984	180.817	180.817
Hansen J	0.000	0.000	0.000	0.000
Farmer FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Instrument	Yes	Yes	Yes	Yes

Table 4: 2nd stage IV log estimations for the impact of insurance on the revenue distribution

6.2 Heterogeneous analysis : Size appears to be the main determinant of insurance subscription, despite not increasing benefits

6.2.1 Probit estimation

Table 5 shows the results for Equations (16)-(18), that is the Probit regression for the determinants of static insurance subscription, entry and exit. Because we are using log transformed predictors in a Probit regression, the interpretation of the size of the coefficients in terms of absolute values is not straightforward (WOOLDRIDGE, 2010), unlike in a Logit model. We can, however, interpret the signs and compare the size of the coefficients between them as long as they are significant. For the dynamic models (columns 2 and 3), the signs of the coefficient can be interpreted as the effect of the variable on the probability to exit (enter) the insurance market in a given year.

The size of the farm, as measured by turnover [15], seems to be the main strong positive determinant of insurance subscription, both for the static and dynamic models. Unsurprisingly, having experienced weather events seems to be correlated with a higher probability of insurance. The lagged values appear to have a higher impact, which is expected, however the current values are also mostly significant. This would suggest that farmers anticipate weather fluctuations for the coming year and insure with this prior. Depending on whether this is purely an anticipatory effect or a substitution effect with protection behavior, we might suspect adverse selection or moral hazard.

Other negative factors include the diversity of the farm, which makes sense considering that crops can only be insured if they make up more than 13% of the total surface (the coefficient is positive, but the index reaches 1 when a farm is not diversified). Production subsidies also seem to have a negative impact, meaning that there might be a crowding out effect between insurance and subsidies. Finally, as we suspected, owning a greenhouse or raising cattle drastically decreases the probability to take up insurance, suggesting substitution behaviors.

[15] Turnover and Surface share a 0.4 correlation

	(1) Static	(2) Exit	(3) Entry
Turnover (log)	0.178***	-0.345***	0.064
	(0.058)	(0.086)	(0.072)
Total work hours (log)	0.004	-0.076***	-0.122***
	(0.020)	(0.018)	(0.017)
Total surface of the farm (log)	0.002	0.042**	0.097***
	(0.022)	(0.020)	(0.019)
Greenhouse dummy	-0.331***	0.080	0.173**
	(0.104)	(0.079)	(0.074)
Cattle dummy	-0.291***	-0.064***	-0.158***
	(0.021)	(0.023)	(0.022)
Organic agriculture dummy (1= at least partial)	-0.093***	0.044***	0.040***
	(0.012)	(0.010)	(0.010)
Real costs for gas/oil (log)	0.009*	-0.001	0.014*
	(0.005)	(0.007)	(0.008)
Real costs of crop protection products (log)	0.129***	0.039***	0.074***
	(0.015)	(0.009)	(0.013)
Debt (log)	0.023***	0.010	-0.004
	(0.007)	(0.009)	(0.008)
Rent (log)	0.022***	0.001	-0.007
	(0.006)	(0.006)	(0.006)
Diversification index (1=Not diversified)	0.942***	0.307***	0.185***
	(0.059)	(0.057)	(0.055)
Cold GDDs (log)	-0.061***	0.016	-0.002
	(0.006)	(0.015)	(0.015)
L.Cold GDDs (log)	-0.009*	-0.022	0.076***
	(0.005)	(0.014)	(0.014)
L2.Cold GDDs (log)	-0.009*	0.034**	0.027*
	(0.006)	(0.015)	(0.014)
L3.Cold GDDs (log)	0.051***	0.004	-0.006
	(0.006)	(0.015)	(0.014)
Hot GDDs (log)	0.143***	0.133**	-0.131**
	(0.028)	(0.063)	(0.059)
L.Hot GDDs (log)	0.183***	-0.061	-0.027
	(0.032)	(0.073)	(0.074)
L2.Hot GDDs (log)	-0.007	-0.033	-0.206***
	(0.025)	(0.054)	(0.058)
L3.Hot GDDs (log)	-0.126***	-0.024	-0.147**
	(0.031)	(0.063)	(0.066)
Number of floods (log)	0.073***	0.029	0.095**
	(0.021)	(0.050)	(0.046)
L.Number of floods (log)	0.090***	-0.032	0.045
	(0.022)	(0.052)	(0.047)
L2.Number of floods (log)	0.097***	-0.002	-0.027
	(0.021)	(0.050)	(0.049)
L3.Number of floods (log)	0.086***	-0.016	-0.011
	(0.022)	(0.051)	(0.050)
Number of droughts (log)	0.082***	0.066	0.103**
	(0.019)	(0.046)	(0.045)
L.Number of droughts (log)	0.096***	-0.002	0.026
	(0.020)	(0.050)	(0.048)
L2.Number of droughts (log)	0.050***	-0.024	-0.081*
	(0.019)	(0.048)	(0.046)
L3.Number of droughts (log)	0.042**	0.018	-0.023
	(0.021)	(0.048)	(0.046)
Subsidies received (log)	-0.045***	0.023***	0.006
	(0.004)	(0.008)	(0.007)
Constant	-3.320***	-0.304	-3.119***
	(0.361)	(0.483)	(0.398)
Observations	71524	71524	71524
Chi2	2136.854	314.277	766.611
Population average	Yes	Yes	Yes
Year FE	No	No	No

Table 5: Probit results : The determinants of insurance subscription

6.2.2 Heterogenous effects : The benefits of insurance do not increase with size

Table 6 shows the second-stage results for turnover (mean and variance) of Equation (19). The results seem highly surprising at first, as it seems that smaller farms benefit much more from insurance than larger farms, with the proportional (log) benefit being 7 times higher for farms in the first quartile compared to those in the last quartile.[16]

There are two ways to solve this paradox. On one hand, it might be the case that smaller farms only insure when the benefit is blatantly obvious, whereas larger farms take insurance as a given option and might need less incentive. In this case, it would mean that small farms lack either the information required or the necessary legal skills to insure until the benefits become too large to ignore. This interpretation would suggest that the true effect of insurance is in fact comparable for smaller and larger farms, with simply a selection into treatment bias. On the other hand, it is also possible that insurance subscription incurs higher up-front costs for smaller farms, due to a lower negotiation power or a lack of research on the market (i.e. information barriers). While the first option is difficult to estimate directly, we use descriptive analysis to check whether the second option applies.

To shed more light on the issue, we also perform the heterogeneous regression on quartiles of size and diversity to uncover potential non-linear treatment effects. The results of this regression can be found in Figures 3 and 4. While we still find that the benefits of insurance decrease with size, it appears that the middle of the sample benefits the least from insurance, while the extremes (smallest and largest) farms benefit the most. This fits with our first explanation if farms in the center of the distribution only insure "by default" whereas smaller farms insure when the benefits are highly obvious. Larger farms, on the other hand, derive small but consistent benefits, probably due to lower barriers to entry.

In terms of diversification, the effect appears more linear and in line with the Probit. That is, the least diversified farms are both the ones who insure the most, and the ones who benefit the most from insurance (selection into treatment). This makes sense as diversification can be seen as a protection strategy which would be a substitute for insurance uptake. Highly diversified farms also incur more costs as they need to insure each crop separately.

Policy-wise, this means that targeting smaller farms (with a high return on insurance but a low probability to insure) may be the best use of public funds. While low-diversity farms also benefit from insurance, they appear to already know that and have high subscription rates.

[16] Because we investigate the impact of farm size on insurance benefits, we cannot use turnover as an independent variable (since EBIDTA is the dependant variable and consists of net turnover including subsidies, taxes, etc.), and therefore use surface as a proxy (0.4 correlation coefficient), controlling for productivity (turnover/Ha) to ensure we capture a size effect.

	With insurance subsidies		Without insurance subsidies	
	(1)	(2)	(3)	(4)
Dummy for crop insurance status (1=insured)	2.093***	-0.296**	1.715***	-0.143
	(0.397)	(0.129)	(0.329)	(0.097)
Insurance status X Surface	-0.199***	0.030**	-0.163***	0.015
	(0.039)	(0.013)	(0.032)	(0.010)
Surface (log)	0.153***	-0.014***	0.142***	-0.007**
	(0.014)	(0.004)	(0.012)	(0.003)
Observations	69790	69790	69006	69006
Weak Ident.	26.376	26.376	30.783	30.783
Hansen J	0.000	0.000	0.000	0.000
Farmer FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Instrument	Yes	Yes	Yes	Yes

Table 6: IV estimations for turnover, with surface interactions

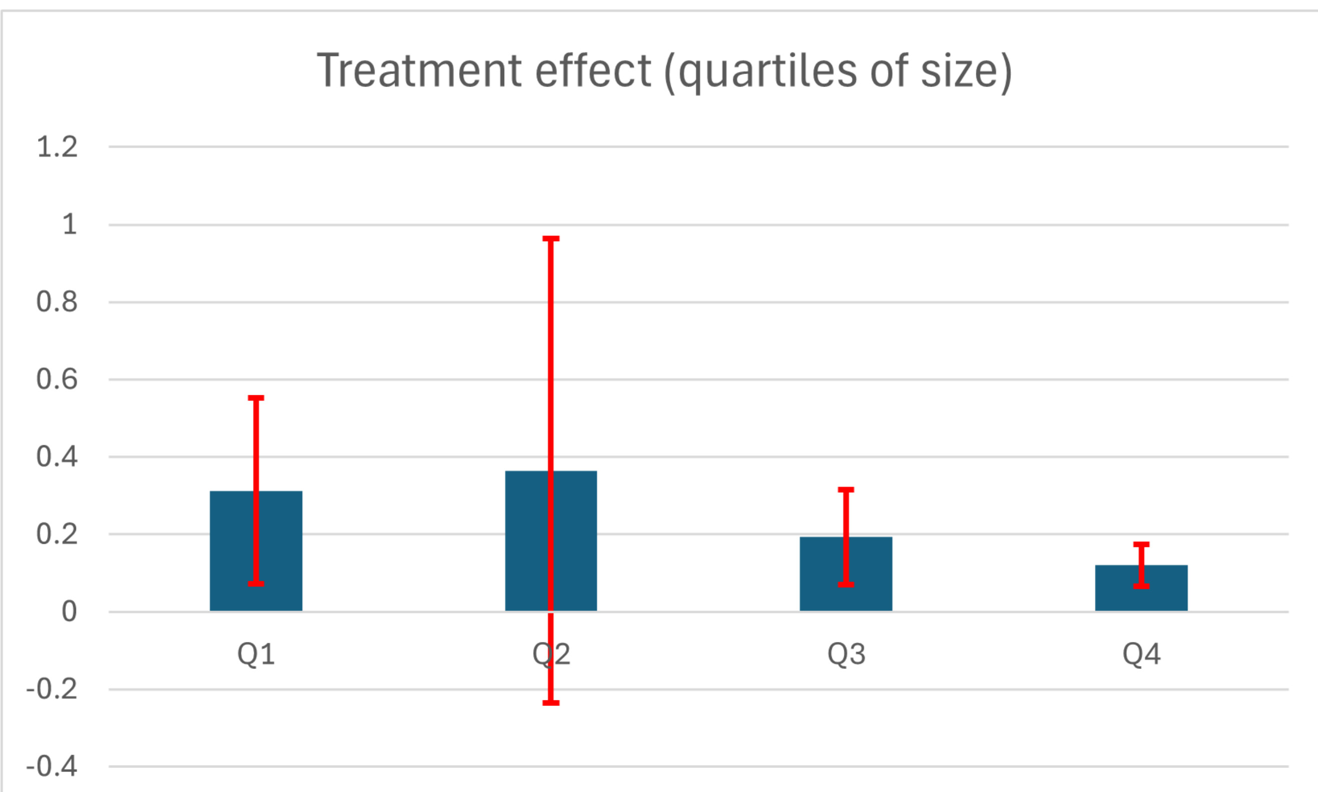

Figure 3: Treatment effect by quartile of size

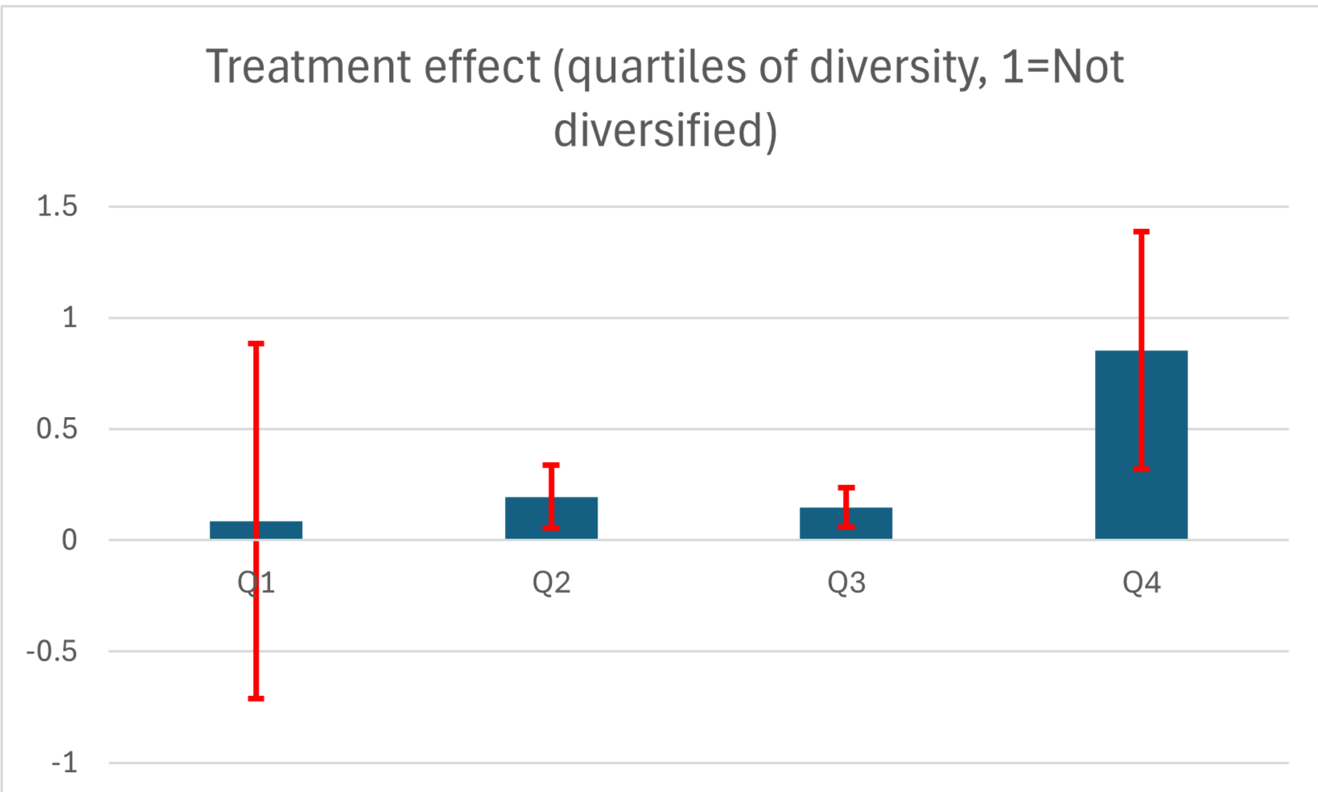

Figure 4: Treatment effect by quartile of diversity

6.3 MTE : Insurance benefits are highly heterogeneous and choices may be influenced by unobserved resistance to treatment

The highest benefits from insurance arise at the extremes of the resistance to treatment scale, while the middle still benefits, albeit to a much lesser extent, as suggested by the MTE results for the semi-parametric estimation which can be found in Figures 3 and 4. Results are only significant for the mean revenues and not the variance, which is coherent with the results from the IV regression. In other words, this means that, controlling for farm size and other parameters, farmers who are the most willing to subscribe to insurance and farmers who are the most resistant benefit the most, while farmers who are relatively indifferent benefit the least.

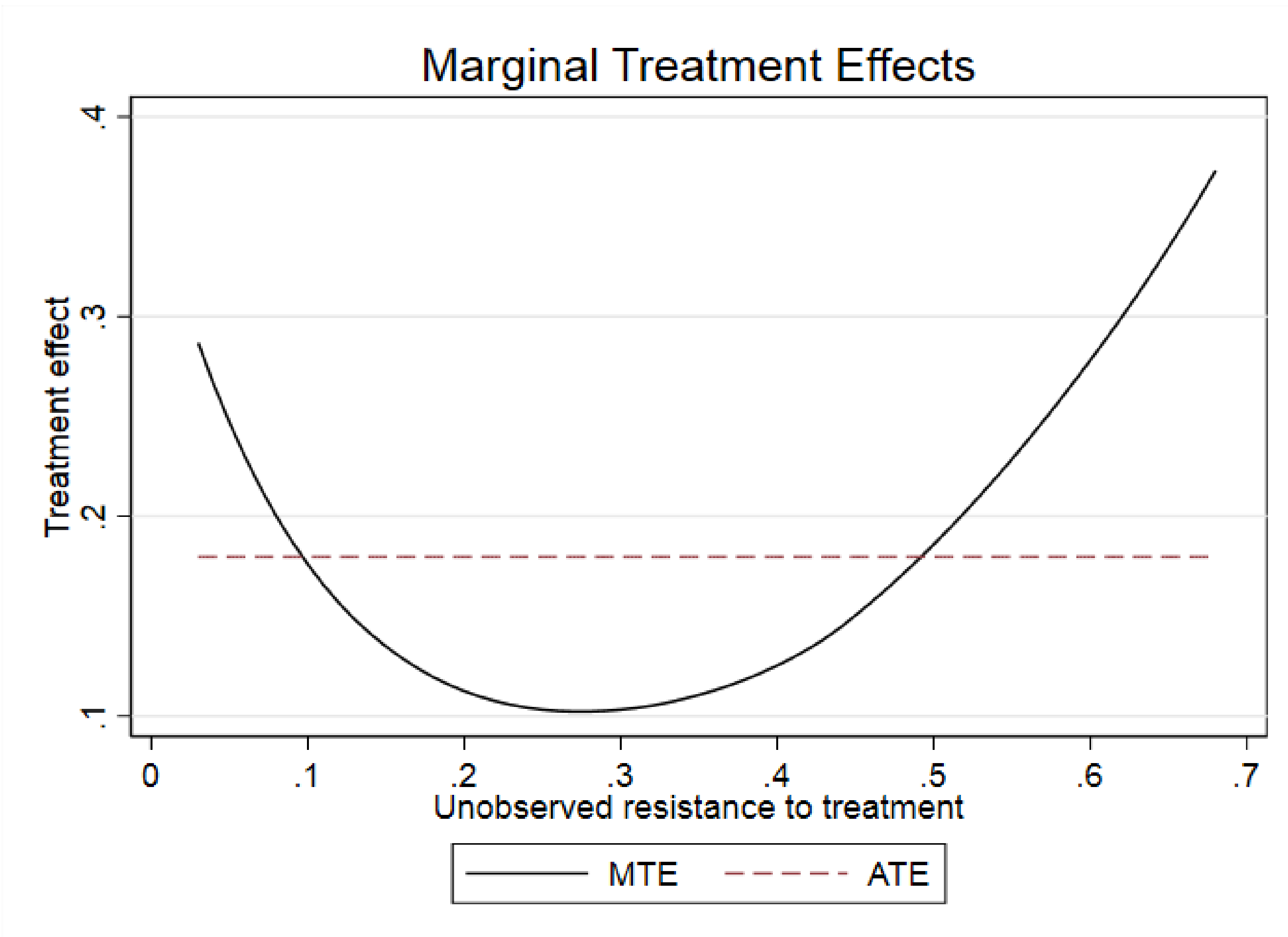

Figure 5: MTE curve for expected EBIDTA with controls

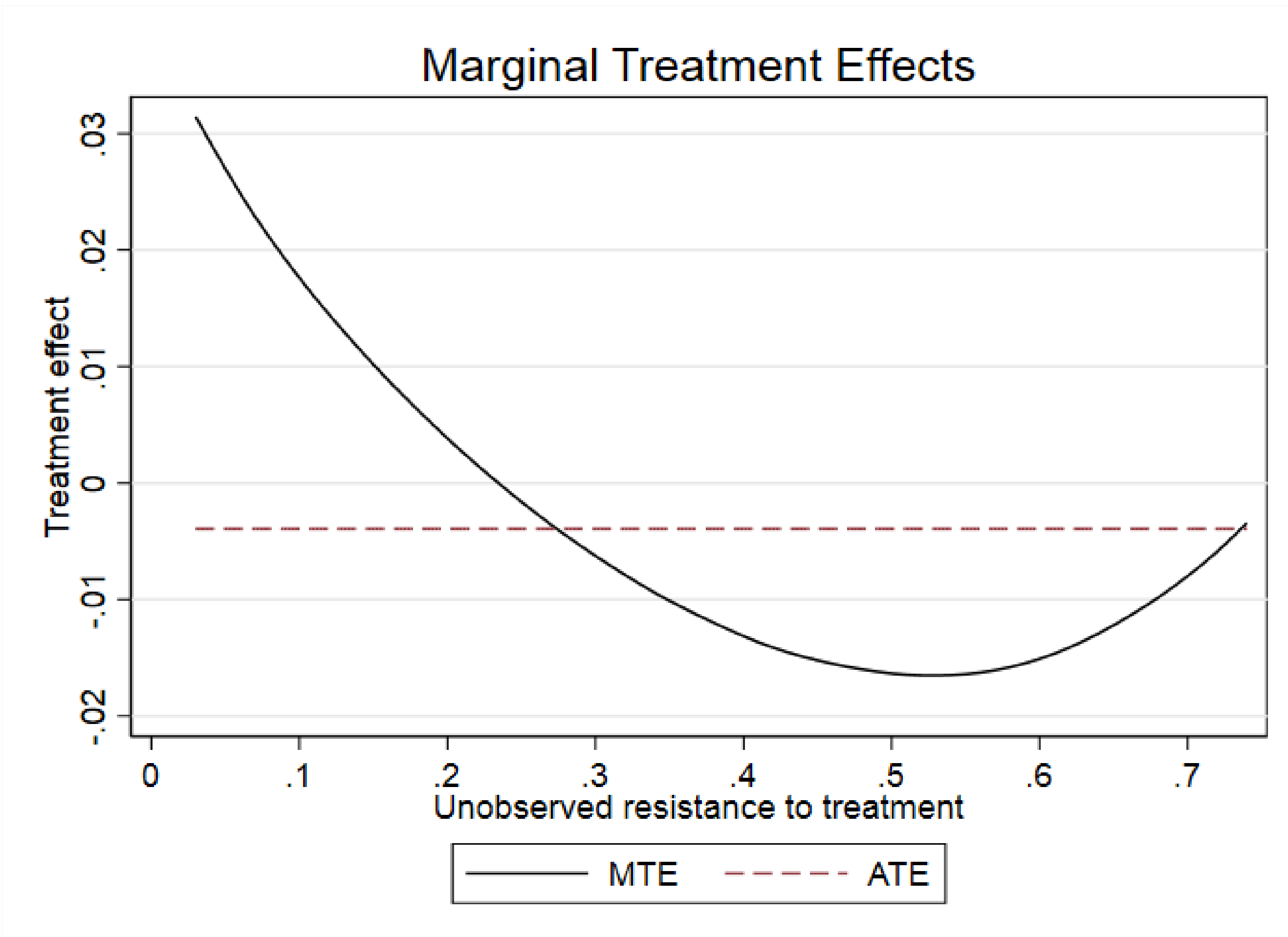

Figure 6: MTE curve for variance of EBIDTA with controls

While this is not a standard result, we propose the following explanation. On the one hand, farmers on the left-side of the distribution are "good managers" who are aware that insurance is an optimal choice and indeed mostly choose to take it. Farmers on the right side are "bad managers" ; risk-taking and probably less informed about the insurance subsidies. Finally, and more interestingly, farmers at the center are risk-averse ; they employ a variety of protection behaviors, which may or may not include insurance, to protect their crops. Those who choose insurance therefore only gain a marginally small benefit because their crops are already protected. This explanation fits with the heterogeneous analysis on size, as the mechanism may be the same ; the lowest benefits of insurance are found among the middle of the pack in terms of size.

7 Policy analysis

The results from the heterogeneous and MTE analysis show that while an increase in insurance uptake is generally desirable, it might not be efficient to increase insurance subsidies across the board, as a risk may arise ; the MTE analysis shows that increasing insurance subsidies may only affect those farmers with low resistance (who already massively subscribe to insurance, which would create an opportunity effect) or middling resistance who mostly should not insure. High resistance farmers are relatively unaffected by insurance subsidies, despite having the most benefit to be insured.

This fits with our heterogeneous analysis where we found that farms in the middle of the sample (for size) derived the least benefits from insurance, as they insure "by default". If we consider size as a proxy for management ability, both explanations reveal the importance of cost-benefit analysis when choosing insurance contracts, rather than doing it by default.

After we design the policy, we discuss the PRTE method and complement it with the MPRTE estimator to provide accurate results that can accomodate the lack of common support in the data.

7.1 Welfare impact of an increase in subsidy rate

7.1.1 Policy design

To emphasize that point, we perform a counterfactual analysis using policy relevant treatment effects (PRTE), using the methodology described in CARNEIRO, HECKMAN, and VYTLACIL (2011). Simply put, we perform the MTE analysis using an increased subsidy rate, which causes overall propensity scores to increase and an influx of individuals switching into insurance. The PRTE then measures the average marginal treatment effect at each point of new the propensity score distribution. Note that the shape of the MTE curve should not change between MTE and PRTE, as only the level of $K(p)$ is affected, not the derivative.

We perform the counterfactual analysis for an overall and equal 20% increase in the subsidy budget (note that this is a multiplication by 1.2 of the budget, not a 20pp increase of the rate). This figure is partly arbitrary, but we tried to strike a balance between a significant enough and a realistic enough policy change.

Table 10 shows the parameters and effects of this policy. Note that the average subsidy per farmer only increases by 11%, because while the budget is 20% larger, more farmers are sharing the insurance subsidies. The uptake rate, however, only increases by 6%, which suggests a very low elasticity of uptake to insurance subsidies. This is consistent with the interpretation that cost is not the main barrier to insurance subscription. Finally, note that because the increase is multiplicative, the standard deviation increases for the distribution of insurance subsidies.

	Mean	SD
Average subsidy per insured farmer (baseline)	645.687	1624.166
Average subsidy per insured farmer (20% increase)	720.621	1948.999
Uptake rate (baseline)	0.271	0.445
Uptake rate (20% increase)	0.288	0.157

Table 7: Parameters of the counterfactual policy

7.1.2 The full unit support hypothesis

Note that the proper interpretation of PRTE requires a full unit support which we do not have, as stated in CARNEIRO, HECKMAN, and VYTLACIL (2011). This is because it might be the case that the PRTE curve changes course outside of the common support, which would bias the PRTE estimator. In other words, we cannot verify the monotonocity assumption outside of the common support, and while this is not required for the base MTE, it is a condition for proper PRTE estimation.

However, we argue that while the average PRTE cannot be trusted, along with its exact level, the comparison between the PRTE and the MTE curve over the common support (in our case 0-0.7) is still relevant for policy analysis. Indeed in our case only the right tail of the common support is missing (i.e. the most resistant farmers), which means that an increase in the subsidy rate would likely have a very marginal impact, especially considering that, as shown in the previous subsection, the elasticity of insurance uptake to the subsidy rate is extremely low over the full available common support. It therefore stands to reason that it would be even lower in the right tail. In that sense our policy design can be approximated by the MPRTE, which does not require full support in order to be properly estimated.

All this means that the actual level of the PRTE probably cannot be interpreted, however the comparison between the PRTE and the MTE curve still provides valuable policy insights.

7.1.3 Results

Figure 5 shows the results of the counterfactual analysis. The PRTE appears to be lower than the MTE (as stated previously, the negative sign cannot be interpreted). This means that the newly insured farmers actually reap lower benefits from their subscription than those who were already insured. This fact, combined with the insubstantial increase in uptake for such a large budget increase, shows that a direct increase in insurance subsidies is not desirable from a welfare perspective.

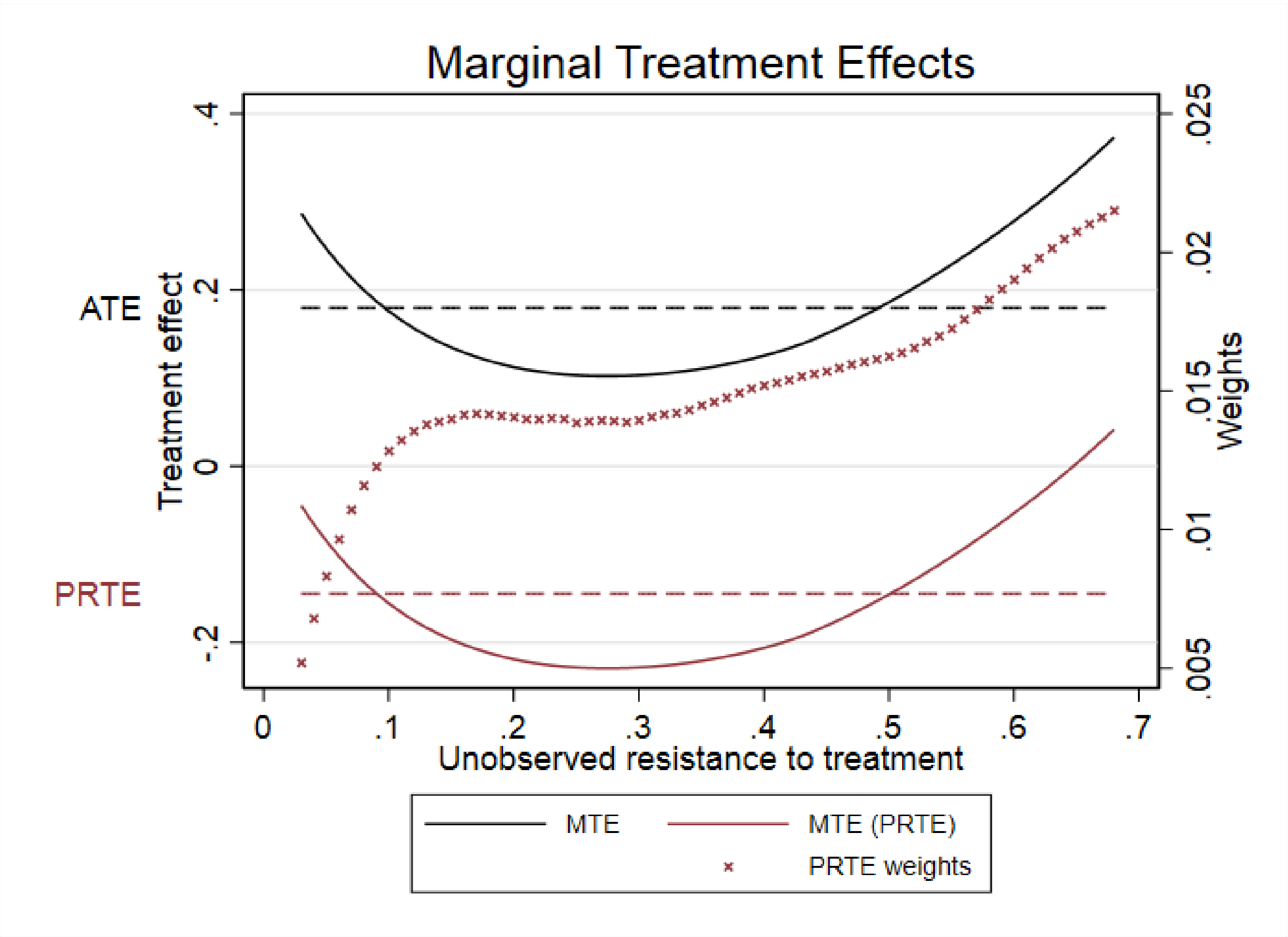

Figure 7: PRTE for 20% increase in subsidy rate

7.2 Impact of a marginal increase in the propensity score

7.2.1 Policy design

CARNEIRO, HECKMAN, and VYTLACIL (2011) define a method to provide accurate and relevant policy insights in the absence of common support ; the marginal policy relevant treatment effects (MPRTE), which corresponds to the average marginal benefit of uptake derived from a marginal increase in propensity score. Compared to the PRTE, this allows for an estimation of the impact of a policy that would target the propensity score directly without actually increasing the instrument.

In practice, this can be an information campaign on the benefits of insurance which targets the entire population, a speech at the national level, etc. These are "soft" policy which are likely to be less costly than an insurance subsidy increase.

Note that CARNEIRO, HECKMAN, and VYTLACIL (2011) provide three ways of computing the MPRTE. These three ways are mostly equivalent and differ in the weight composition. For transparency, we provide all three estimators. Finally, all three MPRTE estimators do not require additional hypotheses compared to the MTE. This is their main strength over the PRTE estimator, and this means that we can interpret them fully.

7.2.2 Results

The MPRTE estimators are very similar to the ATE, which means that a direct marginal increase in propensity score would likely result in a similar average effect for the newly subscribed as the past effect for the already subscribed. This means that, in practice, any policy that can decrease resistance to insurance uptake will be far more efficient than policies that target the subsidy rate.

	(1) EBIDTA net of insurance subsidies (log)	(2) Variance
effects		
ate	0.180	-0.004
att	0.169	0.015
atut	0.189	-0.013
late	0.101	0.007
mprte1	0.185	0.002
mprte2	0.163	-0.000
mprte3	0.182	-0.009
Observations	100329	70213

Table 8: MPRTE estimators (semiparametric)

8 Discussion and Conclusions

8.1 Results

Put together, the results of this study allow us to finely understand the determinants and impacts of crop insurance uptake in France, as well as the policy measures that may be applied to increase overall welfare. Using a highly detailed pseudo panel dataset, we are able to accurately identify and quantify the practical benefits of crop insurance for French farmers, that is an increase of mean revenue. We also shine a light on the paradox underlying the French insurance market ; despite being an arguably rational choice, crop insurance uptake is still surprisingly low. By identifying the size of the farm as one of the main determinants of crop insurance subscription, we can hypothesize the existence of both regulatory and financial barriers to entry in the market. We show that the effects of insurance are highly heterogeneous, which means that pursuing a goal of 100% of insured farmers might not be a feasible or desirable outcome, and that instead smaller, less diversified farms should be targeted. Finally, we use a counterfactual policy analysis to show that the level of insurance subsidies is not the issue causing the low insurance subscription, as increasing insurance subsidies would not cause a large increase in uptake, and those newly insured farmers would actually derive little benefit from their contracts. Instead, overcoming the non-financial barriers to insurance (i.e. information) by targeting the propensity score directly appears to be the optimal way of tackling this issue.

Our results are generally in line with the previous literature (WANG, REJESUS, and AGLASAN, 2021; DI FALCO et al., 2014; ANNAN and SCHLENKER, 2015). The results, however, are distinct in several regards ; first, we show that, in France, not everyone benefits from insurance. Larger farms and those engaged in other protection behaviors notably draw much smaller, if not negative impacts from their subscriptions. Second, we find lower and more realistic impacts across the board than DI FALCO et al. (2014), which might come from the fact that our IV is better suited to the analysis. Finally, we go further than other studies by providing concrete policy recommandations.

8.2 Policy implications

Policy-wise, this study comes at a time of reform for the French crop insurance market, with the 2022 reform still in the works at the time of writing. It is clear that there is a need to simplify the market to, at least partially, eliminate the barriers to entry that we have identified. Furthermore, there is a need to incentivize smaller farmers to take on an insurance subscription through price mechanisms and more adapted contracts, which the reform seems to do with the new tier system. Increasing insurance subsidies for every farm might not be the best way to improve welfare ; instead, figuring out the characteristics, such as size, that improve the benefits of insurance and targeting the farmers with the highest resistance to treatment may be more beneficial. One of the main messages of this study is that insurance, while an important tool, is not a one-size-fits-all remedy for climate shocks.

Specifically, we propose three legally feasible reform pathways to maximize the welfare impacts of crop insurance in the future

- First, insurance subsidies need to be targeted more towards smaller farms to ensure a higher takeup for those that need it the most. Rather than determining insurance bases by crop, we propose a continuous tier-based subsidy rate based on the surface of the farm, while keeping the overall insurance rate the same as today. For example, the first 20 Ha may benefit from a 90% subsidy rate, the next 100 from a 60% rate, etc.. Such a scale would be compatible with the 2022 reform by adapting it over the contract tier dimension. One objection to this might be the equality principle that prevents subsidy dicrimination between firms. However, as outlined by BARROIS DE SARIGNY (2020), this principle tolerates exceptions as long as *"it forsakes equality for reasons of general interest, if the treatment difference that results from it is in direct correlation with the goal of the established norm"*.[17]

[17] Translation by the authors of this paper. Original : il déroge à l'égalité pour des raisons d'intérêt général pourvu que, dans l'un comme l'autre cas, la différence de traitement qui en résulte soit en rapport direct avec l'objet de la norme qui l'établit

This means that a policy may disregard the equality principle if two conditions are met ; first, the reason needs to be the common good (here maximizing welfare), and second, the violation needs to be scientifically justified (as in this study) and actually achieve the goal. We argue that both these conditions would be met here, which makes our proposition politically feasible.

- Second, improving insurance uptake in a welfare-maximizing way means targeting the propensity score (i.e. resistance to insurance) directly, which translates into information campaigns targeting those farmers that would benefit the most from insurance (smaller, less diversified). This would not only be cheaper than increasing insurance subsidies, but it would also be far more efficient.
- Third, financial barriers to entry need to be minimized ; aside from informational issues, the timing of the insurance subsidies needs to be reviewed so that farmers don't have to pay in advance. While funding statistics are not available, it is safe to assume that many smaller farms in France live year to year with very little cash flow available. This measure would cost very little to the State (essentially the interest rates for the 8 months) but would drastically increase uptake.

8.3 Extensions and limits

While this study tries to provide a complete picture of the issue, further research and investigation will still be needed in two key areas. First, the analysis of the behavioral mechanisms underlying insurance impacts could benefit from a more theoretical approach such as the one outlined in WU, GOODWIN, and COBLE (2020) for prevented planting in the US. A more thorough analysis of specific input usage or farm-level decisions (such as the transition to organic agriculture) could also lead to an accurate depiction of these mechanisms. Second, network effects and the transferable beliefs of farmers are ignored in this study. While fixed effects and the instrumental variable approach takes care of most of the bias these could cause for the base framework, they would still be interesting to observe through survey or political data to provide a more accurate description of the determinants of insurance. KOENIG et al. (2022) does this through mainly qualitative data, but adapting this methodology to a large-scale study such as ours could vastly enrich the political discussion around crop insurance.

References

ADDEY, Kwame Asiam, John Baptist D. JATOE, and George Tsey-Mensah KWADZO (2021). "Adoption of crop insurance in Ghana: an application of the complementary log-log truncated Poisson double-hurdle model". *Agricultural Finance Review* 81, pp. 76–93.

ANDERSON, Robyn, Philipp E. BAYER, and David EDWARDS (2020). "Climate change and the need for agricultural adaptation". *Current opinion in plant biology* 56, pp. 197–202.

ANDRESEN, Martin Eckhoff (2018). "Exploring marginal treatment effects: Flexible estimation using Stata". *The Stata Journal* 18, pp. 118–158.

ANDREWS, Isaiah, James H. STOCK, and Liyang SUN (2018). "Weak instruments and what to do about them". *Lectures, NBER Summer Institute*.

ANGRIST, Joshua D. and Alan B. KRUEGER (2001). "Instrumental variables and the search for identification: From supply and demand to natural experiments". *Journal of Economic perspectives* 15, pp. 69–85.

ANNAN, Francis and Wolfram SCHLENKER (2015). "Federal crop insurance and the disincentive to adapt to extreme heat". *American Economic Review* 105, pp. 262–66.

ANTLE, John M. (1983). "Testing the stochastic structure of production: a flexible moment-based approach". *Journal of Business & Economic Statistics* 1, pp. 192–201.

BABUSIAUX, Chistian (2000). "L'assurance recolte et la protection contre les risques en agriculture". *Rapport pour le Ministere de l'Economie des Finances et de l'Industrie et le Ministere de l'Agriculture et de la Peche*.

BARROIS DE SARIGNY, Cécile (2020). "Le principe d'égalité dans la jurisprudence du Conseil constitutionnel et du Conseil d'État". *Titre VII* 4, pp. 18–25.

BIRTHAL, Pratap S., Jaweriah HAZRANA, Digvijay S. NEGI, and Ashok K. MISHRA (2022). "Assessing benefits of crop insurance vis-a-vis irrigation in Indian agriculture". *Food Policy* 112, p. 102348.

BLANC, Elodie and Wolfram SCHLENKER (2017). "The use of panel models in assessments of climate impacts on agriculture". *Review of Environmental Economics and Policy* 11, pp. 258–279.

BOOGAARD, H., J. SCHUBERT, A. DE WIT, J. LAZEBNIK, R. HUTJES, and G. VAN DER GRIJN (2022). *Agrometeorological indicators from 1979 to present derived from reanalysis, version 1.0.*

BOSELLO, Francesco and Jian ZHANG (2005). "Assessing climate change impacts: agriculture". *Fondazione Eni Enrico Mattei (FEEM)*.

BRAVE, Scott and Thomas WALSTRUM (2014). "Estimating marginal treatment effects using parametric and semiparametric methods". *The Stata Journal* 14, pp. 191–217.

BRISSON, Nadine (2010). "Changement climatique et cultures de mais et sorgho grains: l'essentiel des impacts". *Climator 2010*.

CAPITANIO, Fabian, Maria BIELZA DIAZ-CANEJA, Carlo CAFIERO, and Felice ADINOLFI (2011). "FAUX TITRE - Crop insurance and public intervention in the risk management in agriculture: does farmers really benefit?" *Applied Economics* 43, pp. 4149–4159.

CARNEIRO, Pedro, James J. HECKMAN, and Edward J. VYTLACIL (2011). "Estimating marginal returns to education". *American Economic Review* 101, pp. 2754–2781.

CCR (Accessed in 2023). *Caisse Centrale de Réassurance*.

CHARDON, Olivier, Yves JAUNEAU, and Joelle VIDALENC (2020). "Les agriculteurs: de moins en moins nombreux et de plus en plus d'hommes". *Insee Focus*, p. 65.

CHATELLIER, Vincent (2020). "Le paiement redistributif et le plafonnement des aides directes: deux outils de la PAC favorables aux petites exploitations agricoles francaises?" *Économie rurale. Agricultures, alimentations, territoires*, pp. 137–151.

CHATELLIER, Vincent and Hervé GUYOMARD (2020). *PAC et revenus agricoles*.

Communes de france - Base des codes postaux (2020).

CONNOR, Lawson, Roderick M. REJESUS, and Mahmut YASAR (2022). "Crop insurance participation and cover crop use: Evidence from Indiana county-level data". *Applied Economic Perspectives and Policy* 44, pp. 2181–2208.

DELAY, Nathan (2019). "The impact of federal crop insurance on the conservation reserve program". *Agricultural and Resource Economics Review* 48, pp. 297–327.

DI FALCO, Salvatore, Felice ADINOLFI, Martina BOZZOLA, and Fabian CAPITANIO (2014). "Crop insurance as a strategy for adapting to climate change". *Journal of Agricultural Economics* 65, pp. 485–504.

FADN (2023).

FANG, Lan, Rong HU, Hui MAO, and Shaojian CHEN (2021). "How crop insurance influences agricultural green total factor productivity: Evidence from Chinese farmers". *Journal of Cleaner Production* 321, p. 128977.

FOLUS, Didier, Pierre CASAL RIBEIRO, Bruno LEPOIVRE, and Antoine ROUMIGUIÉ (2020). "L'assurance et la protection financiere de l'agriculture". *Annales des Mines-Realites industrielles*. 1. FFE, pp. 30–38.

GAMMANS, Matthew, Pierre MEREL, and Ariel ORTIZ-BOBEA (2017). "Negative impacts of climate change on cereal yields: statistical evidence from France". *Environmental Research Letters* 12, p. 054007.

GÉRY, Claire, Vincent HECQUET, and Félix LUCAS (2023). *Le compte provisoire de l'agriculture pour 2022*.

GOODWIN, Barry K., Monte L. VANDEVEER, and John L. DEAL (2004). "An empirical analysis of acreage effects of participation in the federal crop insurance program". *American Journal of Agricultural Economics* 86, pp. 1058–1077.

GRIGORIEVA, Elena, Andreas MATZARAKIS, and Christopher DE FREITAS (2010). "Analysis of growing degree-days as climate impact indicator in a region with extreme annual air temperature amplitude". *Climate Research* 42, pp. 143–154.

HECKMAN, James J. and Edward J. VYTLACIL (2007). "Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments". *Handbook of econometrics* 6, pp. 4875–5143.

HORTON, Keyle (2018). "Addressing Climate Change Through a New Lens: A Case Study of the Sangamner Region of Western India". PhD thesis. University of Colorado at Boulder.

JOHNSON, S. R. and Gordon C. RAUSSER (1971). "Effects of misspecifications of log-linear functions when sample values are zero or negative". *American Journal of Agricultural Economics* 53, pp. 120–124.

KAMHÖFER, Daniel A., Hendrik SCHMITZ, and Matthias WESTPHAL (2019). "Heterogeneity in marginal non-monetary returns to higher education". *Journal of the European Economic Association* 17, pp. 205–244.

KOENIG, Richard, Marielle BRUNETTE, Philippe DELACOTE, and Camille TEVENART (2022). "Assurance récolte en France: spécificité du régime et déterminants potentiels". *Economie rurale* 380, pp. 7–25.

KORRES, Nicholas E., Jason K. NORSWORTHY, Parsa TEHRANCHIAN, Thomas K. GITSOPOULOS, Dimitra A. LOKA, Derrick M OOSTERHUIS, David R. GEALY, Stephen R. MOSS, Nilda R. BURGOS, M. Ryan MILLER, et al. (2016). "Cultivars to face climate change effects on crops and weeds: a review". *Agronomy for Sustainable Development* 36, pp. 1–22.

LIGNEAU, Laurence, Sylvain TILLY, Franck BARAER, Vincent DUBREUIL, and Valérie BONNARDOT (2020). "Observatoire du changement climatique pour l'agriculture: resultats preliminaires en Bretagne". *Changement Climatique et Territoires*, pp. 427–432.

LOI DU 2 MARS 2022 (2022). *Loi du 2 mars 2022 d'orientation relative à une meilleure diffusion de l'assurance récolte en agriculture et portant réforme des outils de gestion des risques en agriculture*.

LUO, Qunying (2011). "Temperature thresholds and crop production: a review". *Climatic change* 109, pp. 583–598.

MAA (2022). *L'Assurance Multirisque Climatique des Récoltes*.

MASA (2022). *La gestion des risques en agriculture*.

MATHIAS, Le et al. (2022). "L'adaptation des Economies au changement climatique: les enseignements tirés de la recherche Economique". *Bulletin de la Banque de France*.

MELLON, Jonathan (2022). "Rain, Rain, Go Away: 192 Potential Exclusion-Restriction Violations for Studies Using Weather as an Instrumental Variable". *Available at SSRN 3715610*.

RICA (2022). *Accès aux données individuelles du RICA*.

ROLL, Kristin H. (2019). "Moral hazard: the effect of insurance on risk and efficiency". *Agricultural Economics* 50, pp. 367–375.

RÖTTER, Reimund and Siebe C. VAN DE GEIJN (1999). "Climate change effects on plant growth, crop yield and livestock". *Climatic change* 43, pp. 651–681.

SANTERAMO, Fabio Gaetano, Barry K. GOODWIN, Felice ADINOLFI, and Fabian CAPITANIO (2016). "Farmer participation, entry and exit decisions in the Italian crop insurance programme". *Journal of Agricultural Economics* 67, pp. 639–657.

SCHLENKER, Wolfram, W. Michael HANEMANN, and Anthony C. FISHER (2007). "Water availability, degree days, and the potential impact of climate change on irrigated agriculture in California". *Climatic Change* 81, pp. 19–38.

SHUKLA, Priyadarshi R, Jim SKEA, E. CALVO BUENDIA, Valérie MASSON-DELMOTTE, Hans Otto PÖRTNER, DC ROBERTS, Panmao ZHAI, Raphael SLADE, Sarah CONNORS, Renée VAN DIEMEN, et al. (2019). "IPCC, 2019: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems".

SPINONI, J., J. VOGT, and P. BARBOSA (2015). "European degree-day climatologies and trends for the period 1951–2011". *International Journal of Climatology* 35, pp. 25–36.

Tableaux de l'économie française (2020).

USDA (2022). *Reinsurance Agreements*.

VELANDIA, Margarita, Roderick M. REJESUS, Thomas O. KNIGHT, and Bruce J. SHERRICK (2009). "Factors affecting farmers' utilization of agricultural risk management tools: the case of crop insurance, forward contracting, and spreading sales". *Journal of agricultural and applied economics* 41, pp. 107–123.

VIE PUBLIQUE (2022). *Agriculture française : une puissance mondiale qui décline*.

WALTHALL, Charles L., J. HATFIELD, P. BACKLUND, L. LENGNICK, E. MARSHALL, M. WALSH, S. ADKINS, M. AILLERY, E. A. AINSWORTH, C. AMMANN, et al. (2013). *Climate change and agriculture in the United States: Effects and adaptation*. United States Department of Agriculture, Agricultural Research Services.

WANG, Ruixue, Roderick M. REJESUS, and Serkan AGLASAN (2021). "Warming temperatures, yield risk and crop insurance participation". *European Review of Agricultural Economics* 48, pp. 1109–1131.

WOOLDRIDGE, Jeffrey M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

WORTHINGTON, Christine M. and Chad M. HUTCHINSON (2005). "Accumulated growing degree days as a model to determine key developmental stages and evaluate yield and quality of potato in Northeast Florida". *Proceedings of the Florida state horticultural society*. Vol. 118, pp. 98–101.

WU, Shenan, Barry K. GOODWIN, and Keith COBLE (2020). "Moral hazard and subsidized crop insurance". *Agricultural Economics* 51, pp. 131–142.

Appendix

A OLS estimates of Equation (3)

	EBIDTA with insurance subsidies		EBIDTA without insurance subsidies	
	(1) Mean	(2) Variance	(3) Mean	(4) Variance
Dummy for crop insurance status (1=insured)	0.004***	-0.000	0.003**	0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Cold GDDs (log)	0.003***	0.000	0.003***	0.000
	(0.001)	(0.000)	(0.001)	(0.000)
L.Cold GDDs (log)	-0.005***	-0.001***	-0.005***	-0.001***
	(0.001)	(0.000)	(0.001)	(0.000)
L2.Cold GDDs (log)	-0.000	-0.001**	-0.000	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
L3.Cold GDDs (log)	-0.003***	0.001***	-0.003***	0.001***
	(0.001)	(0.000)	(0.001)	(0.000)
Hot GDDs (log)	-0.014***	-0.001	-0.015***	-0.001
	(0.003)	(0.001)	(0.003)	(0.001)
L.Hot GDDs (log)	0.003	0.000	0.003	0.000
	(0.003)	(0.001)	(0.003)	(0.001)
L2.Hot GDDs (log)	0.011***	-0.001	0.011***	-0.001
	(0.002)	(0.001)	(0.002)	(0.001)
L3.Hot GDDs (log)	-0.012***	-0.002***	-0.012***	-0.002**
	(0.003)	(0.001)	(0.003)	(0.001)
Number of floods (log)	-0.006***	-0.000	-0.007***	-0.000
	(0.002)	(0.001)	(0.002)	(0.000)
L.Number of floods (log)	-0.001	0.001	-0.001	0.001
	(0.002)	(0.001)	(0.002)	(0.001)
L2.Number of floods (log)	0.002	-0.000	0.002	0.000
	(0.002)	(0.000)	(0.002)	(0.000)
L3.Number of floods (log)	-0.004**	-0.001	-0.004**	-0.001
	(0.002)	(0.001)	(0.002)	(0.001)
Number of droughts (log)	0.002	-0.002	0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.001)
L.Number of droughts (log)	0.005**	0.000	0.006***	0.000
	(0.002)	(0.001)	(0.002)	(0.001)
L2.Number of droughts (log)	0.004**	0.001	0.004**	0.001
	(0.002)	(0.001)	(0.002)	(0.001)
L3.Number of droughts (log)	0.003	-0.000	0.003*	-0.000
	(0.002)	(0.001)	(0.002)	(0.001)
Observations	71524	71524	70750	70750
ρ	1	0	1	0
Farmer FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Table 9: OLS log estimations for the impact of insurance on the revenue distribution

B Recovered effects from the MTE analysis and common support

B.1 Parametric

	(1) EBIDTA net of insurance insurance subsidies (log)	(2) Variance
effects		
ate	0.321***	-0.018*
	(0.033)	(0.010)
att	-0.156***	-0.001
	(0.016)	(0.006)
atut	0.501***	-0.025
	(0.049)	(0.016)
late	0.020**	-0.005*
	(0.009)	(0.003)
mprte1	0.092***	-0.009**
	(0.013)	(0.004)
mprte2	0.014	-0.012***
	(0.012)	(0.004)
mprte3	0.192***	-0.016**
	(0.023)	(0.008)
Observations	100834	70565

Table 10: Recovered estimators from the MTE framework

B.2 Semi Parametric

	(1) EBIDTA net of insurance subsidies (log)	(2) Variance
effects		
ate	0.180	-0.004
att	0.169	0.015
atut	0.189	-0.013
late	0.101	0.007
mprte1	0.185	0.002
mprte2	0.163	-0.000
mprte3	0.182	-0.009
Observations	100329	70213

Table 11: Recovered estimators from the MTE framework

B.3 Parametric MTE and common support

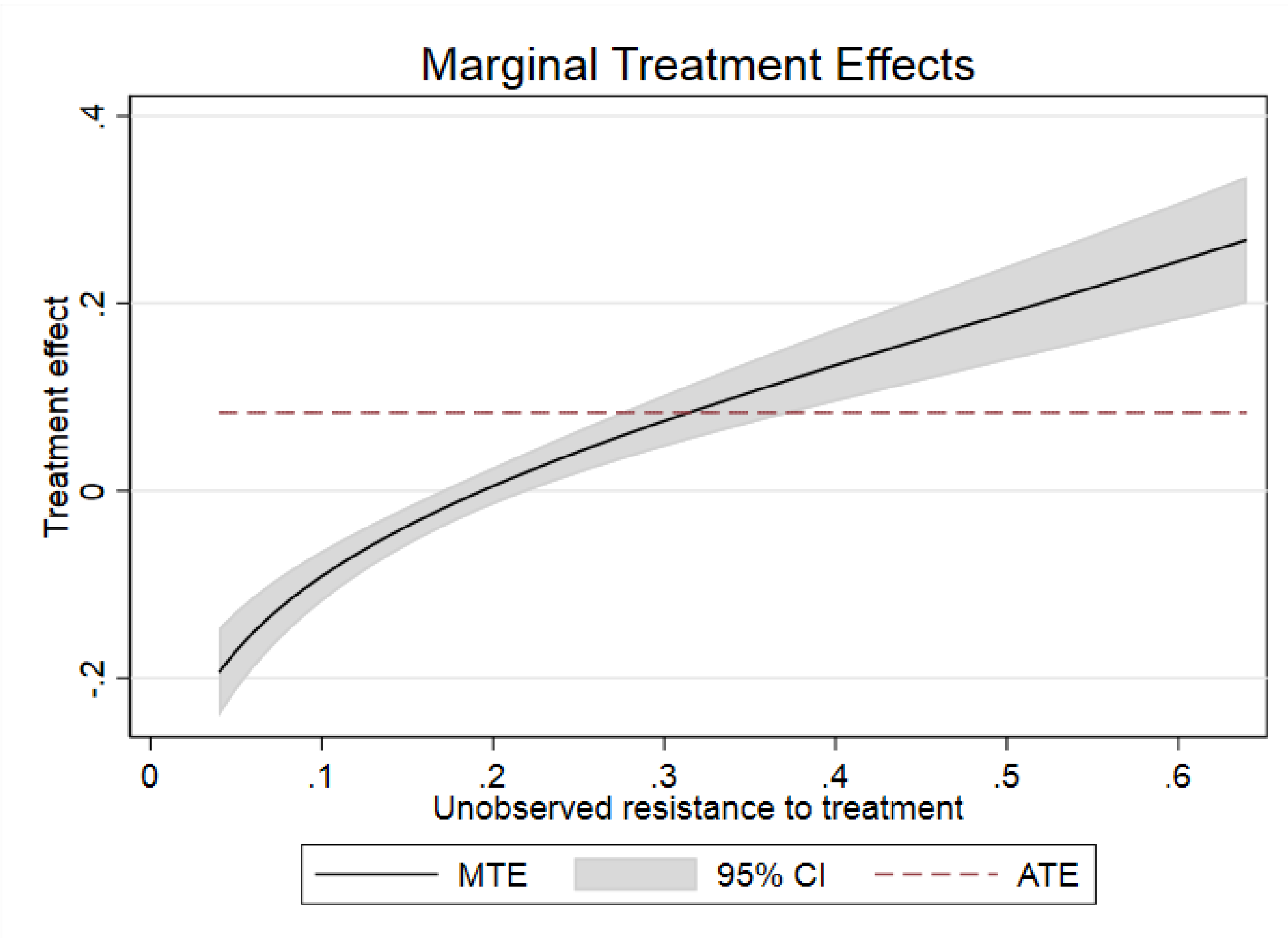

Figure 8: Quadratic MTE curve for mean of EBIDTA net of insurance subsidies

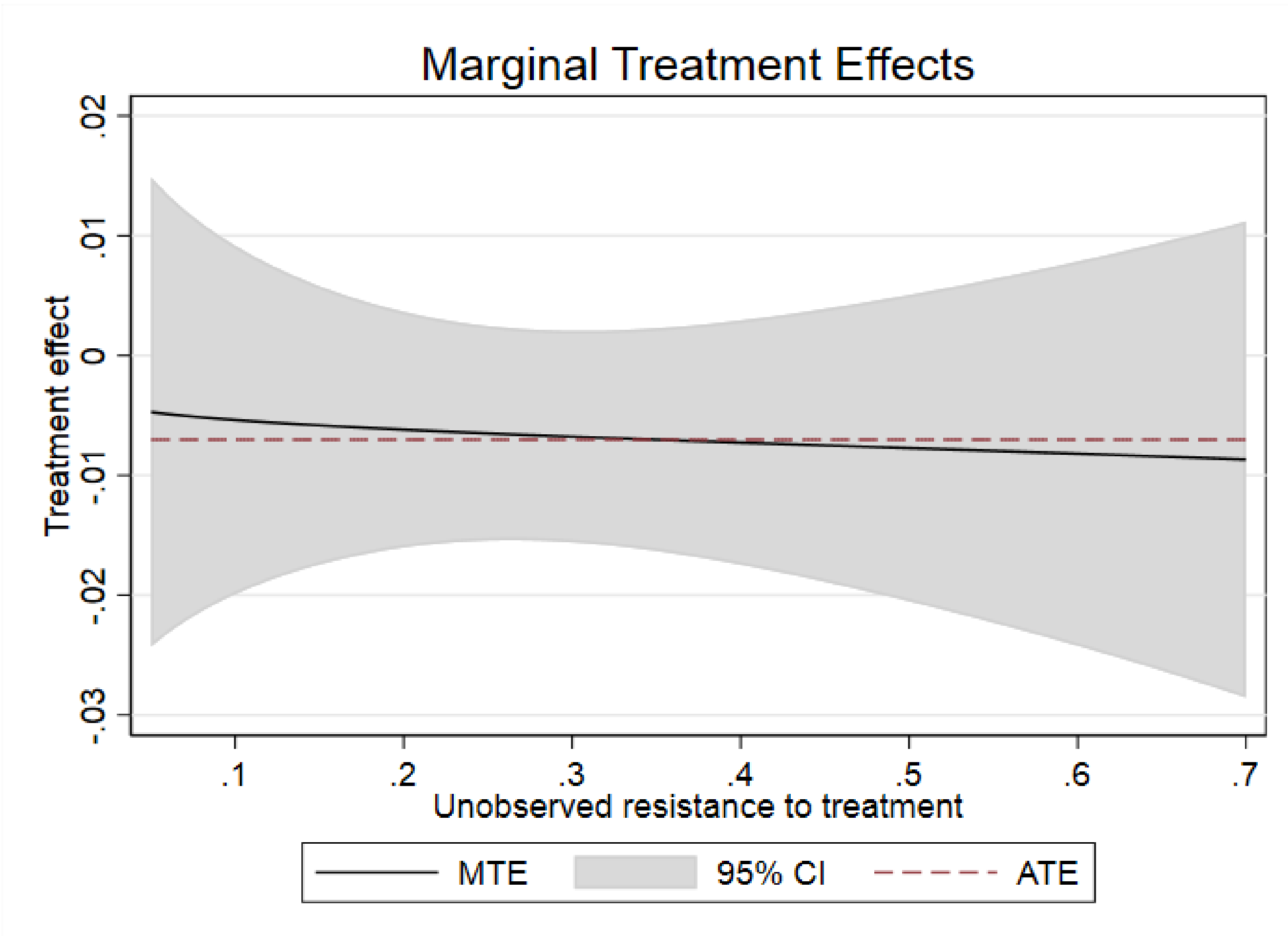

Figure 9: Quadratic MTE curve for variance of EBIDTA net of insurance subsidies

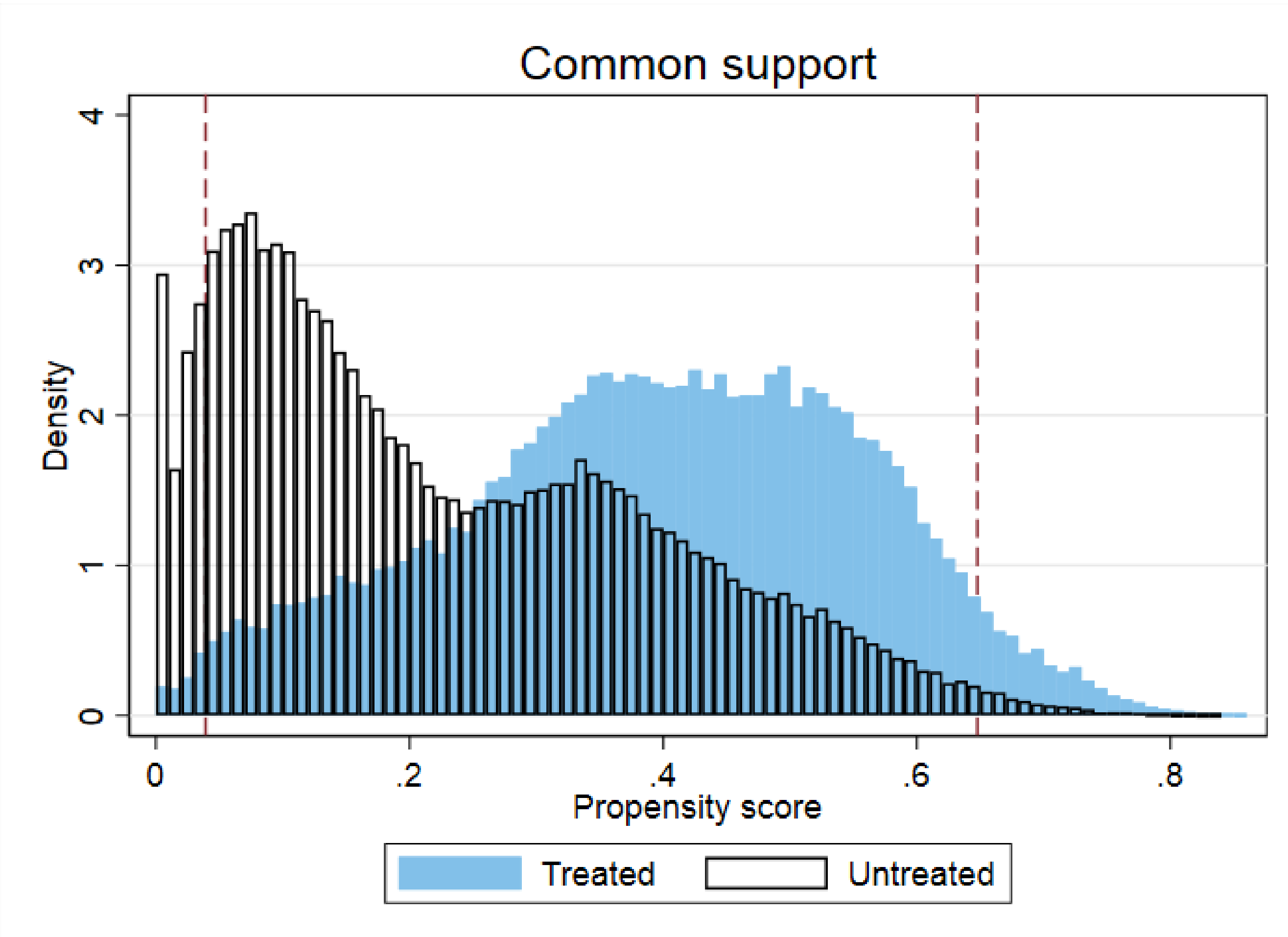

Figure 10: Common support for all MTE graphs

C Robustness checks

In addition to the instrumental variable approach, we perform a series of robustness checks both on the indicators used, and on a sub-sample of the data to ensure the validity of our results.

C.1 Alternative regression for Equation (3) with continuous effects

	EBIDTA with insurance subsidies		EBIDTA without insurance subsidies	
	(1)	(2)	(3)	(4)
Insurance spending (log)	0.046***	-0.003	0.039***	-0.001
	(0.007)	(0.002)	(0.006)	(0.002)
Cold GDDs (log)	0.002*	0.000	0.002**	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Hot GDDs (log)	-0.020***	-0.001	-0.020***	0.000
	(0.004)	(0.001)	(0.004)	(0.001)
Number of floods (log)	-0.010***	-0.000	-0.010***	-0.000
	(0.003)	(0.001)	(0.003)	(0.001)
Number of droughts (log)	0.003	-0.001	0.002	-0.000
	(0.003)	(0.001)	(0.003)	(0.001)
Observations	69790	69790	69006	69006
Weak Ident.	72.028	72.028	77.879	77.879
Hansen J	0.000	0.000	0.000	0.000
Farmer FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Instrument	Yes	Yes	Yes	Yes

Table 12: IV estimations for the impact of insurance on the revenue distribution

C.2 IHS transformation

The log transformation traditionnaly poses an issue with 0 and negative values. While our sample has mostly positive values for the log transformed variables (mainly EBIDTA), some negative values had to be taken into account. We followed the classic method of adding the minimum+1 to all variables, ensuring nothing got dropped and the log sample starts at 0. However, this transformation can cause problems in terms of elasticity interpretations (JOHNSON and RAUSSER, 1971), biasing the results. We therefore test the inverse hyperbolic sine transformation on our main variables (EBIDTA and insurance spending) using the same IV specification as in section 4. The coefficients retain the same signs with some changes in scale (noticeably higher), and are still statistically significant.

	With insurance subsidies		Without insurance subsidies	
	(1)	(2)	(3)	(4)
Dummy for crop insurance status (1=insured)	1.780*	-8.370	0.961**	-0.553
	(1.039)	(13.033)	(0.380)	(1.739)
Cold GDDs (log)	0.192***	-0.948*	0.093***	-0.094
	(0.044)	(0.517)	(0.016)	(0.070)
L.Cold GDDs (log)	0.147***	-1.456***	0.045***	-0.226***
	(0.042)	(0.507)	(0.016)	(0.069)
L2.Cold GDDs (log)	0.238***	-2.693***	0.089***	-0.344***
	(0.044)	(0.535)	(0.017)	(0.075)
L3.Cold GDDs (log)	-0.179***	1.775***	-0.085***	0.227***
	(0.045)	(0.539)	(0.017)	(0.072)
Hot GDDs (log)	0.034	-1.279	-0.072	-0.109
	(0.129)	(1.753)	(0.049)	(0.244)
L.Hot GDDs (log)	0.086	-2.637*	0.012	-0.357*
	(0.134)	(1.572)	(0.051)	(0.209)
L2.Hot GDDs (log)	0.174*	0.105	0.080**	0.046
	(0.102)	(1.314)	(0.039)	(0.176)
L3.Hot GDDs (log)	-0.031	0.105	-0.032	0.003
	(0.150)	(1.942)	(0.057)	(0.268)
Number of floods (log)	-0.496***	4.927***	-0.246***	0.643***
	(0.100)	(1.220)	(0.037)	(0.163)
L.Number of floods (log)	-0.101	0.633	-0.036	0.086
	(0.094)	(1.138)	(0.035)	(0.157)
L2.Number of floods (log)	-0.001	0.111	-0.005	0.004
	(0.093)	(1.184)	(0.035)	(0.167)
L3.Number of floods (log)	-0.258***	2.981**	-0.128***	0.419**
	(0.097)	(1.200)	(0.036)	(0.166)
Number of droughts (log)	0.214**	-2.301**	0.075**	-0.353**
	(0.088)	(1.104)	(0.033)	(0.151)
L.Number of droughts (log)	0.076	-1.062	0.042	-0.176
	(0.087)	(1.045)	(0.033)	(0.146)
L2.Number of droughts (log)	0.085	-0.168	0.041	-0.066
	(0.091)	(1.125)	(0.034)	(0.160)
L3.Number of droughts (log)	0.251***	-2.338**	0.100***	-0.318**
	(0.089)	(1.068)	(0.034)	(0.151)
Observations	69790	69790	69006	69006
Weak Ident.	168.984	168.984	180.817	180.817
Hansen J	0.000	0.000	0.000	0.000
Farmer FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Instrument	Yes	Yes	Yes	Yes

Table 13: IHS results for Equation (3)

C.3 Alternative instruments

We perform the previous regression using a different set of instruments. It might be the case that our preferred instrument (national subsidy rate by crop) might be endogenous if the decisions of farmers affect specific subsidy rates, which in turn affect both insurance intake and revenues. While unlikely, considering the fragmentation of the French agricultural sector, we nonetheless perform this robustness check using two additional instruments ; the 2005 and 2016 reforms. As discussed in the institutional context, the 2005 reform created subsidies to multirisk crop insurance, whereas the 2016 reform expanded the definition of weather shocks to make the contracts more protective (MAA, 2022). To account for the 2005 reform, we also expand our sample to include the period 2002-2022. The results of this test can be found in table 11.

The results are still significant and keep the same sign, but have slightly higher values. Because the instruments are now discrete, we only focus on a sub-sample of the data, that is farmers who changed their insurance behavior due only to the reforms in 2005 and 2016. This means that this regression might introduce an upwards bias since these reforms drastically improved the conditions of crop insurance and - especially in the case of the 2005 reform - created an entirely new family of insurance subsidies. This means that farmers with a lot to gain from insurance could now access the market. Regardless, while the subsidy rate remains our preferred instrument because it reduces the upward bias and has the distinct advantage of being continuous, this robustness test can be viewed as a confirmation of the results.

	EBIDTA with insurance subsidies		EBIDTA without insurance subsidies	
	(1)	(2)	(3)	(4)
Dummy for crop insurance status (1=insured)	0.288***	0.013*	0.280***	0.010*
	(0.022)	(0.008)	(0.022)	(0.006)
Cold GDDs (log)	0.011***	0.000*	0.011***	0.000
	(0.001)	(0.000)	(0.001)	(0.000)
L.Cold GDDs (log)	0.006***	0.000	0.007***	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)
L2.Cold GDDs (log)	0.018***	-0.001**	0.017***	-0.000*
	(0.001)	(0.000)	(0.001)	(0.000)
L3.Cold GDDs (log)	-0.001	-0.000	-0.000	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Hot GDDs (log)	-0.030***	-0.002	-0.030***	-0.001
	(0.004)	(0.001)	(0.004)	(0.001)
L.Hot GDDs (log)	-0.018***	-0.000	-0.017***	0.000
	(0.004)	(0.001)	(0.004)	(0.001)
L2.Hot GDDs (log)	0.014***	0.001	0.014***	0.001
	(0.004)	(0.001)	(0.004)	(0.001)
L3.Hot GDDs (log)	0.004	-0.001	0.003	-0.001
	(0.004)	(0.001)	(0.004)	(0.001)
Number of floods (log)	-0.015***	-0.001	-0.016***	-0.000
	(0.003)	(0.001)	(0.003)	(0.001)
L.Number of floods (log)	-0.003	0.000	-0.003	0.001
	(0.003)	(0.001)	(0.003)	(0.001)
L2.Number of floods (log)	-0.002	-0.000	-0.002	0.000
	(0.003)	(0.001)	(0.003)	(0.001)
L3.Number of floods (log)	-0.002	-0.000	-0.002	-0.000
	(0.003)	(0.001)	(0.003)	(0.001)
Number of droughts (log)	-0.003	-0.002	-0.003	-0.001
	(0.003)	(0.001)	(0.003)	(0.001)
L.Number of droughts (log)	0.008***	0.000	0.009***	-0.000
	(0.003)	(0.001)	(0.003)	(0.001)
L2.Number of droughts (log)	0.003	0.001	0.003	0.001
	(0.003)	(0.001)	(0.003)	(0.001)
L3.Number of droughts (log)	0.003	0.001	0.004	0.001
	(0.003)	(0.001)	(0.003)	(0.001)
Observations	69862	69862	69078	69078
Weak Ident.	140.615	140.615	142.756	142.756
Hansen J	163.553	8.906	167.234	6.443
Farmer FE	Yes	Yes	Yes	Yes
Year FE	No	No	No	No
Controls	Yes	Yes	Yes	Yes
Instrument (reforms)	Yes	Yes	Yes	Yes

Table 14: Alternative IV framework

C.4 Real subsidy rates

Additionnaly to the alternative instrument, we also provide results for a sub-sample tests using only the real subsidy rates from official the official documents from 2016 onwards, rather than subsidy rates estimated from the data as explained in section 3. The results keep the same sign but due to the reduced sample lose a lot of their significance. Nonetheless, the coefficients confirm that our subsidy rates do not differ too much from the real ones on the 2016-2021 period.

	EBIDTA with insurance subsidies		EBIDTA without insurance subsidies	
	(1)	(2)	(3)	(4)
Dummy for crop insurance status (1=insured)	0.295	-0.102	0.289	-0.112
	(0.192)	(0.067)	(0.208)	(0.075)
Cold GDDs (log)	0.002	0.002^{*}	0.001	0.001
	(0.003)	(0.001)	(0.003)	(0.001)
L.Cold GDDs (log)	0.002	0.000	0.002	0.000
	(0.003)	(0.001)	(0.004)	(0.001)
L2.Cold GDDs (log)	0.008^{**}	0.001	0.007^{**}	0.002
	(0.003)	(0.001)	(0.003)	(0.001)
L3.Cold GDDs (log)	-0.004^{*}	0.001	-0.004^{*}	0.000
	(0.002)	(0.001)	(0.002)	(0.001)
Hot GDDs (log)	0.003	-0.002	0.002	-0.001
	(0.006)	(0.002)	(0.006)	(0.002)
L.Hot GDDs (log)	-0.009^{*}	-0.000	-0.010^{*}	-0.000
	(0.005)	(0.002)	(0.005)	(0.002)
L2.Hot GDDs (log)	-0.010	0.004	-0.010	0.005
	(0.012)	(0.004)	(0.013)	(0.004)
L3.Hot GDDs (log)	-0.002	0.007	-0.002	0.007
	(0.016)	(0.005)	(0.017)	(0.005)
Number of floods (log)	-0.012^{**}	-0.002	-0.013^{**}	-0.002
	(0.005)	(0.002)	(0.005)	(0.002)
L.Number of floods (log)	0.003	-0.000	0.003	0.001
	(0.005)	(0.002)	(0.005)	(0.002)
L2.Number of floods (log)	0.009^{**}	-0.002	0.009^{**}	-0.001
	(0.004)	(0.001)	(0.004)	(0.001)
L3.Number of floods (log)	-0.005	-0.002	-0.006	-0.001
	(0.005)	(0.002)	(0.005)	(0.002)
Number of droughts (log)	-0.003	-0.002	-0.003	-0.000
	(0.004)	(0.002)	(0.004)	(0.002)
L.Number of droughts (log)	-0.001	0.002	-0.001	0.001
	(0.004)	(0.002)	(0.004)	(0.002)
L2.Number of droughts (log)	-0.002	0.002	-0.001	0.001
	(0.005)	(0.002)	(0.005)	(0.002)
L3.Number of droughts (log)	0.000	-0.000	0.001	-0.002
	(0.006)	(0.002)	(0.005)	(0.002)
Observations	24239	24239	23984	23984
Weak Ident.	4.872	4.872	4.030	4.030
Hansen J	0.000	0.000	0.000	0.000
Farmer FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Instrument (Real rate)	Yes	Yes	Yes	Yes

Table 15: IV with real subsidy rates