Don't bet the Farm on Crop Insurance Subsidies: A Marginal Treatment Effect Analysis of French Farms[1]

Céline Grislain-Letrémy[2], Bertrand Villeneuve[3], Marc Yeterian[4]

July 2024, WP 956

Abstract

Crop insurance is one of the most important protections against climate-related risks for farmers. Despite being heavily subsidized, insurance take-up in France remains surprisingly low. The goal of this paper is twofold; first, we explain this paradox by analyzing the heterogeneous effects of taking up crop insurance, and second, we provide concrete welfare-enhancing policy recommendations to increase insurance take-up. Using a micro-level panel of 17,000 French farmers over 20 years, we first use a moment-based regression to identify the local average treatment effects (LATE) of insurance on expected revenues and variance. Then we investigate the factors causing the heterogeneity in these effects, both observable through interaction terms and unobservable through a marginal treatment effect (MTE) design. We conclude that insurance subsidies have very little impact on crop insurance demand, especially for those who would benefit the most. Finally, we suggest cost-efficient ways to increase insurance take-up based on administrative simplification, information and imitation.

Keywords: Insurance; Agriculture; Marginal Treatment Effects; Instrumental Variable.

JEL classification: G22; Q10; Q12; Q14

[1] The authors thank Eric Bonsang, Raja Chakir, Salvator di Falco, Loïc Henry and Céline Nauges for their insightful comments. This paper has also benefited from comments by participants at the Modeling Agricultural TransitionS (MATS) seminar at Collège de France, the European Association of Environmental and Resource Economics (EAERE) conferences in 2023 and 2024, the Journées de Microéconomie Appliquées (JMA) in 2023 and 2024, the Association Française de Sciences Economiques conference (AFSE) in 2023, the Agricultural Economics Society conference in 2024, the CGEMP (resource economics) internal seminar at Dauphine. and at the Banque de France research seminar. They thank Alfred Razol for his excellent initial research assistant work.

[2] Banque de France and CREST-ENSAE, celine.grislain-letremy@banque-france.fr

[3] Université Paris-Dauphine, PSL Research University, bertrand.villeneuve@dauphine.psl.eu.

[4] Université Paris-Dauphine, PSL Research University, marc.yeterian@dauphine.psl.eu.

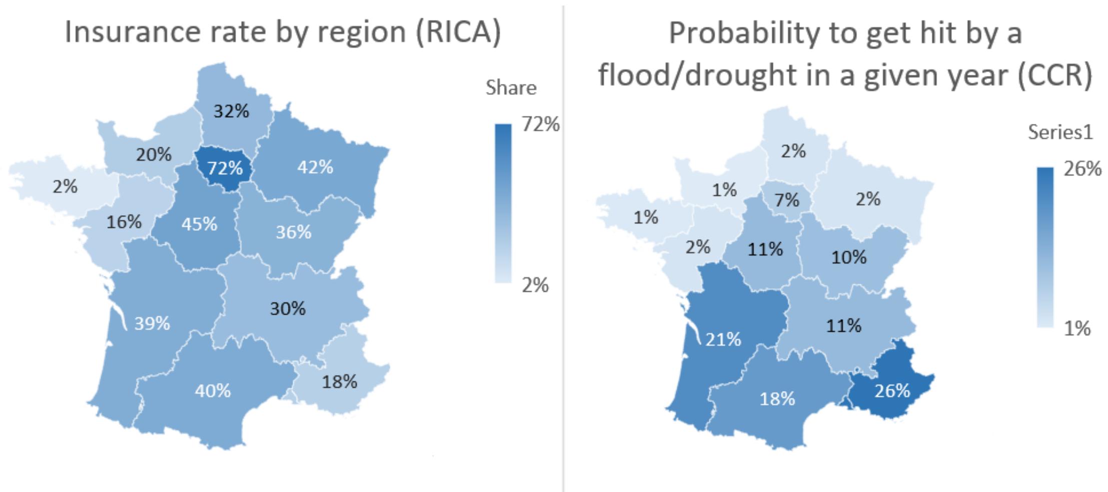
NON-TECHNICAL SUMMARY

The penetration rate of crop insurance is low in France. Only 13.3% of farms were insured in 2020, and the numbers are stable, perhaps in slight increase, from 12% in 2016. Yet, farming is a risky activity and the climate, a driver of the yearly yield variations, is changing towards more frequent extremes. This begs the question: why is adoption so surprisingly limited? Not only is the need for protection insurance is clear, but insurance premiums are highly subsidized and insurance has a net positive impact on earnings, according to the literature. Providing insight into this paradox is the main motivation behind this paper.

First, we provide a new evaluation of the impact of crop insurance on farmers' revenues mean and variance to determine to what extent subscription is an attractive choice. Second, we explain the low take-up through an analysis of the observable determinants of crop insurance subscription. We go beyond treatment effect by exploring the heterogeneity inasmuch as it is partly explained by observables, as expected, but also with structured unobservables. Third, we perform counterfactual analyses of policies to explore their efficiency in both increasing the take-up rate and yielding high benefits for farmers.

We combine a variety of econometric methods, including the most flexible and adequate selection model to analyze, jointly and distinctly, choice and expected benefits: the Marginal Treatment Effect (MTE) framework à la Heckman & Vytlacil. The MTE approach has, to our knowledge, never been used in the context of crop insurance. It allows counterfactual analyses and provides key insights regarding the right policies to maximize insurance take-up and social welfare.

Figure 1. Map of insurance rate and risk exposure by region



Data sources: RICA, Caisse Centrale de Reassurance; Authors' production

Thanks to our methodology and data, our results are therefore more precise, realistic and actionable than those of previous studies. Our results are generally in line with the previous literature as far as averages are concerned (i.e. insurance take-up is generally beneficial to farmers), yet the details are extremely interesting and they have been overlooked thus far. 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, benefits from their subscriptions. Second, we find a "contrarian" selection where, in most cases, farmers who would benefit the most from insurance tend to insure the least. This is true both across and within observable characteristics, i.e., even when conditioning on observables, the contrarian selection still occurs, meaning that there are unobservable barriers to subscriptions (beliefs, non-financial barriers, etc.). Third, we 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 take-up, and those newly insured farmers would actually derive little benefit from their contracts. Instead, overcoming the non-financial barriers to insurance (i.e. information, paperwork) by directly targeting the propensity score (probability to insure) appears to be the optimal way of tackling this issue. Pursuing a goal of 100% of insured farmers might not be a feasible nor a desirable outcome, and smaller, more specialized farms should be targeted instead.

To perform this large-scale analysis across mainland France over a 20-year period, we produce a unique and highly granular dataset composed from individual data on farmers. This dataset combines several sources; it includes agronomic and financial variables coming from the French "Réseau d'Information Comptable Agricole" (part of the European Farm Accountancy Data Network), weather data at a 0.1°×0.1° resolution (from Copernicus) and administrative data for climate disasters (from the French public reinsurer, Caisse Centrale de Réassurance).

Assurance récolte pour les agriculteurs français : Une analyse des effets marginaux de traitement de subventions mal ciblées

RÉSUMÉ

L'assurance récolte est l'une des protections les plus importantes contre les risques liés au climat pour les agriculteurs. Bien qu'elle soit fortement subventionnée, la souscription d'une assurance en France reste étonnamment faible. L'objectif de cet article est double : premièrement, nous expliquons ce paradoxe en analysant les effets hétérogènes de la souscription d'une assurance récolte et, deuxièmement, nous formulons des recommandations de politiques publiques concrètes afin d'augmenter la souscription d'assurance en visant à améliorer le bien-être social. À l'aide d'un panel très détaillé de 17000 agriculteurs français sur 20 ans, nous réalisons tout d'abord une régression basée sur les moments pour identifier les effets de traitement moyens locaux (LATE) de l'assurance sur la moyenne et la variance des revenus. Nous étudions ensuite les facteurs à l'origine de l'hétérogénéité de ces effets, à la fois observables via des termes d'interaction et non observables grâce à un modèle d'effets marginaux de traitement (MTE). Nous concluons que les subventions à l'assurance n'ont que très peu d'impact sur la demande d'assurance récolte, en particulier pour ceux qui en bénéficieraient le plus. Enfin, nous suggérons des moyens rentables d'augmenter l'adhésion à l'assurance en se basant sur la simplification administrative, l'information et l'imitation.

Mots-clés : assurance, agriculture, effets marginaux de traitement, variable instrumentale.

1 Introduction

The penetration rate of crop insurance is low in France. Only 13.3% of farms were insured in 2020 (VIE PUBLIQUE, 2022b), and the numbers are stable, perhaps in slight increase from 12% in 2016 (MINISTÈRE DE L'AGRICULTURE, 2022b). Yet, farming is a risky activity and the climate, a driver of the yearly yield variations, is changing towards more frequent extremes. This begs the question: why is adoption so surprisingly limited? Not only is the need for protection insurance clear, but insurance premiums are highly subsidized and insurance has a net positive impact on earnings, according to the literature (e.g. DI FALCO et al., 2014). Providing insight into this paradox is the main motivation behind this paper.

First, we provide a new evaluation of the impact of crop insurance on farmers' revenues mean and variance to determine to what extent subscription is an attractive choice. Second, we explain the low take-up through an analysis of the observable determinants of crop insurance subscription. We go beyond treatment effect by exploring the heterogeneity inasmuch as it is partly explained by observables, as expected, but also with structured unobservables. Specifically, we identify the contrarian selection that makes the biggest benefactors of insurance the most reluctant to subscribe. Third, we perform counterfactual analyses of policies to explore their efficiency in both increasing the take-up rate and yielding high benefits for farmers.

We combine a variety of econometric methods, including the most flexible and adequate selection model to analyze, jointly and distinctly, choice and expected benefits: the Marginal Treatment Effect (MTE) framework à la Heckman & Vytlacil (HECKMAN and VYTLACIL, 2007). The MTE approach has, to our knowledge, never been used in the context of crop insurance. It allows counterfactual analyses and provides key insights regarding the right policies to maximize insurance take-up and social welfare.[1]

Thanks to our methodology and data, our results are therefore more precise, realistic and actionable than those of previous studies. Our results are generally in line with the previous literature (DI FALCO et al., 2014; ANNAN and SCHLENKER, 2015; WANG, REJESUS, and AGLASAN, 2021) as far as averages are concerned (i.e. insurance take-up is generally beneficial to farmers), yet the details are extremely interesting and they have been overlooked thus far. 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 a "contrarian" selection where, in most cases, farmers who would benefit the most from insurance tend to insure the least. This is true both across and within observable characteristics, i.e., even when conditioning on observables, the contrarian selection still occurs, meaning that there are unobservable barriers to subscriptions (beliefs, non-financial barriers, etc.). Third, we 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 take-up, and those newly insured farmers would actually derive little benefit from their contracts. Instead, overcoming the non-financial barriers to insurance (i.e. information, paperwork) by directly targeting the propen-

[1] It has been used for health insurance by KOWALSKI (2023).

sity score (probability to insure) appears to be the optimal way of tackling this issue. Pursuing a goal of 100% of insured farmers might not be a feasible nor a desirable outcome, and smaller, more specialized farms should be targeted instead.

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. For the average effects, we use a parametric moments-based instrumental variable 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 Local Average Treatment Effect (LATE) of insurance on the revenue distribution.[2] For the selection on observables, we combine a Probit regression on the probability to be insured and an interacted regression on the benefits of insurance uptake on revenues using the take-up variables identified in the Probit. For the main part of the paper, i.e. the selection on unobservables, we use the MTE framework to showcase the contrarian selection effect. For the policy analysis, we reuse the marginal treatment effect estimates to evaluate two counterfactual policies: a 2pp increase in insurance subsidies, and information campaign through the use of the Policy Relevant Treatment Effects (PRTE). Finally, to explain our results, we provide both an *ex ante* theoretical analysis and an *ex post* empirical exploration of the potential channels (moral hazard, shielding, and other endogeneity in farming practices) that might drive them.

To perform this large-scale analysis across mainland France over a 20-year period, we combine several sources to produce a unique and highly granular dataset composed from individual data on farmers, including agronomic and financial variables coming from the French "Réseau d'Information Comptable Agricole" (part of the European Farm Accountancy Data Network), weather data at a 0.1°×0.1° resolution (from Copernicus) and administrative data for climate disasters (from the French public reinsurer, Caisse Centrale de Réassurance).

The paper is structured as follows. Section 2 provides context on the crop insurance market in France, Section 3 a literature review of previous works on the topic. Section 4 provides the theoretical framework for the paper. Section 5 discusses the estimation strategy. Section 6 presents the data and summary statistics. Section 7 provides and discusses the results. Section 8 provides the counterfactual policy analysis, and Section 9 concludes the paper.

2 Context

2.1 Agriculture in France

France has a 17% share of the 27-country European Union's GDP. It is the leading agricultural producer in Europe and the fifth globally (VIE PUBLIQUE, 2022a). The sector contributed 3% to the national GDP in 2022 and employed 1.5% of the workforce (GÉRY, HECQUET, and LUCAS, 2023). However, the sector is declining. The number of farmers has rapidly fallen over the past 40 years,

[2] Classically, the LATE is defined as the treatment effect for those that responded to the instrument, i.e. farmers who subscribe (or cancel) insurance in reaction to a change in subsidies.

with their share in overall employment being almost divided by five (CHARDON, JAUNEAU, and VIDALENC, 2020).

In the meantime, the French agricultural sector has been largely affected by climate change. Numerous studies have attempted to quantify the impact of climate change on agricultural yields, focusing on both variability (i.e., yield risk) and the mean, with contrasted results. On a macro level for the EU, the cumulative effects of climate change—including human behavior and all ecological impacts—on productivity loss seem to range between 10% and 15% (BOSELLO and ZHANG, 2005). These averages conceal heterogeneous realities and massive variability. While globally, macroeconomic results show that French wheat and corn productions have lost 3.8% to 5.5% between 1990 and 2008 (LÉ, 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 hectare per year (+1.2% on average), primarily due to a shortening of the growth period (LIGNEAU et al., 2020).

In addition to its effect on yields, climate change has decreased crop resistance to floods and droughts as a result of the more intensive harvest rate, which in turn has accentuated revenue variability. Specifically in France, GAMMANS, MEREL, and ORTIZ-BOBEA (2017) use a fixed effects model combined with the RCP8.5 (high emissions) scenario to estimate that with constant technology, overall productivity of wheat and barley will decline by 17%-33% by 2100. Other micro-level studies include BAREILLE and CHAKIR (2023) and BAREILLE and CHAKIR (2024), which both show the technical difficulties (omitted variables, adaptation, etc.) with forecasting the impact of climate change on yields. One contributing factor 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, as workers are able to work fewer hours at a reduced productivity rate.[3]

2.2 Institutional context

Though classical worries about the fragility of farms finance has led the French Government to subsidize crop insurance schemes from 2005 onwards (SÉNAT, 2003), this changing context reinforces the interest in the support schemes. Unless otherwise specified, "crop insurance" in this paper refers to revenue insurance (price insurance also exists in France but is not subsidized). These insurance subsidies have become an important part of both the French and European aid to agriculture, and have been integrated with the Common Agricultural Policy (CAP) in 2016 (MINISTÈRE DE L'AGRICULTURE, 2022a). KOENIG et al. (2022) provides a good point of departure for a description of the system.

The public agricultural disaster insurance scheme. The "Dispositif des Calamités Agricoles" (DCA) covers farmers against losses caused by exceptional climatic events. If the farmer's losses

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

are small (less than 30% of annual production), the payouts can be directly financed by a national fund ("Fonds National de Gestion des Risques en Agriculture", FNGRA), which is itself co-financed by farms (one third) and the State (two thirds). If losses are larger, compensation is funded by the European agricultural fund for rural development. This scheme is limited by design and needs to be complemented with private insurance, because of the low payouts (rate of compensation below 45%) and because of threshold effects that discourage diversification and higher yields (BABUSIAUX, 2000). Indeed, 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 proportionally less compensated. Besides, almost every farmer in France receives direct aids from the European Common Agricultural Policy (CAP) scheme; these aids correspond on average to 88% of the revenues of farmers in France (CHATELLIER et al., 2021). However, these aids do not replace insurance, since they are not conditioned on output, hazards nor climate shocks that may occur.

Subsidized crop insurance. The private insurance sector offers more customizable and diversified products (FOLUS et al., 2020). These include crop insurance (covering 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. Since 2009, crop insurance is publicly subsidized, at 65% for a first tranche of the insured value and at 45% for a second tranche of the insured value, the remaining one being unsubsidized. The subsidies only cover parts of the insurance contract; in practice, a farmer chooses to insure the first tranche only, the first two ones or to be fully insured. Besides, the limits of tranches of insured value change every year, while remaining generally lower than the market price of the crop (see Appendix B.2 for details and examples). This is why the actual subsidy rate, defined as the ratio of the subsidy to premium paid, varies between 0% and 47% depending on the crop and on the year (see Table 2). Because they are highly subsidized, these insurance 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 and non-subsidized tiers might include market insurance (i.e. protection against price drops or demand losses) or compensation for supply chain issues. Subsidies are directly given to the farmer months after they actually paid for the contract and may slightly differ from those announced due to budget restrictions, on which farmers based their decision to insure or not.[4]

The 2023 reform. Our studied period (2005-2021) does not include the changes introduced by the 2022 reform in France. This reform, in force since 2023, is based on a three-tier system in which

[4]The budget allocated to insurance subsidies is decided by decree (for example, JOURNAL OFFICIEL (2015)). While the subsidy rate is fixed, if the total amount of subsidies that should be paid exceed the budget, the subsidy rate is reduced. For example, in 2013, the allocated budget was €77M for a subsidy demand of €105M (MINISTÈRE DE L'AGRICULTURE, 2018). This has led to the subsidy rate being reduced at 43% for cereals in this given year (remaining at 65% for vegetables). These adjustments are different every year and can vary by crop.

each of the players assumes a share of the risks. Low-intensity risks are borne by the farmer (1st tier). Medium-intensity risks are covered by the insurer if the farmer has taken out a subsidizable crop insurance policy (2nd tier). This level is applied above a threshold of 20% (deductible). For exceptional risks (3rd tier), a national solidarity allowance is paid, financed by the State and the insurer or the farmer, depending on his situation. This three-tier system is accompanied by an increase in the subsidy to 70%.[5]

International comparison. In many developed countries, coverage of farmers against climatic events relies on a hybrid public-private system with varying degrees of public intervention.[6] In the US the public insurance scheme takes the form of reinsurance scheme as a last resort, completed by a highly subsidized private system (USDA, 2022), just like in Italy and France, where a public ex-post payments and a subsidized private sector coexist (MINISTÈRE DE L'AGRICULTURE, 2022a; CAPITANIO et al., 2011). The US system is particularly interesting as their crop insurance subsidies appear to be inefficiently distributed, which is an issue we explore in our setting in details in Section 8. Two reports from the US Government Accountability Office (GAO) have indeed come out in 2023 (GAO, 2023b; GAO, 2023a), showing that subsidy recipients are mostly concentrated towards the largest firms: 1% of policyholders accounted for 22% of total subsidies received, with high-income policyholders (over $900k per year) earning almost double the subsidies (per farm) compared to the rest of the sample. Using descriptive analysis, those reports estimate that while the crop insurance program costs over $17Bn per year to the Government, they could reduce this bill by an order of $100M without impacting either crop insurance subscription or wealth creation.

Supply. Supply is highly regulated. Insurance companies offering subsidized products are approved by the Ministry and they commit to comply with quality and price standards. This situation limits market power in a context of high concentration.

3 Literature review

This paper fits in the body of literature analyzing the empirical links between climate change, agricultural yields and risk management strategies (VELANDIA et al., 2009; WALTHALL et al., 2013). 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 specialization, 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.

[5]This reform is still too recent to analyze its effects empirically, and at the time of writing only theoretical work has been performed (e.g. ROZAN and SPAETER, 2024).

[6]See BURKE and EMERICK (2016) for a study of the (slow, if any) adaptation in the US.

A few core papers have contributed to the research on the impacts of crop insurance, especially in Europe. 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 farmers' 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 it 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, while financial variables such as the liability ratio yield smaller elasticities (Enjolras and Sentis, 2011; Enjolras, Capitanio, and Adinolfi, 2012). Climate variables yield less stable results due to the fact that they may not be measuring proper shocks. Since Di Falco et al. (2014), other papers have refined the methods (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 net of insurance payments, 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. Other studies look at the behavioral impacts of insurance, both empirically and theoretically, and find contradictory results, i.e. crop insurance may decrease protection behavior, which would be akin to moral hazard (Smith and Goodwin, 1996), or it may increase protection behavior (shielding) as in Horowitz and Lichtenberg (1994) or Chakir and Hardelin (2014). Wu (1999) show that providing crop insurance causes several shifts, one of which is the increase of chemical use at the extensive margin. In the same line, Yu, Smith, and Sumner (2018) show that crop insurance subsidies increase acreage. Though these effects are not our core question, we investigate them altogether.

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) show 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.

4 Model

4.1 Baseline framework

In the textbook model of insurance with perfect competition and no friction (no loadings), insurance keeps the expected revenue constant but reduces its variability. Insurance typically provides an improvement in whatever objective the farmer follows.[7] With loadings, the choice is a trade-off between a reduced mean revenue (a minus) and a reduced variability (a plus). The best choice depends on the magnitude of loadings and on the risk aversion of the farmers. Some farmers may estimate that the gain is not worth the cost. With subsidies, insurance may increase expected revenue and decrease variability, implying a clear gain. Not subscribing appears now as completely irrational. Yet behavorial economics has taught us to be careful. Understanding is more important that judging.

Let's model the revenue as a random variable in which the premium and the indemnity are considered directly:

$$R = R_0 - \widetilde{x} + \alpha\widetilde{x} - (1+\lambda)\alpha\mathbb{E}\widetilde{x}, \tag{1}$$

where R_0 is a certain revenue, $\widetilde{x}$ is a positive random variable (the loss), α is the chosen coverage rate, λ is the loading factor.

If the State reimburses a fraction τ of the insurance premium, then the subsidy is $\tau(1+\alpha)\mathbb{E}\widetilde{x}$, the farmer actually pays $(1-\tau)(1+\lambda)\alpha\mathbb{E}\widetilde{x}$ instead of $(1+\lambda)\alpha\mathbb{E}\widetilde{x}$. Thus the expected revenue is

$$\mathbb{E}R = R_0 - (1+\lambda\alpha)\mathbb{E}\widetilde{x} + \tau(1+\lambda)\alpha\mathbb{E}\widetilde{x}. \tag{2}$$

We define now R_{NS} as R net of subsidies.

$$\mathbb{E}R_{\text{NS}} = R_0 - (1+\lambda\alpha)\mathbb{E}\widetilde{x}. \tag{3}$$

4.2 Comparative statics

Under perfect competition, the revenues and benefits of insurance cancel out (just set λ and τ to 0). In general, the impact of insurance on average revenue depends explicitly on λ and τ, as evidenced by the derivatives of $\mathbb{E}R$ and $\mathbb{E}R_{\text{NS}}$ wrt. α:

$$\frac{\partial \mathbb{E}R}{\partial\alpha} = (\tau(1+\lambda)-\lambda)\mathbb{E}\widetilde{x} + (\tau(1+\lambda)\alpha - (1+\lambda\alpha))\frac{\partial\mathbb{E}\widetilde{x}}{\partial\alpha}, \tag{4}$$

$$\frac{\partial \mathbb{E}R_{\text{NS}}}{\partial\alpha} = -\lambda\mathbb{E}\widetilde{x} - (1+\lambda\alpha)\frac{\partial\mathbb{E}\widetilde{x}}{\partial\alpha}. \tag{5}$$

We examine now the signs of these derivatives.

[7] One may call this effect second-order stochastic dominance.

No moral hazard. $\frac{\partial \mathbb{E}\widetilde{x}}{\partial \alpha} = 0$ is the total absence of moral hazard: the coverage doesn't change the expected loss. Then, the revenue net of subsidies R_{NS} decreases with coverage, because more resources are absorbed by loadings, and thus not returned as indemnities. For the revenue R, the effect of insurance can be an increase if the subsidy is sufficiently generous. This situation is plausible in our context.

Moral hazard. In the case of moral hazard where $\frac{\partial \mathbb{E}\widetilde{x}}{\partial \alpha} > 0$, the decrease of expected revenue is accentuated for R_{NS}. For the revenue R, the effect of moral hazard is negative because $\tau(1+\lambda)\alpha - (1+\lambda\alpha) < 0$ for relevant values of the parameters.[8] Consequently, the expected revenue can increase with respect to insurance coverage if the subsidy is strong enough and moral hazard limited. This case is not the most likely in our empirical study.

Shielding. We refer to the case where $\frac{\partial \mathbb{E}\widetilde{x}}{\partial \alpha} < 0$ as "shielding."

In economic terms, insurance and other protective measures are complements in that case. A "behavioral" intuition explains this effect if, for example, *attention to risk* is either triggered or not. If it is, the farmer uses all sorts of ways to limit risk (insurance and other protective measures). This risk reduction may go with an increased revenue. In that case, insurance is causal. Alternatively, a selection effect may play a role, and the statistical association is not causal: the most risk averse have a higher propensity to take insurance and are more protective at the same time. Our econometric approach disentangles causal and selection effects through the analysis of the marginal treatment effects, as discussed in Section 5.4. Indeed we estimate that R_{NS} increases with insurance on average.

The following proposition suggests plausible interpretation of the empirical results.

Proposition 1 *A necessary condition for the revenue net of subsidies to increase with coverage is that farmers exhibit shielding behavior. I.e.*

$$\frac{\partial \mathbb{E}R_{NS}}{\partial \alpha} > 0 \Rightarrow \frac{\partial \mathbb{E}\widetilde{x}}{\partial \alpha} < 0. \tag{6}$$

Mispricing. Up until now, we have assumed that insurers adjust premia to the actual risk. In particular they price moral hazard or, more likely, shielding (i.e., they know that $\widetilde{x}$ is endogenous). If, on the other hand, moral hazard (or shielding) is ignored in the premium, then insurers set premia as if $\mathbb{E}\widetilde{x}$ was freezed at its historical value denoted by $\mathbb{E}^*$. In other terms, they reason (wrongly) as if $\frac{\partial \mathbb{E}\widetilde{x}}{\partial \alpha} = 0$. Then

$$R = R_0 - \widetilde{x} + \alpha\widetilde{x} - (1+\lambda)\alpha\mathbb{E}^* + \tau(1+\lambda)\alpha\mathbb{E}^*, \tag{7}$$

$$R_{\text{NS}} = R_0 - \widetilde{x} + \alpha\widetilde{x} - (1+\lambda)\alpha\mathbb{E}^*. \tag{8}$$

[8] Indeed, the expression (4) is maximal for $\tau = 1$ (we exclude subsidies of more than 100%). This maximum $\alpha - 1$ is in turn negative for plausible values of the coverage rate α.

The expected values are

$$\mathbb{E}R = R_0 - (1-\alpha)\mathbb{E}\widetilde{x} - (1-\tau)(1+\lambda)\alpha\mathbb{E}^*, \tag{9}$$

$$\mathbb{E}R_{\text{NS}} = R_0 - (1-\alpha)\mathbb{E}\widetilde{x} - (1+\lambda)\alpha\mathbb{E}^*. \tag{10}$$

The derivatives are

$$\frac{\partial \mathbb{E}R}{\partial \alpha} = \mathbb{E}\widetilde{x} - (1-\tau)(1+\lambda)\mathbb{E}^* + (\alpha-1)\frac{\partial \mathbb{E}\widetilde{x}}{\partial \alpha}, \tag{11}$$

$$\frac{\partial \mathbb{E}R_{\text{NS}}}{\partial \alpha} = \mathbb{E}\widetilde{x} - (1+\lambda)\mathbb{E}^* + (\alpha-1)\frac{\partial \mathbb{E}\widetilde{x}}{\partial \alpha}. \tag{12}$$

The expected revenue can be positively affected by the increase of the insurance coverage if for example moral hazard is not too strong ($\frac{\partial \mathbb{E}\widetilde{x}}{\partial \alpha} > 0$ but limited) or if there is shielding ($\frac{\partial \mathbb{E}\widetilde{x}}{\partial \alpha} < 0$), while the underestimation of the risk is substantial ($\mathbb{E}^* < \mathbb{E}\widetilde{x}$). In that case, the farmers will receive more from insurance than what they paid on average because the damages paid by the insurance overcompensate the endogenous risk. In the same vein, a large subsidy can also cause this overcompensation. The opposite may be true. If farmers engage in shielding behavior but the insurance overestimates risk and/or the subsidy is small, the negative impact of insurance on revenues can even be higher than the effect of the loading factor in Equation (4).

This proposition is intended to give alternative (but not exclusive) explanations of why we observe an expected revenue increasing with coverage. The econometric analysis indeed measures such an increase.

Proposition 2

1. *If $\mathbb{E}R$ increases with respect to coverage, at least of the following conditions is true: (i) the subsidy is large, (ii) moral hazard is weak, (iii) farmers exhibit shielding behavior, (iv) insurance is underpriced.*

2. *If $\mathbb{E}R_{NS}$ increases with respect to coverage, at least one of the above conditions (except (i) which is irrelevant).*

The proof is a direct interpretation of Equations (11) and (12).

Variance. Remark that, in the case of mispricing, the revenue net of subsidies has the same variance as the revenue itself since the subsidy is not random.

$$\mathbb{V}R = \mathbb{V}R_{\text{NS}} = (1-\alpha)^2\mathbb{V}\widetilde{x}. \tag{13}$$

Therefore

$$\frac{\partial \mathbb{V}R}{\partial \alpha} = \frac{\partial \mathbb{V}R_{\text{NS}}}{\partial \alpha} = -2(1-\alpha)\mathbb{V}\widetilde{x} + (1-\alpha)^2\frac{\partial \mathbb{V}\widetilde{x}}{\partial \alpha} \tag{14}$$

The first term is the direct variance reduction due to insurance. The strongest effect is for low levels of insurance (α small). The second term comes from the indirect impact of α on the behavior of the farmer. The variance could increase or decrease, depending on moral hazard and/or shielding.

The impact of insurance on the second moment is not as clear intuitively as the impact on the first moment.

4.3 Heterogeneity analysis

The benefits of insurance are heterogeneous and explainable by individual characteristics. Smaller and larger farms, for example, will not derive the same benefits from insurance because their risk profiles, shielding and moral hazard behaviors might differ. The type of crops, the degree of specialization, macro-choices like being organic, etc., also matter. We provide a simple framework to analyze this heterogeneity.

We start with with the case of actuarial pricing with loadings and subsidies, as in Equations (4) and (5).

The variable X represents some observable characteristics. We then just calculate the cross-derivatives:

$$\frac{\partial^2 \mathbb{E}R}{\partial\alpha\partial X} = (\tau(1+\lambda)-\lambda)\frac{\partial \mathbb{E}\widetilde{x}}{\partial X} + (\tau(1+\lambda)\alpha - (1+\lambda\alpha))\frac{\partial^2 \mathbb{E}\widetilde{x}}{\partial\alpha\partial X} \tag{15}$$

and

$$\frac{\partial^2 \mathbb{E}R_{\text{NS}}}{\partial\alpha\partial X} = -\lambda\frac{\partial \mathbb{E}\widetilde{x}}{\partial X} - (1+\lambda\alpha)\frac{\partial^2 \mathbb{E}\widetilde{x}}{\partial\alpha\partial X}. \tag{16}$$

Let's consider that X measures size. The first derivative $\frac{\partial \mathbb{E}\widetilde{x}}{\partial X}$ is likely to be positive. If we believe in moral hazard, then the two cross-derivatives in the right-hand side could be positive if moral hazard increases with size. The effect would be that people with more insurance lose money when they get insured, and all the more so as they are bigger. On the contrary, if we believe in shielding, the first derivative is still positive while the cross derivative could be negative. It is unclear from the theory alone which effect dominates which, which means that empirical proof is needed. The results are discussed in Section 7.2.

5 Estimation strategy

5.1 Overview

We wish to estimate the impact of insurance on revenue. In principle, the impacts are mainly a decrease in expected value, due to the loadings, and a decrease in variance, due to the indemnities. Besides, subsidies increase revenue, and this effect is easy to control for with accounting data. The last channel is the change in agronomic methods that insurance induce. If having insurance implies a discrete move along the production frontier, knowing what input changes are concerned will be necessary to evaluate globally the impact of insurance (or the absence thereof). For example insurance could be associated with more fertilizers.

The issue we address now is empirical. Insurance is chosen by farmers, not randomly assigned. We need to correct for the biases this causes using instrumental variables. We will argue that crop insurance subsidy rates are a valid instrument. Besides their value as an instrument, crop insurance subsidies are a common policy tool whose efficiency at moving subscription rate we wish to estimate. What is the effect of subsidies in terms of propensity to insure? What is their effect in terms of revenue for those switching to insurance?

Heterogeneity among farmers being the norm, effects cannot be summed up as mere means. We have sufficiently many controls to explain differences via observable variables. Yet, unobservable heterogeneity matters, especially because responses to crop insurance subsidies are affected by selection biases. The heterogeneity of treatment effects can be estimated conditional on observables and on the propensity score. The Marginal Treatment Effects, which are literally functions, are the basic ingredients of all measures of treatment effects generally used (ATE, ATT, ATU, IV, OLS, etc.; see e.g. HECKMAN and VYTLACIL, 2005; HECKMAN and VYTLACIL, 2007). They can be used to evaluate counterfactual policy with an unmatched degree of precision, at least in principle.

Our final exercise is a cost-benefit analysis of crop insurance policies. We propose alternative counterfactual policies and we use our previous estimates to track material and financial consequences of the changes in insurance adoption.

5.2 Average impact of insurance subscription on revenues

A fixed effect model estimates revenues with and without insurance subsidies from insurance status (D), inputs and individual characteristics $\boldsymbol{X}$, and climate shocks $\boldsymbol{\Lambda}$. The squared error term is also regressed on the same variables (moment-based approach à la ANTLE, 1983) to estimate the effects of these on variance:[9]

$$R_{it} = \alpha + \beta_{11} D_{it} + \boldsymbol{X_{it}}\beta_{21} + \boldsymbol{\Lambda_{it}}\beta_{31} + \boldsymbol{\Lambda_{it-1}}\beta_{41} + \theta_i + \theta_t + \epsilon_{it}, \tag{17}$$

$$\epsilon_{it}^2 = \alpha' + \beta_{12} D_{it} + \boldsymbol{X_{it}}\beta_{22} + \boldsymbol{\Lambda_{it}}\beta_{32} + \boldsymbol{\Lambda_{it-1}}\beta_{42} + \theta_i' + \theta_t' + \epsilon_{it}', \tag{18}$$

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 row vector of individual characteristics (including subsidies and dummies for cattle/greenhouses, see Subsection 6.2 for a complete list), $\boldsymbol{\Lambda_{it}}$ the row vector of climate variables (sum of out-of-bound hot and cold Growing Degree-Days—or OOB—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.

The inclusion of OOBs stems from the agronomic literature the effect of climate on revenue. They are the best synthetic indicators of the weather experienced by crops, as is detailed in Section 6. The production variables include costs for gas/oil, costs for crop protection products, water used for irrigation, total work hours, total surface area of the farm and production subsidies. The choice to put all the variables in log is justified by the high heterogeneity of the sample for both the dependent and independent variables, as shown in the descriptive statistics in Subsection 6.2. Furthermore, this choice is in line with the rest of the literature (DI FALCO et al., 2014; WANG, RE-

[9]While this method can be used to analyze the third order (i.e. skewness), our preliminary analyses show that the impact of insurance on the skewness of revenue appears insignificant.

JESUS, and AGLASAN, 2021). For negative values, we add to the entire sample the minimum value + 1 before taking the natural log. We show in the Appendices F.1 and F.2 two alternative specifications where the dependent variable is expressed in direct form or with the Inverse Hyperbolic Sine (IHS) function.

This OLS specification is likely to yield biased effects, mainly due to the omitted variables and endogeneity. It is entirely plausible that farmers with a better business sense subscribe more to crop insurance, which would cause an overestimation of the insurance coefficient. Beliefs and the social environment, while partially taken into account through fixed effects, might also play a role in both revenue and insurance subscription. As the analysis of heterogeneity will clarify, the selection is in the unexpected direction: farmers with the highest potential benefit of insurance are underrepresented among the insureds. 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 dependent variable other than through the instrumented variable.

In the instrumental variable approach, we control for inputs and individual farm characteristics, climate (current and past), and we include individual and time fixed effects, as in the OLS specification

$$D_{it} = \alpha + \beta_{11}\mathbb{E}(S|t,c) + \boldsymbol{X_{it}}\beta_{21} + \boldsymbol{\Lambda_{it}}\beta_{31} + \boldsymbol{\Lambda_{it-1}}\beta_{41} + \theta_i + \theta_t + \epsilon_{it}, \tag{19}$$

$$R_{it} = \alpha' + \beta_{12}D^*_{it} + \boldsymbol{X_{it}}\beta_{22} + \boldsymbol{\Lambda_{it}}\beta_{32} + \boldsymbol{\Lambda_{it-1}}\beta_{42} + \theta'_i + \theta'_t + \epsilon'_{it}, \tag{20}$$

$$\epsilon'^2_{it} = \alpha'' + \beta_{13}D^*_{it} + \boldsymbol{X_{it}}\beta_{23} + \boldsymbol{\Lambda_{it}}\beta_{33} + \boldsymbol{\Lambda_{it-1}}\beta_{43} + \theta''_i + \theta''_t + \epsilon''_{it}, \tag{21}$$

with D^*_{it} the first-stage prediction of D_{it}.

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 learning and belief in persistent effects. A farmer hit by a flood in year $t-1$ might have built a tarp to protect the crop. This tarp remains in year t and affects revenues as well as the insurance strategy. Secondly, weather shocks are persistent beyond a year. In terms of impact, 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). Current weather variables explain partly current revenue and are included in the second-stage accordingly. For consistency, they are included in the first stage (insurance decision) (WOOLDRIDGE, 2010).

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 for us because that they do not affect farmers' revenues through any other channel than insurance. Discrete 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 collinearity, 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 annual average subsidy rate for insurance for each type of crops as an instrument. Insurance subsidy rates are decided at the EU level every year (Ministère de l'Agriculture, 2022b) since 2015 (and at the French level beforehand) and are differentiated between crops (Section 6).

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 former condition, 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 the instrument is sufficiently correlated with the variable of interest (Andrews, Stock, and Sun, 2018). While the exclusion restriction cannot be tested, we use the fact that our policy-based instrument is decided at the national level (EU level after 2015) before the beginning of the contracting season. It is unlikely that it would affect revenue in any other way than through insurance take-up.

One worry is that we do not know exactly how subsidy rates are determined. If farmers had a way of influencing the decision process, they might push to increase insurance subsidies for their specific crops. If that were the case, we would be capturing the impact of influence rather than insurance. However, because we are also using fixed effects (both farmer and year), this would only matter if the influence of farmers of a specific crop 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 to a significant extent.

"All included". Finally, despite the two-way fixed effects, the controls and the instrument, a simultaneity bias may remain. Indeed, insurance decision is not taken in a vacuum but as part of a broader protection strategy. For example, it is plausible that farmers substitute insurance for pesticides, or, contrarywise, that the protection insurance offers pushes farmers to produce at a higher scale, with more inputs overall, protective ones included. In the latter case, insurance would go with more expenditures in fertilizers and pesticides. If that were the case, then the estimated impact of insurance on revenue would include that substitution/complementarity effect. In that sense, pesticides would be a "bad control," and including it would lead to that simultaneity bias, while leaving it out leads to an omitted variable bias. In fact, with a single instrument, not including the control but being aware that there is a simultaneous choice is the best strategy.[10]

Our theoretical model and our interpretations always include the full behavioral impact of insurance. In other words, we do not want to only look at the benefits of insurance "everything else being equal," but as those benefits including any behavioral change that may occur (the global effect). The coefficients of the IV/LATE capture this global effect, but it is important to keep in mind that we cannot split it into substitution and pure effect to analyze it further with our current strategy. We examine the behavioral changes associated with insurance later in Section G.

5.3 Heterogeneity

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 theoretical framework through Equations (15) and (16), this would mean that Equation (4) might yield different results depending on the observables $\boldsymbol{X}$. The aim of this subsection is to identify the critical variables. Looking into the determinants of insurance subscription, we can check whether the criteria that influence insurance decisions are the same as those determining insurance benefits. They are not.

The insurance decision. To estimate the variables of interest, we run a Probit regression with fixed effects, using the same production and weather variables as in the base framework, this time to assess their effect on the probability to take out crop 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:

$$P(D_{it}=1) = \phi(\mathbb{E}(q|c,t), \boldsymbol{X_{it-1}}, \boldsymbol{\Lambda_{it-n}}, R_{it}, \theta_t, \epsilon_{it}), \tag{22}$$

$$P(D_{it}=0|D_{it-1}=1) = \phi(\mathbb{E}(q|c,t), \boldsymbol{X_{it-1}}, \boldsymbol{\Lambda_{it-n}}, R_{it}, \theta_t, \epsilon_{it}), \tag{23}$$

$$P(D_{it}=1|D_{it-1}=0) = \phi(\mathbb{E}(q|c,t), \boldsymbol{X_{it-1}}, \boldsymbol{\Lambda_{it-n}}, R_{it}, \theta_t, \epsilon_{it}), \tag{24}$$

with D_{it} the dummy for crop insurance. The production and climate variable are the same as in Equation (17). We additionally include R_{it} as the revenue variable (EBITDA) and $\mathbb{E}(q|c,t)$ as the price of insurance in the region by crop.[11]

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

[10]As is standard, we still provide a robustness check in the Appendix (Table 26), which shows the results when controls are not included, which confirms that the main estimates are unchanged.

[11]See Appendix B.3 for the construction of the price variable

of these determinants is not relevant to the research question and does not necessitate the use of an IV. The current climate is excluded as it is not a part of the insurance decision.

Heterogeneous benefits of insurance. By abuse, we call $\boldsymbol{X}$ the main variables identified in the previous subsection. We can then run the same regression as for Equations (19)-(21) (using the IV) but this time with an interaction term between our variable of interest and insurance take-up. This gives us a regression that we can interpret through the cross derivative given in Equation (15). 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

$$D^*_{it} = \alpha + \beta_{11}\mathbb{E}(S|t,c) + \boldsymbol{X_{it}}\beta_{21} + \boldsymbol{\Lambda_{it}}\beta_{31} + \boldsymbol{\Lambda_{it-1}}\beta_{41} + \theta_i + \theta_t + \epsilon_{it}, \tag{25}$$

$$(D^*_{it}\boldsymbol{x_{it}})^* = \alpha + \beta_{12}\mathbb{E}(S|t,c) + \beta_{22}\mathbb{E}(S|t,c)\boldsymbol{x_{it}} + \boldsymbol{X_{it}}\beta_{32} + \boldsymbol{\Lambda_{it}}\beta_{42} + \boldsymbol{\Lambda_{it-1}}\beta_{52} + \theta'_i + \theta'_t + \epsilon'_{it}, \tag{26}$$

$$R_{it} = \alpha'' + \beta_{13}D^*_{it} + \beta_{23}(D^*_{it}\boldsymbol{x_{it}})^* + \boldsymbol{X_{it}}\beta_{33} + \boldsymbol{\Lambda_{it}}\beta_{43} + \boldsymbol{\Lambda_{it-1}}\beta_{53} + \theta''_i + \theta''_t + \epsilon''_{it}, \tag{27}$$

$$\epsilon''^2_{it} = \alpha'' + \beta_{14}D^*_{it} + \beta_{24}(D^*_{it}\boldsymbol{x_{it}})^* + \boldsymbol{X_{it}}\beta_{34} + \boldsymbol{\Lambda_{it}}\beta_{44} + \boldsymbol{\Lambda_{it-1}}\beta_{54} + \theta'''_i + \theta'''_t + \epsilon'''_{it}, \tag{28}$$

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 changes. Formally, we can differentiate Equation (27) with respect to D to get

$$\frac{\partial \mathbb{E} R_{it}}{\partial D} = \beta_{13} + \beta_{23}\boldsymbol{x_{it}}. \tag{29}$$

A negative β_{23} would mean that increasing the characteristics in $\boldsymbol{X}$ 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 (15).

5.4 The MTE framework

Many summaries, reinterpretations, parameterizations and econometric implementations (comprising software packages) of the MTE have been proposed in the literature. Our aim is to limit our presentation to the essential notions and notation as well as to our restrictions. We will show the parametric and semi-parametric versions we use to clarify our strategy. Our focus is definitely on the exploitation of the unique rich dataset that we have.

Selection model. The estimates of subscription probabilities and the IV methods are useful to explore the heterogeneity of responses to the incentives to crop insurance. The selection model we use now connects tightly these two approaches. The model distinguishes the outcome R^s depending on whether the farmer is treated ($s = 1$) or not ($s = 0$), the effects depending on observable variables $\boldsymbol{X}$ and unobservables $\boldsymbol{U_1}$ and $\boldsymbol{U_0}$. The treatment is taken depending on a propensity score equation that contains both the observables $\boldsymbol{X}$ and an instrument:

$$R^1 = \mu_1(\boldsymbol{X}) + \boldsymbol{U_1}, \tag{30}$$

$$R^0 = \mu_0(\boldsymbol{X}) + \boldsymbol{U_0}, \tag{31}$$

$$D^* = \mu_D(\boldsymbol{Z}) - V, \text{ where } D = \mathbb{1}[D^* > 0] = \mathbb{1}[\mu_D(\boldsymbol{Z}) > V]. \tag{32}$$

D^* the latent desire to take up insurance that depends on $\boldsymbol{Z}$ ($\boldsymbol{X}$ plus an instrument) and an unobserved V. We observe only $R = DR^1 + (1-D)R^0, D, \boldsymbol{X}, \boldsymbol{Z}$, but not $\boldsymbol{U_1}, \boldsymbol{U_0}, V$. From Equation (32), a monotonous transformation gives $P(\boldsymbol{Z})$ (propensity scores as quantiles of $\boldsymbol{Z}$, a particular quantile being noted p) and U_D (quantiles of V, a particular value is referred to as resistance to treatment).

When insurance subsidies are low, only the farmers that had relatively low "resistance" to insurance are getting insured, and as insurance subsidies increase, more reluctant farmers get insured as well. Remark that the decision is not exactly directed by the outcomes. VYTLACIL (2002) shows the equivalence between this model and the LATE approach that assumes independance and monotonicity. Heckman and Vytlacil had proven earlier that with this generalized Roy model, under hypotheses we show later, the most elementary pieces of information that can be identified and estimated are the Marginal Treatment Effects, or MTE. This notion measures the expected effect of the treatment conditional on X and p:

$$\text{MTE}(\boldsymbol{X}, p) \equiv \mathbb{E}(R^1 - R^0 | \boldsymbol{X}, U_D = p). \tag{33}$$

This is a definition: the MTE gives the expected treatment effect for an individual with resistance to treatment p (as if we knew at what quantile he is situated), or equivalently, with propensity p and who is exactly indifferent between being insured or not.

The interest of the estimated MTE becomes clearer when one considers that all standard measures and estimands of treatment effects are weighted averages of the MTE, the weights being recovered from the principles driving the said measures and estimands. This has been proven and exposed in several papers by HECKMAN and VYTLACIL (e.g. 2005). In other words, the IV as presented before is a weighted average of the MTE, as we shall see. The literature on the LATE by IMBENS and ANGRIST (1994) has made clear that the IV estimate is an average over a particular population: those on which the instrument has had an effect. The MTE are the most elementary (or atomic) effects one can expect to identify in that context.

Most important for applications, having the MTE on hand enables the exploration of varied and justifiable counterfactuals. The effects of alternative policies can be explored via the predictive power of the model.

HECKMAN and VYTLACIL (2005) have shown identification strategies. The MTE can be retrieved empirically using the identity

$$\text{MTE}(\boldsymbol{X}, p) = \frac{\partial \mathbb{E}(R | \boldsymbol{X}, p)}{\partial p}, \tag{34}$$

where the expectation on the right-hand side $\mathbb{E}(R|\boldsymbol{X}, p)$ is directly estimable, while the derivative can be numerically calculated either because p is sufficiently "dense" at the point of interest, or because holes are filled with the help of a parametric or semi-parametric specification. We are in the latter case in this study.

Hypotheses. We use again the average subsidy rate over crops and year as the instrument. The effect of the instrument is assumed to be monotonic: as (crop-year) subsidies increase, more farmers are going to take up insurance. We argue that the assumptions required for the MTE to be unbiased are met.

Despite a highly endogenous treatment (insurance take-up), our instrument is strong enough to yield interpretable estimates. The average insurance subsidy rate is not correlated to the 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, the subsidies are typically given to farmers at the end of the growing season (MINISTÈRE DE L'AGRICULTURE, 2022b), whereas insurance is paid at the beginning or as monthly installments. Furthermore, the subsidy bases are recalculated every year at the national level (before 2015) and the European level (2016 and onwards), which means that farmers could hardly anticipate or influence this choice. This leads us to think that the data satisfy the "as good as random" assumption, i.e. $\boldsymbol{U_0}, \boldsymbol{U_1}, V \perp \boldsymbol{Z}|\boldsymbol{X}$.

The second requisite is, as argued in HECKMAN and VYTLACIL (2007), that the propensity score 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 E that show that $P(\boldsymbol{Z})$ is indeed estimated on the (almost) full spectrum with enough variability. Furthermore, our instrument takes over 100 different values, varying from 0 for some crops (namely vines) up to 45% depending on year and crop.[12] HECKMAN and VYTLACIL (2007) shows that the MTE curve can still be estimated in that context over the common support either because we have a parametric form, or because of the separability on the semi-parametric form.

The other assumption we need to make here is that there are no violations of monotonicity between our point estimates, which we think is fairly weak. Discontinuities due to left-digit attention found in behavioral economics could matter, but the density of rates is not sufficient to detect such moves, if they exist.

Specification choice and estimation. Following BRINCH, MOGSTAD, and WISWALL (2017, p. 999), we adopt the auxiliary assumption

$$\mathbb{E}(R^j|\boldsymbol{X}, V) = \mu_1(\boldsymbol{X}) + \mathbb{E}(\boldsymbol{U_j}|V), \quad j = 0, 1. \tag{35}$$

[12] Nominal and actual subsidy rates differ because farmers may take up extended guarantees, whereas only standardized tranches are subsidized. We don't observe directly the details of the actual contracts, but only the monetary transfers.

We take linear forms for $\mu_1(\boldsymbol{X}) = \boldsymbol{X}\beta_1$ and $\mu_0(\boldsymbol{X}) = \boldsymbol{X}\beta_0$. Then the MTE is additively separable in $\boldsymbol{X}$ and p:

$$\text{MTE}(\boldsymbol{X}, p) = \boldsymbol{X}(\beta_1 - \beta_0) + \underbrace{\mathbb{E}(\boldsymbol{U_1} - \boldsymbol{U_0}|V = p)}_{k(p)} \tag{36}$$

We perform the regression from Equation (36) using first a parametric approach for $k(p)$ (a degree-4 polynomial, as is standard in the literature, for $K(p)$ where $K'(p) = k(p)$), and second using a semi-parametric approach (Local Instrumental Variable, or LIV) laid out in ANDRESEN (2018) and first explained in HECKMAN and VYTLACIL (2007). In order to fulfill the MTE exclusion restriction, the revenue variable is net of insurance subsidies.

This approach allows us to estimate a local impact of the instrument on the propensity score conditional to the levels of the instrument and observable characteristics.[13] The included controls are the same as in Section 5.2, including year fixed effects. We interpret mainly the semi-parametric regression in the results. We also bootstrap the regression with 100 iterations to provide confidence intervals and cut the common support at the ends due to lack of data (see Figure 11 in the Appendix).

Finally, note that we do not exploit the panel dimension of our data in the estimation strategy for the MTE. The fixed effects present in Equation (19) are therefore not reused in Equation (36) (although we do control for year fixed effects). Indeed, they would drastically reduce the matching potential of the first stage, while adding little value in terms of identification. The MTE framework relies on unobserved heterogeneity, which would be heavily restrained by the use of individual fixed effects. More specifically, it uses a propensity score matching methodology in the first-stage, meaning that using a within estimator at the individual level would only allow matching in time for the same individual (i.e. an individual can only be matched with themselves). This almost completely negates the crop dimension of our instrument and only leaves the temporal variation of the subsidy rate to compute resistances to treatment. These benefits could be worth the cost if individual fixed effects were needed, but they are not: the potential individual endogeneity from unobserved characteristics that the fixed effects would correct (for example good managers earning more revenues and being more likely to insure after a change in subsidies) is already accounted for in the MTE framework through the unobserved resistance to treatment.

6 Data

To perform this large-scale analysis across mainland France over a 20-year period, we produce a unique and granular dataset composed from individual data on farmers, including agronomic and financial variables, weather data at a lat $0.1° \times$ long $0.1°$ resolution and administrative data for climate disasters.

[13] Specifically, we estimate $K(p)$ using the Robinson estimator (ROBINSON, 1988). For a summed up approach and the commands used, see VERARDI and DEBARSY (2012).

Refining WANG, REJESUS, and AGLASAN (2021)'s methodology, we create climate 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.

6.1 Construction of the dataset and key variables

Farm accounting, financial and agronomic data. The financial, input and insurance data of farms comes from the French "Réseau d'Information Comptable Agricole" (RICA, 2022). This is an annual survey-based panel dataset containing 17,743 individual firms observed over the 2002-2021 period and providing individual accounting and financial data (administrative data from farm accounts such as EBITDA, namely Earnings Before Interest, Taxes, Depreciation, and Amortization) as well as agronomic data (farm structure, irrigation, geographical data, fertilizers, etc.). This rich national database is produced and directly managed by the French Government. It is part of the Farm Accountancy Data Network (FADN) 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.[14]

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. EBITDA tracks the farm's performance before any policy change (taxes), which provides a more accurate estimate of the impact of climate change and insurance on revenues than a simpler variable such as operating profit. Additionally, we build EBITDA net of all insurance subsidies.[15] Besides, we build a dummy for crop insurance (1 if the farm is insured in a given year, 0 otherwise) by considering a farm insured if high enough premiums are paid.[16]

Weather 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 (about 6 km in France) for the 1950-2022 period. The data comes from 82 institutions and 22,600 weather stations (Figure 7 in Appendix A). As the station density in France is not large enough to

[14]Access to the RICA is restricted and confidential, and only summary-level data can be extracted and presented in this paper. However we have access to confidential data and are not limited in the possible treatments of the database. Our access was authorized by the French secrecy committee and required written commitments. Our access is managed by the Secure data hub ("Centre d'accès sécurisé à distance"), which operates in a similar way to the center to the The Federal statistical research data center in the United States or the Secure research service in the United Kingdom. Other authorized scholars can replicate our study at the cost of a relatively simple application and some delays.

[15]The MTE framework requires that we use the second indicator (net of insurance subsidies) to satisfy the exclusion restriction (Subsection 5.4).

[16]We consider as insured the farmers paying more than 20 €/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.

collect all the data at such a granular level, a prediction model (reanalysis) is employed by Copernicus to fill the missing data.[17]

We extract temperature data and compute their daily average for the longitudes, latitudes and time ranges relevant to our study. Using the "Base des Codes Postaux" (zip codes) from the French Government (COMMUNES DE FRANCE, 2020), which provides the latitude/longitude coordinates for the center of each French municipality (French "commune"), we then match each weather observation to a municipality with a least squares method. Finally, we match this local temperature variable to the municipality zip codes of each farm provided in the RICA database.

Second, the data on droughts and floods comes from Caisse Centrale de Réassurance (CAISSE CENTRALE DE RÉASSURANCE, 2023), which is an 100%-state-owned reinsurance company. This reinsurance company collects and provides data relative to the interministerial orders recognizing the state of natural disaster, specifying the period of the disaster, the concerned municipalities and the involved natural hazard.[18] We build discrete variables that count the number of floods or droughts which hit the municipality where the farm is located in a given year. Again, we match these disasters data to the municipality zip codes of each farm provided in the RICA database.

Building climatic agronomic indicators. Our choice of climatic indicators relies on the agronomic literature, as common statistical methods to identify temperature extremes (e.g. top 5% of temperatures or deviation from previous means) are not relevant to analyze the non linear effect of temperature on plant growth (SPINONI, BARBOSA FERREIRA, and VOGT, 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).[19] Our extreme temperature indicators are inspired by the index of Growing Degree Days (GDDs) (LUO, 2011; BLANC and SCHLENKER, 2017; KORRES et al., 2016; HORTON, 2018). This index represents heat accumulation and captures the non-linear effects of temperatures on plants over the growing season. Plants do not grow if the mean temperature over a day T is below a certain threshold T_c^b, which depends on the crop type c, and slow their growth above a certain upper threshold T_c^u. GDD is an index of heat accumulation, and the plant needs a certain amount of accumulated heat to grow ($\text{GDD}_c^{\text{opt}}$). These thresholds and limits are different depending on the plant and can be found in the literature (Table 1).

Inspired by the concept of Growing Degree Days, we refine the methodology used in WANG, REJESUS, and AGLASAN (2021) to create an aggregated indicator that capture the heterogeneous and non-linear effects of temperature on various types of crops. For our purposes, we compute

[17] See the dataset documentation in BOOGAARD et al. (2022) for the detailed process.

[18] After an event, the French government decides whether the event is a natural disaster and for what period and municipality. The decision relies on the conclusions of an interministerial commission, which analyzes the phenomenon on the basis of scientific reports. Insured households and firms (including farms) can benefit from the insurance compensation for natural disasters only if an order is published for the event concerned. This compensation covers buildings and furniture and is completely separate from crop insurance.

[19] Not 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), who do not find a negative impact of cold temperatures on farmers' revenue.

Type of crops	T_c^b	T_c^u	$\text{GDD}_c^{\text{opt}}$	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, BARBOSA FERREIRA, and VOGT (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)

Table 1: Agronomic parameters for GDD and OOB computations

a variable of interest for temperature other than GDD: the sum of out-of-bound GDDs (OOB) for cold and hot temperatures, as defined in SCHLENKER, HANEMANN, and FISHER (2007).[20] That is the sum of the differences between the temperatures below (above) T_c^b (T_c^u), which represents the sum of cold (hot) GDDs received throughout the year by the crops,

$$\text{OOB}_{stc}^C = \sum_d (T_c^b - T_{sdc})^+, \tag{37}$$

$$\text{OOB}_{stc}^H = \sum_d (T_{sdc} - T_c^u)^+, \tag{38}$$

with T is the average temperature during day d of year t, s the 6 km$\times$6 km square, c is the crop type, OOB^C the cold OOBs and OOB^H the hot OOBs. The indicator is then built for 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.

Instrumental variable. We use the national subsidy rate per crop and per year as an instrumental variable. Every year, insurance subsidy rates are decided at the EU level (MINISTÈRE DE L'AGRICULTURE, 2022b) since 2015 and at the French level beforehand.[21] They are differentiated between crops. The official documents for insurance subsidies are available since 2015 only and they provide the share of the crop insured value which is subsidized. This is why we compute the actual insurance subsidies paid to farmers as the ratio of subsidies received to their insurance premiums. More precisely, we first sum the individual insurance subsidies received by farmers by year and crop type, then divide it by the sum of the individual insurance premiums paid by year and crop type (Equation 39). Then, we apply this ratio to every farmer, insured or not, with a given crop for a given year

$$S_{tc} = \frac{\sum_i \text{Sub}_{itc}}{\sum_i P_{itc}}, \tag{39}$$

[20]The classic way of computing GDDs can be found in Appendix B.1.

[21]The subsidies, as they were defined until 2022, were introduced in the 2009 reform. In the RICA database, subsidy data is only available since 2009. This is why all regressions using this instrumental variable are estimated over the period 2009-2022.

with S_{tc} the subsidy rate for year t and crop type c, Sub the actual subsidy received and P the premium paid. The crop type corresponds to the main activity of the farmer, which is actually divided into 17 categories according to the OTEX17 nomenclature (MINISTÈRE DE L'AGRICULTURE, 2010).[22]

There are significant annual variations of the subsidy rate (Figure 8 in Appendix B.2). Indeed, two sources of variation, respectively due to the annual modification of the insured value tranches for each crop or a possible budget revision (Section 8 and Appendix B.2), contribute to the variations of our instrument.

The actual subsidy rate, which is decided at the French/European level, constitutes a good exogenous IV, as farmers have more interest to insure if the insurance subsidies are high, and insurance subsidies only affect farmers' revenues through their impact on insurance subscription. After the 2015 reform, official documents detailing the base insured prices were published yearly. To ensure that our measure is consistent with the nominal rates, we perform a robustness test (Table 25 in Appendix F.4) using only the nominal rates on a reduced sample after 2015 and find similar results (albeit less significant) than in our base framework.

6.2 Summary statistics

Tables 2 and 3 show the summary statistics for all the variables used in the regressions. Our sample includes larger farms than other national sources. Indeed, the median surface area is 85 ha in our sample, whereas it is below 50 ha according to the French Institute of Statistics and Economic Studies (INSEE, 2020). It is very likely because we observe farmers in mainland France only, whereas other sources include the French overseas departments, where the average surface area of farms is 5 ha versus 69 ha in mainland France in 2020 (MINISTÈRE DE L'AGRICULTURE, 2022c).

Very heterogeneous farms. As expected with farm data, the sample is highly heterogeneous, with some farms 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 example) and the heterogeneity in inputs (standard deviation 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 contrary, tomatoes grown in greenhouses require a lot more inputs to grow. For EBITDA, stock variations are the main cause of negative values.[23]

The hot and cold OOBs also exhibit a lot of variations, which is normal considering the various

[22] To be classified in an OTEX, a farm needs to use at least two thirds of its surface to produce one type of crops (i.e. winter cereals, fruits, etc.). Farms that do not meet this criteria for any crop (i.e. no one crop occupies over two thirds of the total surface) are classified in the OTEX "Diversified."

[23] 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 self-consumption, which does not mean that their main activity lies in cattle.

	Mean	SD	Q1	Median	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	123,700
Insurance spending per ha (€/ha)	24.22	55.91	0.00	2.32	22.81	0.00	450.00	123,700
EBITDA incl. insurance subsidies (€)	85,670	87,450	35,930	64,184	110,311	-504,040	3,755,931	123,700
EBITDA net of insurance subsidies (€)	85,700	86,940	36,080	64,290	110,320	-504,040	3,755	122,039
Subsidy rate (year, crop)	8.40	9.38	0.00	6.34	15.51	0.00	46.58	123,575
Sum of cold OOBs across the year (°C)	51.42	52.27	15.79	34.87	68.55	1.00	582.41	87,357
Sum of hot OOBs across the year (°C)	1.06	0.39	1.00	1.00	1.00	1.00	37.38	87,357
Number of floods/year	0.05	0.23	0.00	0.00	0.00	0.00	6.00	123,700
Number of droughts/year	0.06	0.27	0.00	0.00	0.00	0.00	4.00	123,700

Table 2: Summary statistics for the main variables

	Mean	SD	Q1	Median	Q3	Min	Max	Count
Number of workers (hours equivalent)	3,922.07	4,262.48	1,600.00	3,200.00	4,600.00	45.00	216,158.00	123,700
Used agricultural surface (ha)	104.21	81.40	46.20	85.42	141.50	0.32	795.49	123,700
Specialization index (1=Highly specialized)	0.48	0.28	0.25	0.46	0.67	0.00	1.00	123,700
Subsidies received (€)	3,6949	30,564	15,750	30,834	50,784	0.00	1,106,312	123,700
Cattle dummy	0.39	0.49	0.00	0.00	1.00	0.00	1.00	123,353
Greenhouse dummy	0.02	0.15	0.00	0.00	0.00	0.00	1.00	123,700
Organic dummy (1 = at least partial)	0.03	0.17	0.00	0.00	0.00	0.00	1.00	123,700
Real costs for gas/oil (€)	6,744	6,592	2,519	4,890	8,835	0.00	172,891	123,700
Real costs for pesticides (€)	12,312	14,809	2,693	7,426	16,614	0.00	311,599	123,700
Agrotourism revenues	77.58	1,292.50	0.00	0.00	0.00	0.00	147,940.00	123,700
Debt	210,971	278,329	60,692	135,906	266,040	0.00	12,118,604	123,700
Rent	15,217	16,619	4,852	10,926	20,064	0.00	654,873	123,700
Main activity: Cereals	0.50	0.50	0.00	1.00	1.00	0.00	1.00	123,700
Main activity: Vine	0.12	0.32	0.00	0.00	0.00	0.00	1.00	123,700
Main activity: Mixed	0.32	0.47	0.00	0.00	1.00	0.00	1.00	123,700
Main activity: Fruits and vegetables	0.06	0.24	0.00	0.00	0.00	0.00	1.00	123,700

Table 3: Summary statistics for the control variables

needs of the plants and the highly heterogeneous climate in France.[24] 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 2002-2021 period.

Insurance subscription. The insurance subscription rate is relatively stable but with movements in and out (Table 4), its reaches 27% over the period, twice as what would be expected considering the national average of 13%. This high figure can be explained by the under-representation of very small farms (who are typically under-insured) in the sample.

We illustrate the geographical distribution of crop insurance take-up by drawing regional maps of insurance subscription rates and of probability to get hit by a flood or drought in a given year over the entire time sample over 2002-2021 (Figure 1). The lack of correlation between take-up and exposure is very apparent, with the largest take-up being by far the Île-de-France (Paris) region (72%), despite having a low risk exposure (7%). The most exposed region (Provence-Alpes-Côte

[24] Cold OOBs seem to be a lot more numerous than hot OOBs, 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).

Status of insurance compared with the previous year	Frequency	Percent
Kept insurance	43,794	47.88
Canceled insurance	3,256	3.56
Opted into insurance	4,050	4.43
Stayed uninsured	40,361	44.13

Table 4: Distribution of movements within the full sample

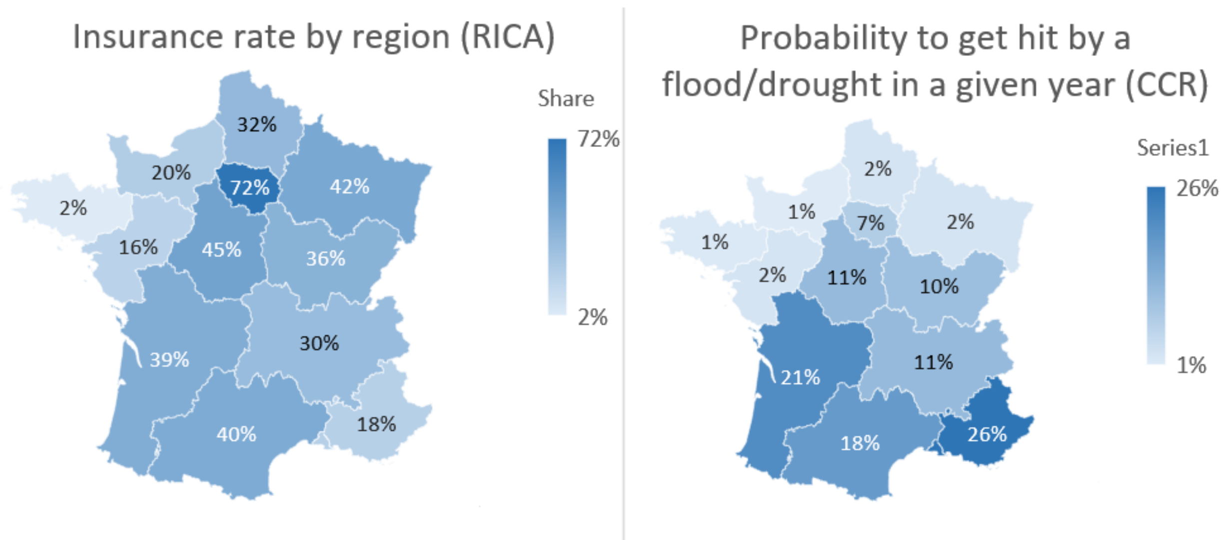


Figure 1: Map of insurance rate and risk exposure by region. Data sources: RICA, Caisse Centrale de Reassurance; Authors' production

d'Azur, 26%) also has a below average take-up (18%).

7 Results

7.1 Insurance increases revenues on average

The results of the IV estimation confirm that crop insurance is indeed an attractive choice. Subscribing to crop insurance increases revenues on average by about 23% with insurance subsidies, and 20% without insurance subsidies. Table 5 shows the results for the IV estimation (Equations 20-21), 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.[25]

The similar increases in revenue with and without insurance subsidies mean that there are behavioral impacts beyond the direct financial consequences to being insured. This would confirms a case of our model, specifically the plausibility of "shielding" as said in Proposition 1. The insured

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

farmers actually produce more than the non-insureds, the causality link being that they have more incentive to invest in their fields when they reckon that they will be compensated should climate shocks hurt them. Admittedly, advantageous mispricing could produced the same effect in terms of sign (Proposition 2), but not in terms of magnitude.

We test the hypothesis that indirect effects are dominant by interacting indemnities received with insurance status. The results of this regression can be found in Table 19. We find similar effects: the impact of insurance being mainly behavioral Remark that indemnities reduce revenues due to the fact that receiving indemnities means having been hit by a shock (which are not directly measured). On average, insurance does not fully cover the costs of shocks, which is consistent with the theory.

The increased production also explains why the impact of insurance on variance appears to be non-significant whereas a reduction could be expected. Insurance reduces variance others things equal, but if farmers cultivate riskier, higher-value crops when subscribing to insurance, before insurance revenues see their variance increase. The two effects compensates each other empirically.

Still, the coefficients are high. A 20% increase in revenues just thanks to insurance might appear large. Yet this number is actually lower than 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. Furthermore, a comparison with the estimates obtained with OLS in the Appendix C 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 that the selection bias the IV corrects is strong. It happens that the selection is unusual and contrarian. The MTE estimation exposed below makes this effect very apparent and shows the high heterogeneity in insurance benefits based on unobservable characteristics.

The effect of the weather variables on revenue first two moments is also in line with the agronomic literature. Cold temperatures appear to increase revenues, but only on the short-term, as an increase in past OOBs decreases revenues (-0.7% per OOB for the third lag). Indeed, in the short-term, cold OOBs 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 reflects the fact that meteorological droughts can be compensated with irrigation, while floods are inescapable.

The comparison of the IV 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 here is not sensible if we aim to understand how insurance actually impacts revenues.

7.2 Beyond the average

We show now that the propensity to insure increases with size, but the benefits of insurance don't.

	EBITDA with insur. subsidies		EBITDA w/out insurance subsidies	
	Mean	Variance	Mean	Variance
Dummy for crop insurance status (1=insured)	0.233*** (0.031)	-0.007 (0.010)	0.203*** (0.029)	-0.002 (0.009)
Cold OOBs (log)	0.007*** (0.001)	0.000 (0.000)	0.006*** (0.001)	0.000 (0.000)
L.Cold OOBs (log)	-0.001 (0.001)	-0.001* (0.000)	-0.002 (0.001)	-0.001** (0.000)
L2.Cold OOBs (log)	0.004*** (0.001)	-0.002** (0.001)	0.003** (0.001)	-0.001** (0.001)
L3.Cold OOBs (log)	-0.008*** (0.001)	0.001 (0.000)	-0.007*** (0.001)	0.000 (0.000)
Hot OOBs (log)	-0.025*** (0.005)	-0.001 (0.002)	-0.025*** (0.004)	0.001 (0.002)
L.Hot OOBs (log)	-0.011** (0.005)	0.001 (0.002)	-0.010** (0.005)	0.001 (0.002)
L2.Hot OOBs (log)	0.011*** (0.004)	-0.002* (0.001)	0.011*** (0.004)	-0.002 (0.001)
L3.Hot OOBs (log)	0.008* (0.005)	-0.004*** (0.001)	0.005 (0.005)	-0.003** (0.001)
Number of floods (log)	-0.011*** (0.003)	-0.000 (0.001)	-0.011*** (0.003)	-0.000 (0.001)
L.Number of floods (log)	-0.011*** (0.003)	0.001 (0.001)	-0.010*** (0.003)	0.001 (0.001)
L2.Number of floods (log)	-0.002 (0.003)	0.001 (0.001)	-0.002 (0.003)	0.001 (0.001)
L3.Number of floods (log)	-0.009*** (0.003)	-0.000 (0.001)	-0.008*** (0.003)	-0.000 (0.001)
Number of droughts (log)	0.003 (0.003)	-0.002 (0.002)	0.002 (0.003)	-0.001 (0.001)
L.Number of droughts (log)	0.007** (0.003)	-0.001 (0.001)	0.007*** (0.003)	-0.001 (0.001)
L2.Number of droughts (log)	0.002 (0.003)	0.001 (0.001)	0.002 (0.003)	0.001 (0.001)
L3.Number of droughts (log)	0.007** (0.003)	0.000 (0.001)	0.007** (0.003)	0.000 (0.001)
Subsidy rate (1st stage)	0.004*** (0.000)		0.004*** (0.000)	
Observations	51,142	51,142	50,567	50,567
Weak Ident. (F-test)	137.707	137.707	143.752	143.752
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

*: 90% significance, **: 95% significance, ***: 99% significance. Robust standard errors in parenthesis.

Table 5: 2nd stage IV log estimations for the impact of insurance on the revenue distribution

Propensity to be insured. Farm size is the strongest positive determinant of insurance subscription, both for the static and dynamic models. Among the proxies for size, turnover is the most significant.[26] Unsurprisingly, having experienced weather events is correlated with a higher probability of insurance.

Table 6 shows the results for Equations (22)-(24), that is the Probit regression for the determinants of static insurance subscription, exit and entry. 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.

Weather lagged values have a higher impact, which is expected, however the current values are also mostly significant. Farmers decide to get insured on the basis of their understanding of the risk. This understanding is determined in particular by their experience, and past weather is part of this experience. Current weather is not part of the experience, by definition, yet it enlarges the sample by which we approximate the experience of the farmer. Because this estimate is made more precise (just by adding a year), we get a better predictor of their behavior. This has nothing to do with prescience or adverse selection. Yet, in a changing world where expectations evolve fast, farmers may anticipate weather trends for the years to come insure with this prior.

Owning a greenhouse or raising cattle drastically decreases the probability to take up insurance, suggesting substitution behaviors: greenhouses are protections, and cattle management is very different from cultivation in terms of dynamics and market opportunities. Specialization has a positive impact: without a risk-reducing diversification, insurance is demanded. This result and our interpretation still requires explanations. Given that non-diversified farms are covered by the agricultural disaster scheme, which acts as a substitute for the private crop insurance system, the net effect could go the other way. However two points need to be considered. First, many farms in the sample are excluded from the agricultural disaster public scheme by virtue of their crop type (i.e. cereals). For them, the lack of diversification needs to be covered by private crop insurance. Second, even for farms covered by the agricultural disaster scheme, some losses like hail are not covered and they may want to insure as a complement (MINISTÈRE DE L'AGRICULTURE, 2022d). Crowding-out, though not to be excluded, is limited by nature.

Production subsidies have an observable negative impact on insurance subscription. Subsidies increase the revenue, but being unconditional on actual outcomes, they are not direct substitutes for insurance. Yet, subsidies may avoid the worst a farmer may experience: bankruptcy. This might suffice to partially crowd out insurance. This argument could be counterbalanced by the idea that the best managers are the best at grabbing all subsidies, for production or crop insurance. When observables are controlled for, the crowding out dominates on average.

Finally, the main activity appears to be a determinant as well, with vine growers having higher

[26] Turnover and surface area have a 0.4 correlation coefficient.

than average insurance rates.

Heterogeneous effects: Size matters. The results are highly surprising at first, as it seems that smaller farms benefit much more from insurance than larger farms, with the proportional (log) benefit being 3 times higher for farms in the first quartile compared to those in the last quartile (Figure 2). Because we investigate the impact of farm size on insurance benefits, we cannot use turnover as an independent variable (since EBITDA is the dependent variable and consists of net turnover including subsidies, taxes, etc.), and therefore use surface area as a proxy (0.4 correlation coefficient), controlling for productivity (turnover/ha) to ensure we capture a size effect. Table 7 shows the second-stage results of Equation (25) where the main interest is in the cross effect.

The paradox of size is the following: farmers who would benefit the most are those who subscribe the less. Things appear a little more subtle when we perform the heterogeneous regression on quartiles of size and specialization to uncover potential non-linear treatment effects. The results of this regression can be found in Figure 2. 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 (effect is not statistically different from zero), while the extremes (smallest and largest) farms benefit the most.

The story could be told like this. 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 this case, it would mean that small farms lack either the information required or the necessary managerial 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. This explanation from the demand side is the notion we explore further via the analysis of the MTE in Subsection 7.3.[27]

Heterogeneous effects: Diversification matters. The effect of diversification is non-linear, as seen in Table 8 and Figure 3. We find that the quadratic form captures these effects well, with a downwards slope until a specialization of about 0.6 (i.e. around the Q3, which can be interpreted as one crop taking up 60% of surface), followed by an upwards slope. Diversification can be seen as a protection strategy which would be a substitute for insurance take-up: different crops may grow at different times, they may be sensitive to different adverse shocks, etc. Highly diversified farms also have a more complicated insurance choice to design and implement.

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 specialized farms also benefit from insurance, they appear to already know that and have high subscription rates.

[27] Explanations from the supply side are worthy of interest. It could be that insurance contracts incurs higher up-front costs for smaller farms, due to a lower bargaining power or a lack of research on the market (i.e. information barriers), or simply because there are economies of scale and scope in insurance distribution. Note however that we control for observed heterogeneity, size in particular.

	(1) Static	(2) Exit	(3) Entry
Turnover (log)	0.216***	-0.429***	0.087
	(0.069)	(0.102)	(0.085)
Total work hours (log)	-0.001	-0.094***	-0.143***
	(0.024)	(0.023)	(0.021)
Total surface of the farm (log)	0.012	0.038	0.092***
	(0.028)	(0.025)	(0.024)
Greenhouse dummy	-0.379***	-0.060	0.077
	(0.141)	(0.101)	(0.099)
Cattle dummy	-0.282***	-0.031	-0.126***
	(0.025)	(0.028)	(0.025)
Mean real price of insurance (year, crop)	0.000	-0.000	-0.001***
	(0.000)	(0.000)	(0.000)
Organic agriculture dummy (1= at least partial)	0.282***	0.022	0.248***
	(0.075)	(0.082)	(0.071)
Real costs of crop protection products (log)	0.128***	0.022**	0.062***
	(0.019)	(0.010)	(0.014)
Agrotourism revenues (log)	-0.001	0.010	0.003
	(0.012)	(0.012)	(0.012)
Specialization index (1=Highly specialized)	0.837***	0.327***	0.075
	(0.069)	(0.070)	(0.065)
L.Cold OOBs (log)	-0.015**	-0.030*	0.082***
	(0.007)	(0.016)	(0.015)
L2.Cold OOBs (log)	-0.019***	0.056***	0.020
	(0.007)	(0.017)	(0.017)
L3.Cold OOBs (log)	0.026***	-0.000	-0.012
	(0.007)	(0.016)	(0.015)
L.Hot OOBs (log)	0.281***	-0.006	-0.070
	(0.040)	(0.068)	(0.079)
L2.Hot OOBs (log)	-0.081***	-0.023	-0.288***
	(0.029)	(0.059)	(0.067)
L3.Hot OOBs (log)	-0.153***	0.013	-0.194**
	(0.036)	(0.070)	(0.076)
L.Number of floods (log)	0.097***	-0.051	0.049
	(0.024)	(0.061)	(0.054)
L2.Number of floods (log)	0.079***	0.039	-0.081
	(0.025)	(0.058)	(0.061)
L3.Number of floods (log)	0.098***	0.002	0.028
	(0.026)	(0.060)	(0.058)
L.Number of droughts (log)	0.089***	0.027	0.041
	(0.024)	(0.055)	(0.054)
L2.Number of droughts (log)	0.042*	-0.038	-0.078
	(0.023)	(0.056)	(0.055)
L3.Number of droughts (log)	0.018	0.034	-0.055
	(0.025)	(0.057)	(0.055)
Main activity: Cereals (compared with fruits and vegetables)	-0.047	-0.100	-0.046
	(0.051)	(0.062)	(0.062)
Main activity: Vine (compared with fruits and vegetables)	0.114**	-0.098	0.038
	(0.055)	(0.067)	(0.062)
Main activity: Mixed (compared with fruits and vegetables)	-0.192***	-0.195***	-0.225***
	(0.056)	(0.066)	(0.068)
Constant	-3.440***	0.371	-2.914***
	(0.428)	(0.574)	(0.466)
Observations	52,637	52,637	52,637
Chi2	1683.440	261.296	640.587
Population average	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

*: 90% significance, **: 95% significance, ***: 99% significance. Robust standard errors in parenthesis.

Table 6: Probit results: The determinants of insurance subscription

Heterogeneous effects: Biggest benefits for cereals and vine. All the analyses performed above have been done on the entire sample of the FADN database in France (excluding pure cattle farmers). While crop type is controlled in every regression through the use of individual fixed effects, there may still be some heterogeneity that can add to the story.

We therefore perform our LATE regression from Equations (19)-(21) on the EBITDA including subsidies by dividing the sample into crop-type subsamples. The results can be found in Table 9. The positive impact of insurance on revenue appears to be entirely driven by cereals and vine. While the coefficient for vine is to be expected (grapes is a high value-added product with high risk), the result for cereal appears at odds with our previous analysis. Indeed, cereal farms are generally larger than average but also highly specialized: size and specialization pull in opposite directions. The net effect proves that specialization drives the impact of insurance.

To shed light on this issue, we perform the regression from Equation (25) by subsample (i.e. interacting size and insurance, but this time in crop subsamples). The results can be found in Table 10. They show that the size effect is also negative for cereal farmers. This means that smaller cereal farms benefit the most from insurance, which confirms our general results. Vinegrowers, on the other hand, exhibit the opposite trend.

Finally, the first stage results show that the elasticity to subsidies varies greatly depending on crop type. Diversified farmers, who have low benefits from insurance, are the ones who react the most to subsidies, while cereal and vine growers still have a positive reaction, albeit lower. Fruits and vegetable growers appear to have no significant reactions to subsidy changes. This adds to the argument that increasing subsidies is not the simple remedy to the problem undersubscription.

	EBITDA with insurance subsidies		EBITDA w/out insurance subsidies	
	(1)	(2)	(3)	(4)
Dummy for crop insurance status (1=insured)	1.186*** (0.249)	-0.147* (0.086)	1.042*** (0.223)	-0.074 (0.069)
Crop Insurance × Surface	-0.180*** (0.040)	0.024* (0.013)	-0.158*** (0.036)	0.015 (0.011)
Total surface of the farm (log)	0.154*** (0.016)	-0.012** (0.005)	0.148*** (0.014)	-0.007 (0.004)
Observations	51,142	51,142	50,567	50,567
Weak Ident.	23.176	23.176	24.142	24.142
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

*: 90% significance, **: 95% significance, ***: 99% significance. Robust standard errors in parenthesis. Columns (1 and (3) for mean and (2) and (4) for variance.

Table 7: IV estimations on revenue with surface area interactions

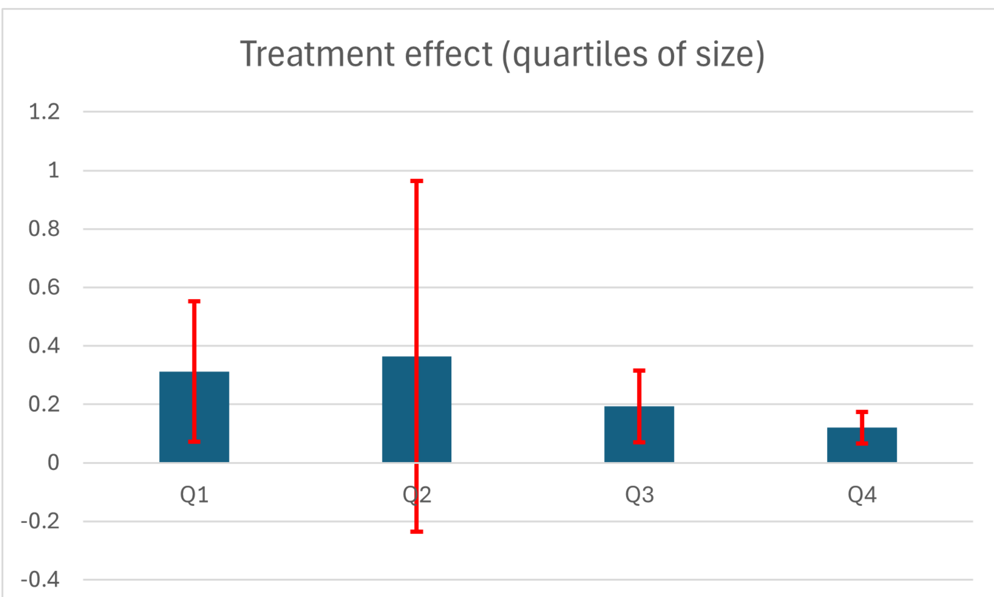


Figure 2: Treatment effect by quartile of size

Note: Red bars are the standard errors. Size corresponds to the surface area of the farm in ha. For example, the top 25% of the largest farms derive a significant treatment effect of 0.1, i.e. about a 10% increase in revenues.

	EBITDA with insurance subsidies		EBITDA w/out insurance subsidies	
	(1)	(2)	(3)	(4)
Dummy for crop insurance status (1=insured)	2.996**	-2.994	2.878**	-2.766
	(1.187)	(1.858)	(1.186)	(1.801)
Insurance × Specialization	-9.524***	9.693*	-9.201**	8.971
	(3.685)	(5.723)	(3.709)	(5.589)
Insurance × Specialization (squared)	7.662***	-7.385*	7.385***	-6.833
	(2.822)	(4.298)	(2.850)	(4.220)
Specialization index (1=Highly specialized)	2.063***	-2.704**	1.938***	-2.449**
	(0.763)	(1.170)	(0.750)	(1.119)
Specialization index (squared)	-1.828***	2.177**	-1.697***	1.960**
	(0.634)	(0.923)	(0.620)	(0.884)
Observations	51,142	51,142	50,567	50,567
Weak Ident.	2.510	2.510	2.299	2.299
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

*: 90% significance, **: 95% significance, ***: 99% significance. Robust standard errors in parenthesis. Columns (1 and (3) for mean and (2) and (4) for variance.

Table 8: IV estimations on revenue with specialization interactions

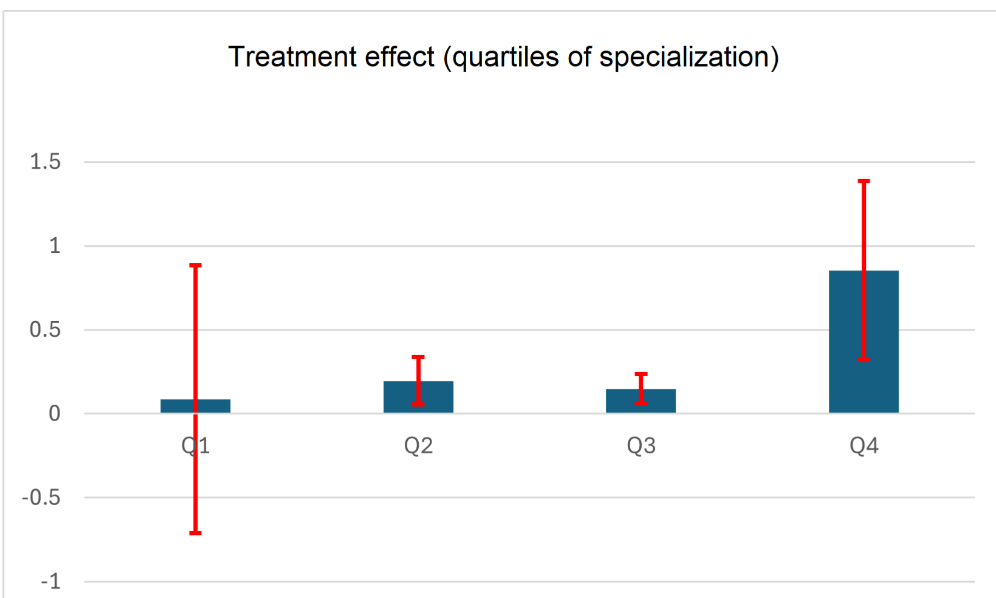


Figure 3: Treatment effect by quartile of specialization

Note: Red bars are the standard errors. Q4 corresponds to the highest quarter of the distribution, i.e. the least diversified farms; Q1 corresponds to the most diversified farms. For example, it appears that the 25% most diversified farms do not derive a significant effect from insurance.

	Cereals	Fruits and vegetables	Mixed	Vine
Dummy for crop insurance status	0.202***	-1.589	-0.461	0.391**
(1=insured)	(0.031)	(6.733)	(0.820)	(0.152)
First stage	0.016***	-0.023	0.052***	0.004***
(subsidies received on insurance rate)	(0.003)	(0.109)	(0.014)	(0.002)
Observations	27,390	2,072	15,810	4,820
Weak Ident.	86.630	0.057	0.410	13.704
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

*: 90% significance, **: 95% significance, ***: 99% significance. Robust standard errors in parenthesis.

Table 9: 2nd stage IV log estimations for the impact of insurance on the revenue distribution based on crop type

	Cereals	Fruits and vegetables	Mixed	Vine
Dummy for crop insurance status	1.694**	0.980	-8.645	-7.030*
(1=insured)	(0.679)	(10.439)	(27.206)	(3.797)
Subsidy rate× Surface	-0.154**	-0.275	0.779	0.849**
	(0.066)	(1.679)	(2.364)	(0.433)
Total surface of the farm (log)	0.172***	0.187	0.032	-0.625*
	(0.025)	(0.516)	(0.073)	(0.359)
Observations	27,390	2,072	15,810	4,820
Weak Ident.	8.049	0.057	0.040	2.273
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

*: 90% significance, **: 95% significance, ***: 99% significance. Robust standard errors in parenthesis. Columns (1 and (3) for mean and (2) and (4) for variance.

Table 10: IV estimations on revenue with surface area interactions by crop type

Effects	(1) EBITDA net of insurance subsidies (log)	(2) SE
ATE	0.149***	0.018
ATT	0.132***	0.014
ATUT	0.158***	0.026
LATE	0.079***	0.009
Observations	72,970	

Note: ATE = Average Treatment Effect, ATT = Average Treatment on the Treated, ATUT = Average Treatment on the Untreated, LATE = Local Average Treatment Effect.
Bootstrapped standard errors, *90% CI, **95% CI, ***99% CI

Table 11: Recovered estimators from the semiparametric MTE framework

7.3 MTE: Insurance benefits and unobserved resistance to treatment

The MTE is clearly U-shaped. 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. The semi-parametric estimates displayed in Figures 4 and 5, and the recovered effects for the mean are displayed in Table 11. 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, 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.

This nonmonotonicity is not standard. We complete with a look at the effects on variance despite the lack of statistical significance We propose the following explanation. Farmers at the center seem to be plainly risk averse. They deploy a variety of protections, which may or may not include insurance, to secure their revenues. Their risk taking hardly increases with insurance. 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. They select the treatment that is profitable, at the cost of taking more risk, to a limited extent though. Farmers on the right side are “bad managers”. They forego revenue, and even a slight reduction of the variance. Are they risk lovers, or less informed? Are they overwhelmed by the many administrative tasks, insurance being just more paperwork that is often eschewed? The latter is the most plausible.

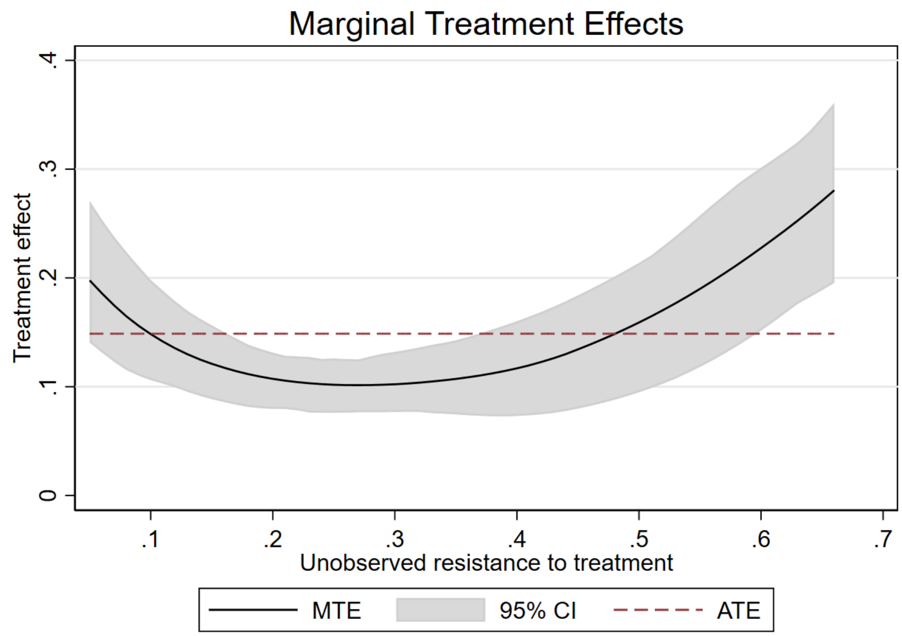


Figure 4: MTE curve for expected EBITDA with controls

Note: The left-side of the distribution describes farmers that are very inclined to insure (i.e. insure even with very low subsidies). Farmers who are the most resistant to treatment derive the highest benefit (0.28 or about a 28% increase in revenues). The gray outline corresponds to a 95% confidence interval obtained from the bootstrap estimation.

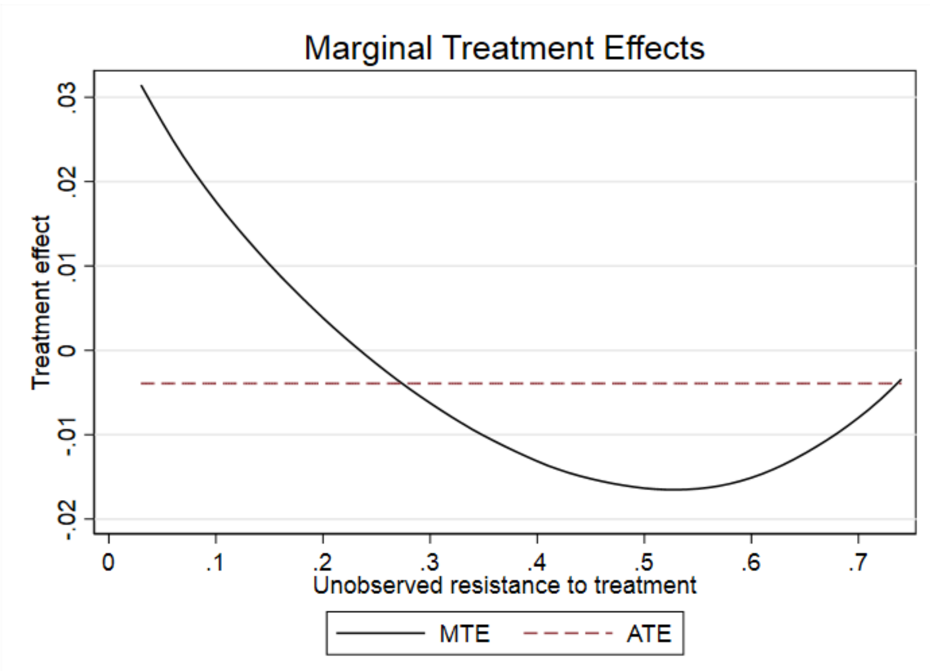


Figure 5: MTE curve for variance of EBITDA with controls

Note: The left-side of the distribution describes farmers that are very inclined to insure (i.e. insure even with very low subsidies). Farmers who are the least resistant to treatment see higher impacts of insurance on their variance (0.03 or about 3%). Confidence interval absent due to computation limitations (will be updated in a newer version)

8 Policy analysis

While an increase in insurance take-up is generally desirable, it might not be efficient to increase insurance subsidies indiscriminately. The MTE analysis shows that increasing insurance subsidies will benefits most those that are already insured and it will convert those with middling resistance whose gains are limited. High resistance farmers are relatively unaffected by insurance subsidies, despite having the highest benefit to be insured. The target is missed.

We design and parameterize alternative pro-insurance counterfactual policies. We employ the PRTE method and complement it with the MPRTE estimators. We suggest improvement in the design of incentive for insurance and discuss the limitations of our extrapolations.

8.1 Increasing the subsidy rate

Policy design. We use the MTE estimates to simulate an increased subsidy rate. The subsequent increase of the propensity scores causes an influx of individuals into insurance. This counterfactual analysis uses the notion of policy relevant treatment effects (PRTE) as exposed for example in CARNEIRO, HECKMAN, and VYTLACIL (2011). The PRTE measures the average marginal treatment effect at each point of new the propensity score distribution.

The baseline counterfactual is an increase of 2pp of the observed subsidy rate, that is, the one we measure with our instrument, not the nominal rate. This is a major gain as it increases it from 8.6% to 10.6% on average. Still, this 2pp increase is realistic, since its budget impact is similar to the what the 2023 reform achieves. As a reminder, subsidy rates vary by year and main crop type.

The full unit support hypothesis. The best case for PRTE requires a full unit support, as stated in CARNEIRO, HECKMAN, and VYTLACIL (2011), which we do not have. The fact is that below some level of the propensity score, we observe virtually no one with insurance, hence the impossibility to evaluate MTE over some range (high resistance to treatment). While full support is not required for the base MTE estimation (a limited range is informative after all), it is a condition for a complete and proper PRTE estimation where the chosen counterfactual may explore propensities outside of the common support.

We have to ignore the effects of the policy for the fraction of the population with high resistance, thus underestimating the impact. We think that the error is small. Indeed in our case only the right tail of the common support is missing (i.e. the most resistant farmers are hardly seen with insurance), 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 take-up to the subsidy rate is extremely low over the full available common support. Yet 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.

	Mean	SD
S_B: Average subsidy per insured farmer (baseline)	720	2202
S_C: Average subsidy per insured farmer (2 pp increase)	1157	3417
P_B: Uptake rate (baseline)	0.28	0.45
P_C: Uptake rate (2 pp increase)	0.31	0.19

Table 12: Parameters of the counterfactual policy

Results. Table 12 shows the parameters and effects of this policy.[28] The average subsidy per farmer increases by 61% over the whole period. The take-up rate increases by 3pp (11%), which illustrates the low elasticity of take-up to insurance subsidies (about 0.5), which is consistent with our previous results. This reinforces the finding that cost is not the main barrier to insurance subscription. The total cost of the subsidy increase can therefore be decomposed in the following way:

$$[N \times (P_C - P_B) \times S_C] + [N \times (S_C - S_B) \times P_B], \tag{40}$$

with N the total number of observations, P_C and P_B respectively the take-up rate for the counterfactual and the baseline, and S_C and S_B the average subsidies in absolute value for the counterfactual and the baseline, all conditioned on the main crop type. The left hand side of the calculation corresponds to the cost of the new entrants, while the right hand side is the pure transfer for those who were already insured. Over the whole population of farmers in France, the new entrants cost €43.2M, while the transfers cost €153.6M, for a total cost of €196.8M.[29] This corresponds to a 38% total subsidy budget increase for 36,000 new contracts.

Figure 6 shows the detailed curve of the counterfactual analysis, while Table 13 shows the estimates of the cost-benefit analysis. The PRTE is significantly lower than the ATE (6% vs. 15%). This means that the newly insured farmers actually reap lower benefits from their subscription than those who were already insured (€6,000 per farmer on average, or €216M in total).[30]. In other words, For 1€spent, about 1 is created, which is not bad form a social point of view. But 0.78€is spent in vain. This fact, combined with the insubstantial increase in take-up for such a large budget increase, shows that a direct increase in insurance subsidies comes mostly with consequences other than a higher take-up rate.

These results highlight the issues with increasing subsidies indiscriminately, which corroborates the intuitions from the GAO (2023a) Report (see Section 2.2 for a highlight of the results): high levels of subsidies may be inefficient due to transfer and composition effects.

[28] The increased take-up rate comes from the Probit prediction. For the increased subsidy per farmer, we make the approximation that, conditioning on main crop type, the newly insured farmers receive the same average subsidies (in absolute value) as the previously insured ones under the counterfactual scenario.

[29] This is scaled up for our sample of 18,000 farmers. To get an estimate of the cost on the whole population, we just multiply by 24 the estimates over the sample: €1.8M for the newly insured, €6.4M for the already insured, and a total €8.2M over the 20 years of the sample.

[30] This is a marginal effect at the average : we apply the treatment effect to the average value of the EBITDA

	Mean
Total budget increase (€M)	196.8
Number of newly insured farmers	36,000
Indirect benefits of the subsidies (€M)	216
Pure transfers to those already insured (€M)	153.6
Pure transfers to the newly insured (€M)	43.2

Table 13: Aggregate results of the counterfactual analysis (Figures scaled up to all farmers in France)

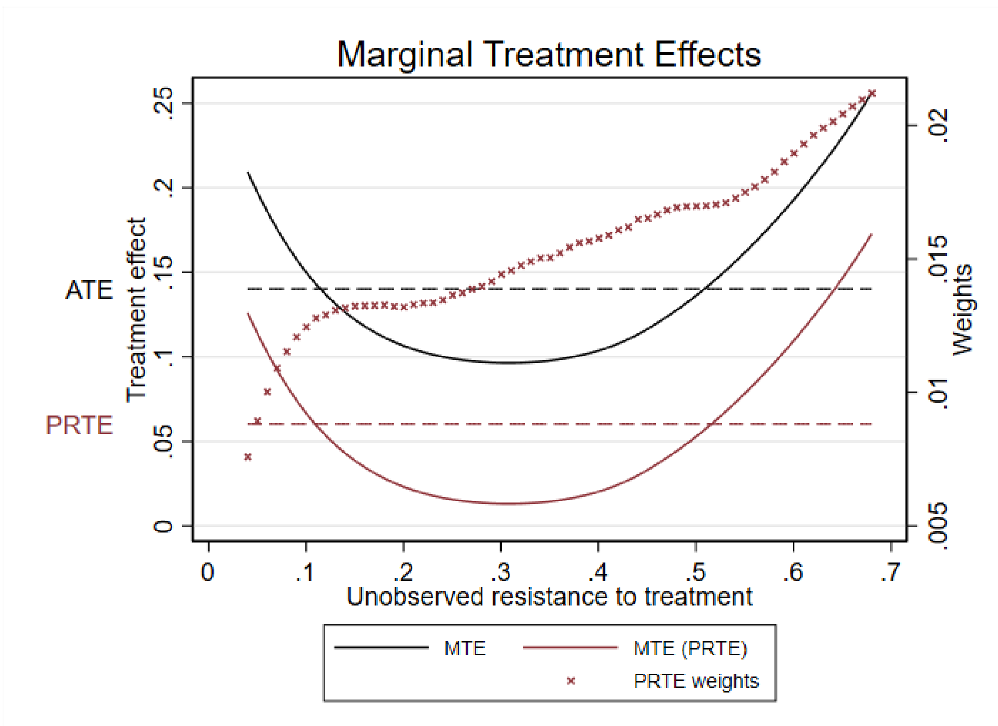


Figure 6: PRTE for a 2pp increase in subsidy rate

Note: The red curve being lower than the black curve means that on average those that insure due to the change in policy derive less benefits than in the baseline. The weights mean that the policy will mostly have an impact on those that are resistant to treatment.

8.2 Increasing the propensity scores

Policy design. We examine now counterfactual policies by which the propensity scores are increased across the population in different ways. We parameterize these policies and we evaluate their consequences. The discussion about how to do all this in reality is left for the Subsection 8.3 ("Policy implications").

The PRTE of CARNEIRO, HECKMAN, and VYTLACIL (2011) defines a method to provide accurate and relevant policy insights in the absence of common support. The marginal policy relevant treatment effects (MPRTE) can be calculated when a counterfactual policy can be continuously parameterized as a perturbation of the baseline policy. The MPRTE is measured as the limit of the average effect when the parameter goes to zero. The small variation is reassuring as it avoids unproven extrapolation. It can be seen as a function derivative à la Gateaux. 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.

It remains to find plausible actions that could have the effect of shifting propensity scores. In our case, 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" policies which are likely to be less costly than an insurance subsidy increase. We follow CARNEIRO, HECKMAN, and VYTLACIL (2011) who provide three types of parameterization of the MPRTE. These three types are mostly equivalent in our case, and their differences in the weight composition lead to little aggregate differences. The 3 MPRTE estimators correspond to 3 different weight distributions. MPRTE1 is an increase with the PRTE weights, MPRTE2 is a fixed upwards shift, MPRTE3 is a proportional upwards shift.

Results. The MPRTE estimates 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 take-up will be far more efficient than policies that target the subsidy rate.

More precisely, the results in Table 14 show that regardless of how the weights are distributed, a marginal increase in the propensity score will have a larger positive impact on the newly insured farmers than increasing subsidies. Therefore, targeting the propensity score directly, rather than through subsidies, appears to be the way to go. The intuition is that farmers who insure due to a direct increase of their propensity score generate much more wealth than farmers who don't, which could stem from the fact that they insure "for the right reasons" once their beliefs have been updated.

To showcase the scale of the results and properly compare them to the PRTE, we perform a simple calculation using our counterfactual policy. Assume that the Government wishes to launch an information campaign with the goal of obtaining the same increase in take-up as our increase 2 pp in subsidies tested in Subsection 8.1, that is a policy that would create 36,000 new contracts.

Effects	(1) EBITDA net of insurance subsidies (log)	(2) SE
MPRTE1	0.159***	0.001
MPRTE2	0.135***	0.002
MPRTE3	0.156***	0.022
PRTE	0.060***	0.001
Observations	100,329	

Note: The 3 MPRTE estimators correspond to 3 different weight distributions: MPRTE1 is an increase with the PRTE weights, MPRTE2 is a fixed upwards shift, MPRTE3 is a proportional upwards shift.
Bootstrapped standard errors, *90% CI, **95% CI, ***99% CI

Table 14: MPRTE estimates (semiparametric)

Assuming a linear marginal rate of return for these contracts, the indirect benefits of this policy for farmers would be between €418M and €492M, [31] double what they were with the increased subsidies. This means that, assuming the same increase in take-up rate, an information campaign could cost double the amount of a subsidy increase policy and still be worth it in terms of wealth creation. Table 13

8.3 Policy implications

The main message of this study is that insurance, though an important tool, faces strong resistance to adoption that classical incentives like subsidies cannot solve, or at a huge cost. On the latter point, we concur with the vigorous claims in the report of the GAO (2023b). Subsidies can lose their initial purpose and just provide unjustified transfers to the biggest farms.

In the absence of specific fields experiment, we can nevertheless conceive 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 area 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 discrimination between firms.[32] 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*

[31] We multiply the MPRTE with the average EBIDTA in our sample, then scale the result to the 36,000 farmers to obtain those figures

[32] Firms are endogenous to the regulatory and fiscal system. Though fiscal optimization may attenuate the efficiency of this solution, the disadvantages of splitting farms (duplication of administrative burden for example) would limite the unintended effect.

in direct correlation with the goal of the established norm".[33] 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, but also supporting traditional farms), and second, the violation needs to be scientifically justified (as with 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 take-up means targeting the propensity score (i.e. resistance to insurance) directly, which translates into information campaigns aimed at those farmers that would benefit the most from insurance (smaller, less diversified). This would not only be cheaper but also far more efficient than increasing insurance subsidies indiscriminately.

  Easier said than done? Actions on the supply side like bonuses given to insurers, to incentivize active underwriting, could be considered. Given that the subsidy is otherwise highly regulated, the risk that bad products are sold is limited.[34] Moreover, most farmers are already in contact with insurers. An alleviation of the paperwork in general and of the farmers in particular, a service for which insurers could be efficient, would be attractive.

  The most adequate form of information could be example based and concrete. Peer effects (imitation) is generally considered as a efficient channel. The French Ministry of Agriculture is already promoting model farms to rein excess use of phytosanitary products (DEPHY network). This optimized protective strategy could well be combined with better financial hedging.

- Third, plain financial hurdles to subscription need to be lowered. Besides informational issues, the timing of the insurance subsidies needs to be reviewed so that farmers don't have to wait months for the support. While money management statistics are not yet available, it is safe to assume that many smaller farms in France live year to year with very little cash available. This measure would cost very little to the State (essentially the interest rates for the growing period) but would drastically increase take-up.

9 Conclusion

This study comes at a time of reform for the French crop insurance market. It was clear from observers and concerned parties that barriers to a general extension of coverage where high, though the remedies are not consensual. One concern was the reliance of the public relief scheme, which offers a kind of free basic insurance. To limit crowding out, the benefits from this scheme are now significantly higher for those with private insurance. The subsidy is thus ex post instead of ex ante,

[33]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."

[34]The role of Freddie Mac and Fanny Mae in the subprime crisis of 2008 may be a scarecrow. However, the nature of the risk (crop losses) doesn't have the systemic nature of the real estate market.

and its form is more striking. It seems that the subscription rate was boosted in the first year of application (2023) but less so afterwards (2024).

While the benefits of crop insurance for farmers cannot be disputed on average, it is clear that insurance is not a cure-all for the climate crisis that the agricultural sector has been facing in the past decades. A goal of 100% coverage is not reachable, and incentives to insure need to be properly targeted to those farmers that need it the most. Going beyond the average effects, we identify that these are mainly small, diversified farms, and as well as farms with a low probability to insure for idiosyncratic (and non observable) reasons.

We have identified mechanisms that are weak and costly (subsidies) and other that are promising yet still fuzzy (information, subscription assistance, nudges). Targeting cognitive and behavioral frictions is likely to have a much larger effect on insurance subscription compared to straight increases of subsidy rates. In that sense the 2022 reform, which simplified the subsidies system, is a welcome addition to the crop insurance landscape. Increasing our understanding of how farmers actually perceive crop insurance, finding ways to improve the way they update their beliefs and designing insurance contracts that can reduce those frictions will be the next challenge facing the sector.

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.

AGRIMER (2022). *Marché du blé dur*. URL: https://www.franceagrimer.fr/content/download/66775/document/Bilan_ble_dur_2020-2021%20web.pdf.

— (2024). *Carotte*. URL: https://rnm.franceagrimer.fr/prix?CAROTTE.

AIHOUNTON, Ghislain B. D. and Arne HENNINGSEN (2021). "Units of measurement and the inverse hyperbolic sine transformation". *Econometrics Journal* 24, pp. 334–351.

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*.

BAREILLE, François and Raja CHAKIR (2023). "The impact of climate change on agriculture: A repeat-Ricardian analysis". *Journal of Environmental Economics and Management* 119, p. 102822.

— (2024). "Structural identification of weather impacts on crop yields: Disentangling agronomic from adaptation effects". *American Journal of Agricultural Economics* 106, pp. 989–1019.

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.

BO AGRI (2020). *Cahier des charges applicable aux entreprises d'assurance pour la prise en charge partielle de primes et cotisations d'assurance récolte 2020*. URL: https://info.agriculture.gouv.fr/gedei/site/bo-agri/document_administratif-eb37c9db-3b80-4abc-a8a4-50402dc6810e.

BOOGAARD, H. et al. (2022). *Agrometeorological indicators from 1979 to present derived from reanalysis, version 1.0*. URL: 10.24381/cds.6c68c9bb.

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

BRINCH, Christian N., Magne MOGSTAD, and Matthew WISWALL (2017). "Beyond LATE with a discrete instrument". *Journal of Political Economy* 125, pp. 985–1039.

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

BURKE, Marshall and Kyle EMERICK (2016). "Adaptation to climate change: Evidence from US agriculture". *American Economic Journal: Economic Policy* 8, pp. 106–40.

CAISSE CENTRALE DE RÉASSURANCE (2019). *Note de synthèse sur le suivi de l'aide à l'assurance multirisque climatique pour l'exercice 2016 par CCR pour le Ministère de l'Agriculture et de l'Alimentation*. Accessed in 2024. URL: https://agriculture.gouv.fr/telecharger/106262.

— (2023). *Catastrophes naturelles*. Accessed in 2023. URL: https://www.ccr.fr/activites/reassurances-et-fonds-publics/catastrophes-naturelles.

CAPITANIO, Fabian, Maria BIELZA DIAZ-CANEJA, Carlo CAFIERO, and Felice ADINOLFI (2011). "Does market competitiveness significantly affect public intervention in agricultural insurance: the case in Italy". *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.

CHAKIR, Raja and Julien HARDELIN (2014). "Crop insurance and pesticides in French agriculture: An empirical analysis of multiple risks management". *Review of Agricultural and Environmental Studies*, pp. 25–50.

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, Cécile DETAND-DESSANDRE, Pierre DUPRAZ, and Hervé GUYOMARD (2021). "Revenus agricoles, aides directes et future PAC : focus sur les exploitations françaises de ruminants et de grandes cultures". *INRAE Productions Animales* 34, pp. 173–190.

COMMUNES DE FRANCE (2020). *Base des codes postaux*. URL: https://www.data.gouv.fr/fr/datasets/communes-de-france-base-des-codes-postaux/.

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.

ENJOLRAS, Geoffroy, Fabian CAPITANIO, and Felice ADINOLFI (2012). "The demand for crop insurance: Combined approaches for France and Italy". *Agricultural Economics Review* 13, pp. 5–22.

ENJOLRAS, Geoffroy and Patrick SENTIS (2011). "Crop insurance policies and purchases in France". *Agricultural Economics* 42, pp. 475–486.

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*, 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.

GAO (2023a). *Crop Insurance: Update on Opportunities to Reduce Program Costs*. URL: https://www.gao.gov/products/gao-24-106086#:~:text=In%5C%20addition%5C%20to%5C%20paying%5C%20companies,total%5C%20cost%5C%20of%5C%20%5C%2417.3%5C%20billion.

— (2023b). *Reducing Crop Insurance Costs Could Fund Other Priorities*. URL: https://www.gao.gov/products/gao-23-106228.

GÉRY, Claire, Vincent HECQUET, and Félix LUCAS (2023). *Le compte provisoire de l'agriculture pour 2022*. URL: https://www.insee.fr/fr/statistiques/7645553#documentation.

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 (2005). “Structural equations, treatment effects, and econometric policy evaluation”. *Econometrica* 73, pp. 669–738.

— (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.

HOROWITZ, John K. and Erik LICHTENBERG (1994). “Crop insurance and agricultural chemical use”. *Economics of Agricultural Crop Insurance: Theory and Evidence*. Springer, pp. 307–322.

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.

IMBENS, Guido W. and Joshua D. ANGRIST (1994). “Identification and estimation of Local Average Treatment Effects”. *Econometrica* 62, pp. 467–475.

INSEE (2020). *Tableaux de l’économie française*. URL: https://www.insee.fr/fr/statistiques/4277860?sommaire=4318291#consulter.

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.

JOURNAL OFFICIEL (2015). *Décret n° 2015-629 du 5 juin 2015 fixant pour l’année 2014 les modalités d’application de l’article L. 361-4 du code rural et de la pêche maritime*. URL: https://www.legifrance.gouv.fr/loda/id/JORFTEXT000030681621/2021-05-04/.

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

KORRES, Nicholas E. et al. (2016). “Cultivars to face climate change effects on crops and weeds: a review”. *Agronomy for Sustainable Development* 36, pp. 1–22.

KOWALSKI, Amanda E. (2023). “Reconciling seemingly contradictory results from the Oregon Health Insurance Experiment and the Massachusetts Health Reform”. *Review of Economics and Statistics* 105, pp. 646–664.

LÉ, Mathias (2022). “L’adaptation des Economies au changement climatique: les enseignements tirés de la recherche Economique”. *Bulletin de la Banque de France*.

LIGNEAU, Laurence et al. (2020). “Observatoire du changement climatique pour l’agriculture: resultats preliminaires en Bretagne”. *Changement Climatique et Territoires*, pp. 427–432.

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

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

MINISTÈRE DE L’AGRICULTURE (2010). *Nomenclature des OTEX*. URL: http://sg-proxy02.maaf.ate.info/IMG/pdf/nomenlature-Otex.pdf.

— (2018). *Aide à l’assurance récolte – Exemple de la France*. URL: https://opera-connaissances.chambres-agriculture.fr/doc_num.php?explnum_id=103577.

MINISTÈRE DE L'AGRICULTURE (2022a). *L'Assurance multirisque climatique des récoltes*. URL: https://agriculture.gouv.fr/telecharger/82960?token=5a25d343f79718a87a4d75d495d59f7c2276197d2e3113798fe931d66584cffe.

— (2022b). *La gestion des risques en agriculture*. URL: https://agriculture.gouv.fr/la-gestion-des-risques-en-agriculture.

— (2022c). *Recensement agricole 2020*. URL: https://agreste.agriculture.gouv.fr/agreste-web/download/publication/publie/Pri2213/Primeur%202022-13_RA2020_%20VersionD%C3%A9finitive.pdf.

— (2022d). *Tout savoir sur le régime des calamités agricoles*. URL: https://agriculture.gouv.fr/tout-savoir-sur-le-regime-des-calamites-agricoles.

RICA (2022). *Accès aux données individuelles du RICA*. URL: https://www.agreste.agriculture.gouv.fr/agreste-web/servicon/I.2/listeTypeServicon/.

ROBINSON, Peter M. (1988). "Root-N-consistent semiparametric regression". *Econometrica: Journal of the Econometric Society*, pp. 931–954.

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.

ROZAN, Anne and Sandrine SPAETER (2024). "Quel équilibre entre protection ex ante et compensations ex post dans la réforme de l'assurance agro-climatique?" *Revue Économique*.

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.

SÉNAT (2003). *Projet de loi de finances pour 2005 : Agriculture, alimentation, pêche et affaires rurales*. URL: https://www.senat.fr/rap/l04-074-303/l04-074-3034.html#toc130.

SMITH, Vincent H. and Barry K. GOODWIN (1996). "Crop insurance, moral hazard, and agricultural chemical use". *American Journal of Agricultural Economics* 78, pp. 428–438.

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

USDA (2022). *Reinsurance Agreements*. URL: https://www.rma.usda.gov/en/Topics/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.

VERARDI, Vincenzo and Nicolas DEBARSY (2012). "Robinson's square root of N consistent semiparametric regression estimator in Stata". *The Stata Journal* 12, pp. 726–735.

VIE PUBLIQUE (2022a). *Agriculture française : une puissance mondiale qui décline*. URL: https://www.vie-publique.fr/en-bref/286593-agriculture-francaise-une-puissance-mondiale-qui-decline.

— (2022b). *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*. URL: https://www.vie-publique.fr/loi/282699-loi-2-mars-2022-reforme-du-regime-de-lassurance-recolte-en-agriculture.

VYTLACIL, Edward J. (2002). “Independence, monotonicity, and latent index models: An equivalence result”. *Econometrica* 70, pp. 331–341.

WALTHALL, Charles L. et al. (2013). *Climate change and agriculture in the United States: Effects and adaptation*. USDA Technical Bulletin 1935.

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, JunJie (1999). “Crop insurance, acreage decisions, and nonpoint-source pollution”. *American Journal of Agricultural Economics* 81, pp. 305–320.

YU, Jisang, Aaron SMITH, and Daniel A. SUMNER (2018). “Effects of crop insurance premium subsidies on crop acreage”. *American Journal of Agricultural Economics* 100, pp. 91–114.

A Weather station density across Europe for the Copernicus datasets

Figure 7 shows the distribution of weather stations across Europe, as mentioned in Section 6.

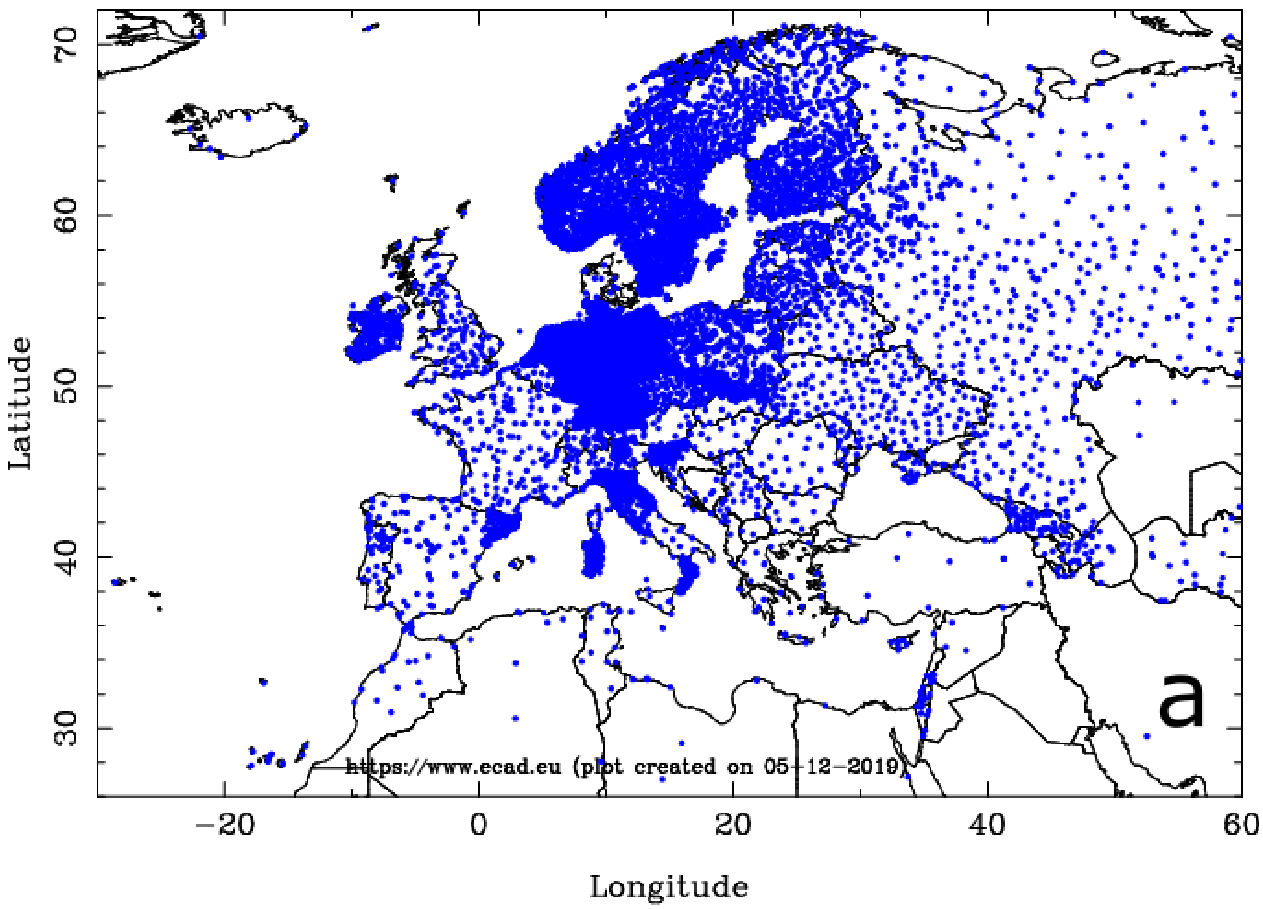


Figure 7: Weather station density across Europe for the Copernicus dataset. Source: Copernicus

B Details on variable construction and data

B.1 GDDs

Below we compute the GDD index over the year using the classic methodology. Although this specific index is not used in the paper, our measure of extreme temperatures derives from it:

$$\mathrm{GDD}_{stc} = \frac{\sum_d \mathrm{GDD}_{sd}}{\mathrm{GDD}_c^{\mathrm{opt}}} \text{ with } \begin{cases} \mathrm{GDD}_{sdc} = 0 & \text{if } T_{sdc} \leq T_c^b \\ \mathrm{GDD}_{sdc} = T_{sdc} - T_{sdc}^b & \text{if } T_c^b < T_{sdc} < T_c^u \\ \mathrm{GDD}_{sdc} = T_c^u & \text{if } T_{sdc} \geq T_c^u \end{cases} \tag{41}$$

with s the 6 km$\times$6 km square (or more generally the geographic unit where the individual is located), d the day of year t and c the crop type. The rationale behind this formula is that when temperatures are too cold, plants do not absorb any energy (first line). When the temperatures are in the right range, they absorb energy linearly based on the temperature (second line). Finally, when temperatures exceed the threshold, they keep absorbing energy but at a less efficient rate (third line). $\mathrm{GDD}_c^{\mathrm{opt}}$ is a normalization.

B.2 Subsidies

Variations of the subsidy rates. Figure 8 draws the subsidy rates by aggregated categories for ease of reading. The graph is flat before 2009 because the subsidies, as they were defined until 2022, were introduced in the 2009 reform. In the RICA database, subsidy data is only available since 2009, and there is no data on pre-2009 subsidies. The fall in 2020 comes from a reduction in overall subsidies spending due to an increased take-up of non-subsidized contracts. This still needs to be explored.

These variations come from two sources of variation, respectively due to the annual modification of the insured value tranches for each crop or a possible budget revision (Section 8). Below are details and examples of the rules governing annual changes to the insured value of crops.

Annual modification of the crop insured value tranches. The subsidized insured value by crops changes every year and is generally lower than the actual price of the crop. In practice, the subsidies most often only cover parts of the insurance contract. We explain here the details of subsidies rules.

Assume a farmer in 2020 who wants to insure their carrot production: they subscribe to the base contract (subsidized at 65%), but the base to be eligible to the subsidy means they can only insure up to 54-264 €/t (BO Agri, 2020), or 0.054-0.264 €/Kg. Considering that the selling price of carrots is 2020 was between 0.50-1 €/kg (depending on where they are sold) (AgriMer, 2024), this means that with this contract the farmer would only insure 11-53% of their production value. Carrots are not necessarily representative of all crops, for example the base value for wheat is

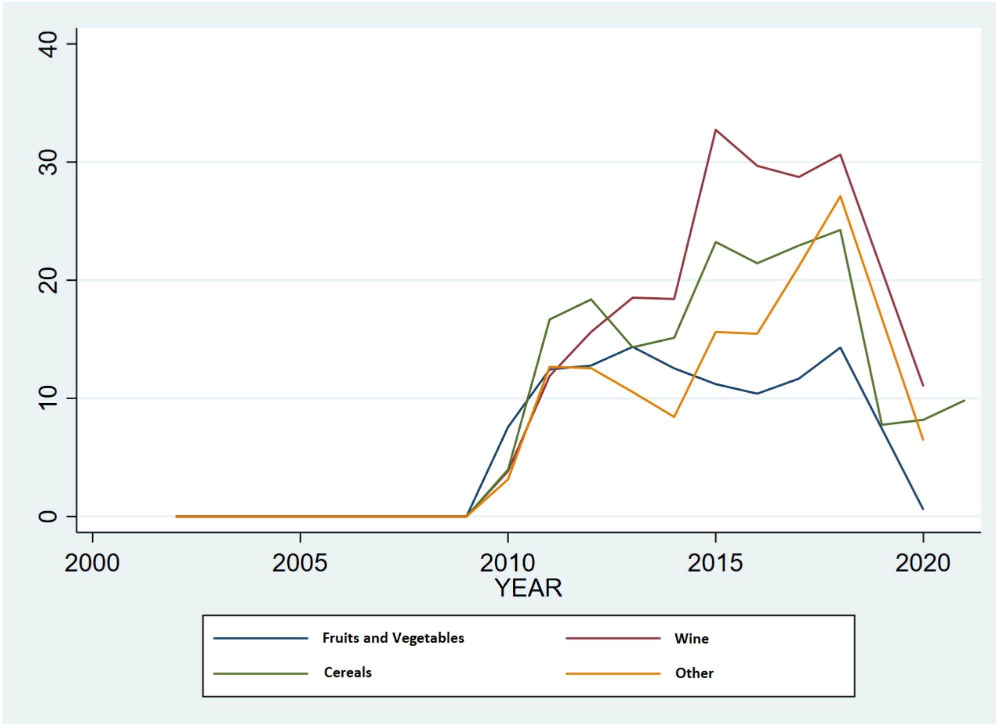


Figure 8: Subsidy rate computed from the RICA data by aggregated categories

252 €/t, closer to the average selling price of 283 in 2022. Table 15 shows different examples of the comparison between the insured price and the market price.

Should farmers want to increase that value, they can subscribe the complementary contract (subsidized at 45%) up to 0.064 €/kg, and an extra unsubsidized protection should they want to cover more.

The computation of "Actual subsidy rate for the market price" is simply a multiplication of the share of the market price subsidized by 65%, as we assume that insurance premiums increase linearly with the insured value.[35] These rates, just like our instruments, are lower than the nominal rate due to the low insured bases.

Crop type	**Subsidies base price €/t for 65%**	**Market price €/t**	**Share of market price subsidized**	**Actual subsidy rate for market price**	**Sources for market price**
Carrots	54-264	500-1000	11-53%	7-24%	AGRIMER (2024)
Winter wheat	252	283	89%	58%	AGRIMER (2022)
Tomatoes	618	1400-2500	25-44%	16-29%	AGRIMER (2024)

For carrots and tomatoes, ranges are given due to the heterogeneous nature of the price. For winter wheat, we use the actual commodity price recorded in France.

Table 15: Examples of comparisons between nominal and observed subsidy rates (2020)

[35]Indeed, as this table is for illustration purposes only, we do not include the portion of the price subsidized at 45%.

B.3 Insurance pricing

We include in our Probit regression (the results of which can be found in Table 6) a measure of price. This measure is computed at the crop and year level following the same logic as for our instrument. We sum the premiums net of subsidies paid by every farmer in a given year and crop and divide by the number of insured farmers assigned to that crop. We then apply this figure to all the farmers, insured or not, corresponding to the year and crop. This effectively gives us an average measure of price. Formally

$$\mathbb{E}(q|t,c) = \frac{\sum_i q_{itc}}{n_{tc}}, \tag{42}$$

with q the premiums paid and n the number of farmers i in year t and for a given crop type c.

B.4 Additional summary statistics

Table 16 shows the global loss ratio (i.e. the ratio of the sum of premiums over payouts) for our entire sample, with and without insurance subsidies. Including insurance subsidies makes insurance worth it for farmers, since they receive about the same amount as they put in.

Table 17 shows the insurance uptake rate by aggregated crop type over the full sample. Vine and cereal growers appear to be the most willing to take up insurance, as confirmed by our Probit regression in Table 6.

	Loss ratio
Net (without insurance subsidies)	91%
Gross (with insurance subsidies)	101%

Table 16: Loss ratio aggregation (total claims/total premiums)

Crop type	Insurance subscription rate
Cereals	33%
Fruits and vegetables	28%
Vine	47%
Other/Mixed	8.5%

Table 17: Subscription rate by aggregated crop category

C OLS estimates of Equation (17)

Table 18 shows the result of the OLS specification of Equation 17 (i.e. the impact of crop insurance uptake on revenue mean and variance without an instrument). The coefficient have the same sign and are highly significant, but are orders of magnitude below those we find in the IV regression (5)

	EBITDA with insur. subsidies		EBITDA w/out insur. 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 OOBs (log)	0.003***	0.000	0.003***	0.000
	(0.001)	(0.000)	(0.001)	(0.000)
L.Cold OOBs (log)	-0.005***	-0.001***	-0.005***	-0.001***
	(0.001)	(0.000)	(0.001)	(0.000)
L2.Cold OOBs (log)	-0.000	-0.001**	-0.000	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
L3.Cold OOBs (log)	-0.003***	0.001***	-0.003***	0.001***
	(0.001)	(0.000)	(0.001)	(0.000)
Hot OOBs (log)	-0.014***	-0.001	-0.015***	-0.001
	(0.003)	(0.001)	(0.003)	(0.001)
L.Hot OOBs (log)	0.003	0.000	0.003	0.000
	(0.003)	(0.001)	(0.003)	(0.001)
L2.Hot OOBs (log)	0.011***	-0.001	0.011***	-0.001
	(0.002)	(0.001)	(0.002)	(0.001)
L3.Hot OOBs (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	71,524	71,524	70,750	70,750
ρ	1	0	1	0
Farmer FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Table 18: OLS log estimations for the impact of insurance on the revenue distribution

D Separation between indemnities and behavioral insurance effects

Table 19 shows the result of the IV specification of Equation 19 with the added interaction term (Insurance $\times$ Indemnities) in an attempt to disentangle the behavioral and financial effects of insurance on revenue. Farmers who take up insurance and earn indemnities appear to have lower revenues than farmers who don't, which confirms the intuition that farmers who tend to protect their crop more (i.e. shielding) gain the most out of insurance. Furthermore, these coefficients show that indemnities do not fully cover farmers' losses.

	EBITDA with insur. subsidies		EBITDA w/out insurance subsidies	
	Mean	Variance	Mean	Variance
Dummy for crop insurance status (1=insured)	0.377***	-0.006	0.329***	0.003
	(0.060)	(0.019)	(0.054)	(0.016)
Insurance× Indemnities	-0.023***	-0.001	-0.020***	-0.001
	(0.004)	(0.001)	(0.003)	(0.001)
Observations	51,142	51,142	50,567	50,567
Weak Ident.	71.041	71.041	76.827	76.827
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

*: 90% significance, **: 95% significance, ***: 99% significance. Robust standard errors in parenthesis.

Table 19: 2nd stage IV log estimations for the impact of insurance on the revenue distribution

E Recovered effects from the MTE analysis and common support

E.1 Parametric

Table 20 show the recovered effects from the parametric MTE analysis. These are very similar, albeit higher, than our semi-parametric estimation, which confirms its robustness.

Effects	(1) EBITDA net of insurance subsidies (log)	(2) Variance
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	100,834	70,565

Table 20: Recovered estimators from the MTE framework

E.2 Parametric MTE and common support

Figures 9 and 10 show the MTE curves of our parametric analysis, which have a similar shape to our semiparametric curves used in the core of the paper.

Figure 11 shows the common support used for all MTE analysis (does not differ between semi-parametric and parametric), with the red dash lines showing where the support has been cut due to a lack of data for the matching.

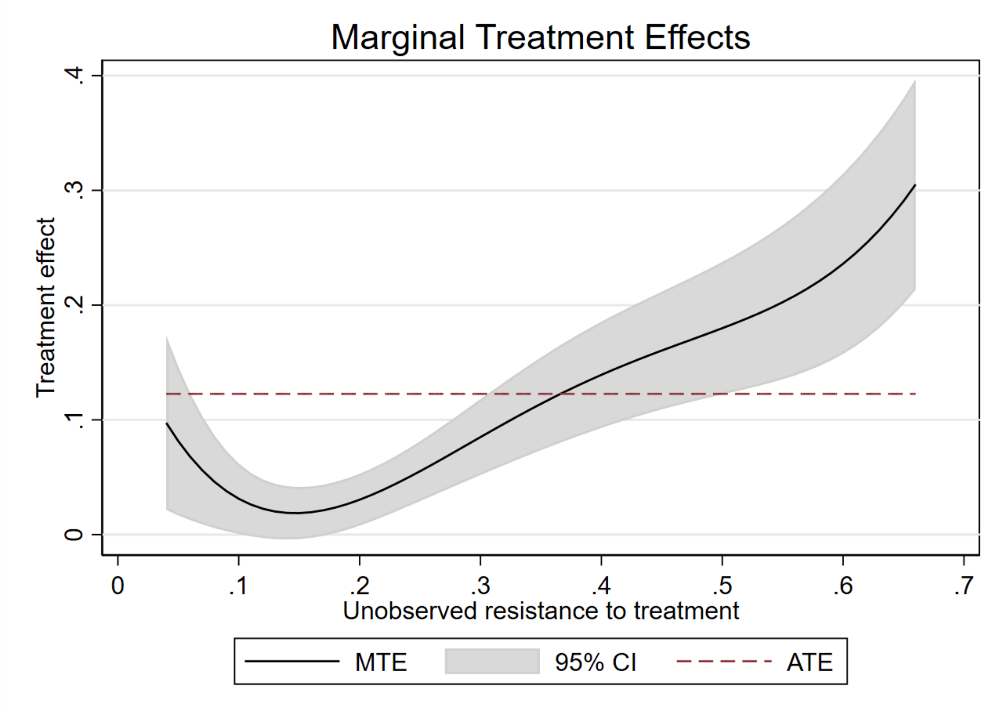


Figure 9: MTE curve for mean of EBITDA net of insurance subsidies (4th degree estimation of $K(p)$)

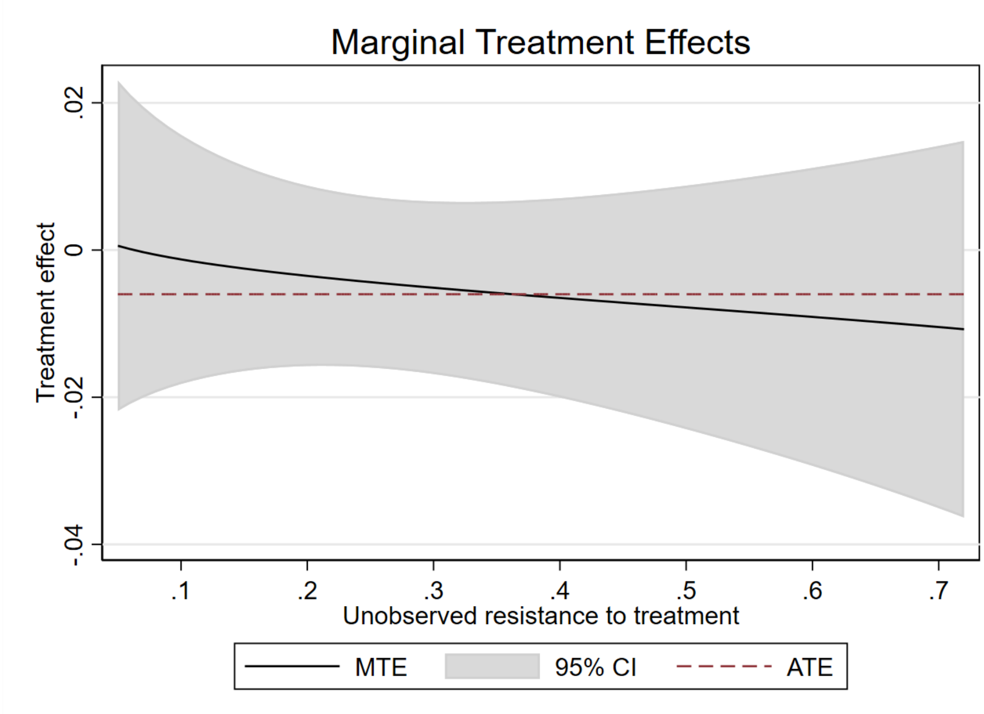


Figure 10: MTE curve for variance of EBITDA net of insurance subsidies (4th degree estimation of $K(p)$)

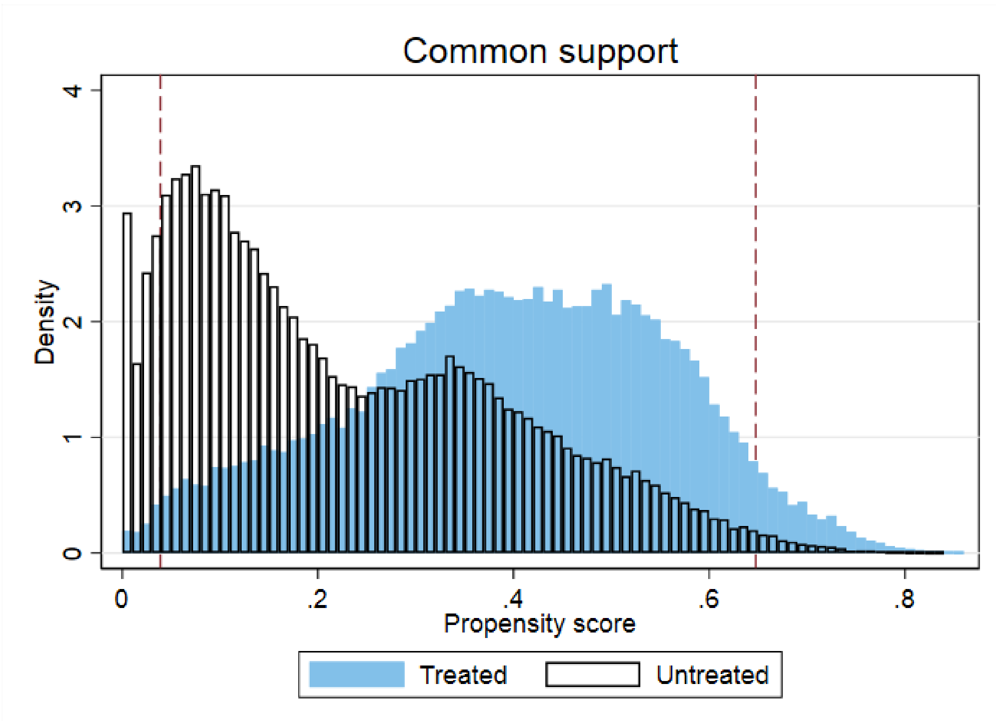


Figure 11: Common support for all MTE graphs

F Robustness checks

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

F.1 Alternative regression for Equation (17) with continuous effects

Table 21 shows the results of Equation (19) using a continuous measure for insurance uptake. Rather than using a dummy like in the main regression, we use a measure of insurance spending, consistent with WANG, REJESUS, and AGLASAN (2021). We find similar orders of magnitude and the same signs and robustness as in the main regression.

F.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 EBITDA), 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 (EBITDA and insurance spending) using the same IV specification as in Section 5.2. The coefficients retain the same signs with some changes in scale (noticeably higher), and are still statistically significant. Table 22 shows the result of this exercise. The sign and significance remains, but the order of magnitude is much higher, which can be explained by the IHS sensitivity to units as showcased in (AIHOUNTON and HENNINGSEN, 2021). One should therefore be careful in interpreting those figures as elasticities, although the sign and standard errors are still a good indication that the log transformation works.

F.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, namely the 2005 and 2016 reforms. As discussed with 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 (MINISTÈRE DE L'AGRICULTURE, 2022a). 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 23.

	EBITDA with insur. subsidies		EBITDA w/out insur. 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 OOBs (log)	0.002*	0.000	0.002**	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Hot OOBs (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	69,790	69,790	69,006	69,006
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 21: IV estimations for the impact of insurance on the revenue distribution

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 subsample 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.

	With insurance subsidies		w/out 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 OOBs (log)	0.192***	-0.948*	0.093***	-0.094
	(0.044)	(0.517)	(0.016)	(0.070)
L.Cold OOBs (log)	0.147***	-1.456***	0.045***	-0.226***
	(0.042)	(0.507)	(0.016)	(0.069)
L2.Cold OOBs (log)	0.238***	-2.693***	0.089***	-0.344***
	(0.044)	(0.535)	(0.017)	(0.075)
L3.Cold OOBs (log)	-0.179***	1.775***	-0.085***	0.227***
	(0.045)	(0.539)	(0.017)	(0.072)
Hot OOBs (log)	0.034	-1.279	-0.072	-0.109
	(0.129)	(1.753)	(0.049)	(0.244)
L.Hot OOBs (log)	0.086	-2.637*	0.012	-0.357*
	(0.134)	(1.572)	(0.051)	(0.209)
L2.Hot OOBs (log)	0.174*	0.105	0.080**	0.046
	(0.102)	(1.314)	(0.039)	(0.176)
L3.Hot OOBs (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	69,790	69,790	69,006	69,006
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 22: IHS results for Equation (17)

	EBITDA with insur. subsidies		EBITDA w/out 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 OOBs (log)	0.011***	0.000*	0.011***	0.000
	(0.001)	(0.000)	(0.001)	(0.000)
L.Cold OOBs (log)	0.006***	0.000	0.007***	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)
L2.Cold OOBs (log)	0.018***	-0.001**	0.017***	-0.000*
	(0.001)	(0.000)	(0.001)	(0.000)
L3.Cold OOBs (log)	-0.001	-0.000	-0.000	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Hot OOBs (log)	-0.030***	-0.002	-0.030***	-0.001
	(0.004)	(0.001)	(0.004)	(0.001)
L.Hot OOBs (log)	-0.018***	-0.000	-0.017***	0.000
	(0.004)	(0.001)	(0.004)	(0.001)
L2.Hot OOBs (log)	0.014***	0.001	0.014***	0.001
	(0.004)	(0.001)	(0.004)	(0.001)
L3.Hot OOBs (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	69,862	69,862	69,078	69,078
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 23: Alternative IV framework

F.4 Nominal subsidy rate

Additionally to the alternative instrument, we also provide results for a subsample tests using only the official documents (CAISSE CENTRALE DE RÉASSURANCE, 2019) from 2015 onwards, rather than subsidy rates estimated from the data as explained in Section 5.2. These documents provide, on a macro-level, the share of premiums eligible to subsidies in total premiums paid for contracts that are at least partially eligible. We take this share and multiply it by 0.65 to get the subsidy rate. The instrument now becomes the base insured price covered by subsidized insurance, which provides another measure of the subsidy coverage.

The results keep the same sign at both stages, but due to the reduced sample, they lose a lot of their significance. Nonetheless, the coefficients from the first stage confirm that our subsidy rates' heterogeneity do not differ too much from the real ones on the 2015-2020 period.

A comparison between the subsidy rates we compute from the data and the nominal subsidy rates in the 2015-2020 period can be found in Table 24. The subsidy rates we find are a lot lower, because our measure includes in the denominators all multirisk crop insurance contracts, even those that are not eligible to the subsidies, whereas the official measure only includes contracts that are at least partially eligible. To showcase this difference, we also include in the Table a measure of the subsidy rate (not used in the regressions) excluding the non-eligible contracts (0s), making the subsamples much more comparable. Still, both measures exhibit enough variation to be robust instruments. [36]

	Measured subsidy rate	Measured subsidy rate excluding non-eligible contracts	Nominal subsidy rate
Cereals	13.4%	39.8%	45.6%
Fruits and vegetables	8.9%	40.3%	47.1%
Vine	24.6%	50.1%	47.4%
Other/Mixed	17.8%	33.9%	47.6%

Table 24: Comparison of nominal and measured subsidy rates (by aggregated OTEX) between 2015-2020

Table 25 shows the result of the first- and second-stage regressions.

[36] The small variation in the last column comes from the fact that we present here an average over the whole subsample.

	EBITDA with insur. subsidies		EBITDA w/out insur. subsidies	
	(1)	(2)	(3)	(4)
Dummy for crop insurance status (1=insured)	0.073	0.043	0.078	0.049
	(0.145)	(0.042)	(0.134)	(0.037)
Cold GDDs (log)	0.006^{***}	0.001	0.006^{***}	0.000
	(0.002)	(0.001)	(0.002)	(0.000)
L.Cold GDDs (log)	0.004	-0.002^{*}	0.004	-0.002^{**}
	(0.003)	(0.001)	(0.003)	(0.001)
L2.Cold GDDs (log)	0.011^{***}	-0.001	0.011^{***}	-0.001
	(0.002)	(0.001)	(0.002)	(0.001)
L3.Cold GDDs (log)	-0.003^{*}	0.000	-0.003^{*}	-0.000
	(0.002)	(0.001)	(0.002)	(0.000)
Hot GDDs (log)	-0.001	-0.000	-0.002	0.001
	(0.005)	(0.002)	(0.005)	(0.001)
L.Hot GDDs (log)	-0.012^{***}	-0.001	-0.013^{***}	-0.001
	(0.005)	(0.001)	(0.005)	(0.001)
L2.Hot GDDs (log)	0.015^{*}	-0.000	0.015^{*}	-0.000
	(0.009)	(0.003)	(0.009)	(0.003)
L3.Hot GDDs (log)	0.017	-0.003	0.017	-0.004
	(0.012)	(0.004)	(0.012)	(0.003)
Number of floods (log)	-0.015^{***}	0.001	-0.016^{***}	0.001
	(0.005)	(0.002)	(0.005)	(0.002)
L.Number of floods (log)	-0.004	0.000	-0.005	0.002
	(0.004)	(0.002)	(0.004)	(0.001)
L2.Number of floods (log)	0.007^{*}	-0.001	0.005	0.000
	(0.004)	(0.001)	(0.004)	(0.001)
L3.Number of floods (log)	-0.010^{**}	-0.001	-0.011^{***}	0.000
	(0.004)	(0.002)	(0.004)	(0.001)
Number of droughts (log)	0.003	-0.004^{*}	0.000	-0.002^{**}
	(0.004)	(0.002)	(0.004)	(0.001)
L.Number of droughts (log)	0.004	0.000	0.005	-0.001
	(0.004)	(0.002)	(0.004)	(0.001)
L2.Number of droughts (log)	-0.003	0.001	-0.002	0.000
	(0.005)	(0.002)	(0.005)	(0.001)
L3.Number of droughts (log)	0.003	0.001	0.001	-0.001
	(0.006)	(0.002)	(0.005)	(0.001)
Observed subsidy rate (first-stage)	0.164^{**}	0.164^{**}	0.180^{**}	0.180^{**}
	(0.083)	(0.083)	(0.084)	(0.084)
Observations	17,860	17,860	17,621	17,621
Weak Ident.	3.915	3.915	4.563	4.563
Hansen J	0.000	0.000	0.000	0.000
AR F-Test				
Farmer FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Instrument (Nominal rate)	Yes	Yes	Yes	Yes

Table 25: IV with a nominal subsidy rate on a subsample

F.5 LATE with substitution effects

One worry with our LATE estimate may be that, due to the simultaneous nature of the insurance decision with other protection behaviors (e.g. pesticide use), inputs use may be viewed as "bad controls", that is a channel rather than a confounder (WOOLDRIDGE, 2010). Indeed, if, for example, pesticide usage and insurance status are decided at the same time, and both behaviors are substitutes, then the insurance decision affects revenue both through its intrinsic value, and through the reduction in pesticide use. In that case, controlling for pesticides may incur a form of simultaneity bias. On the other hand, excluding pesticides will lead to an omitted variable bias.

Nonetheless, as is standard, we perform our LATE regression from Equation (19) without controlling for the main inputs (pesticide usage) and protection behaviors (greenhouses, cattle). We find that the coefficient do not change, which is expected if the equations are well-specified. This means that the effect we capture with the insurance coefficient is a global effect, including any substitution or shielding that may occur. Considering our model takes these behaviors into account, this is actually a desirable outcome. Table 26 shows the results of this no-controls regression.

	EBITDA with insur. subsidies		EBITDA w/out insurance subsidies	
	Mean	Variance	Mean	Variance
Dummy for crop insurance status (1=insured)	0.229***	-0.008	0.204***	-0.002
	(0.029)	(0.010)	(0.027)	(0.008)
Cold OOBs (log)	0.006***	0.000	0.006***	0.000
	(0.001)	(0.000)	(0.001)	(0.000)
L.Cold OOBs (log)	-0.003**	-0.001**	-0.003***	-0.001**
	(0.001)	(0.000)	(0.001)	(0.000)
L2.Cold OOBs (log)	0.003**	-0.002**	0.002*	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
L3.Cold OOBs (log)	-0.008***	0.001	-0.007***	0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Hot OOBs (log)	-0.024***	-0.001	-0.025***	0.001
	(0.005)	(0.002)	(0.005)	(0.002)
L.Hot OOBs (log)	-0.010**	0.001	-0.009*	0.001
	(0.005)	(0.002)	(0.005)	(0.002)
L2.Hot OOBs (log)	0.011***	-0.002*	0.012***	-0.002
	(0.004)	(0.001)	(0.004)	(0.001)
L3.Hot OOBs (log)	0.007	-0.004***	0.004	-0.003**
	(0.005)	(0.001)	(0.005)	(0.001)
Number of floods (log)	-0.011***	-0.000	-0.011***	-0.000
	(0.003)	(0.001)	(0.003)	(0.001)
L.Number of floods (log)	-0.011***	0.001	-0.010***	0.002
	(0.003)	(0.001)	(0.003)	(0.001)
L2.Number of floods (log)	-0.002	0.001	-0.002	0.001
	(0.003)	(0.001)	(0.003)	(0.001)
L3.Number of floods (log)	-0.008***	-0.000	-0.008***	-0.000
	(0.003)	(0.001)	(0.003)	(0.001)
Number of droughts (log)	0.004	-0.002	0.003	-0.002
	(0.003)	(0.002)	(0.003)	(0.001)
L.Number of droughts (log)	0.007**	-0.001	0.008***	-0.001
	(0.003)	(0.001)	(0.003)	(0.001)
L2.Number of droughts (log)	0.002	0.001	0.002	0.001
	(0.003)	(0.001)	(0.003)	(0.001)
L3.Number of droughts (log)	0.007**	0.000	0.007**	0.000
	(0.003)	(0.001)	(0.003)	(0.001)
Subsidy rate (1st stage)	0.004***		0.004***	
	(0.000)		(0.000)	
Observations	51,322	51,322	50,747	50,747
Weak Ident.	156.170	156.170	162.753	162.753
Hansen J	0.000	0.000	0.000	0.000
AR F-Test				
Farmer FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	No	No	No	No
Instrument	Yes	Yes	Yes	Yes

*: 90% significance, **: 95% significance, ***: 99% significance. Robust standard errors in parenthesis.

Table 26: 2nd stage IV log estimations for the impact of insurance on the revenue distribution without controls

G Channels

We showcase the behavioral impacts of insurance through an analysis of the channels through which insurance take-up may increase revenues. The intuition is that, when farmers insure, they make other endogenous decisions simultaneously that will change their revenues. As discussed in the theory section, this could be moral hazard or, on the contrary, shielding. To test this, we select four potential channels: pesticide use per ha, surface area, fertilizer use per ha and specialization.

We perform an instrumental variable regression from Equation (19) using our insurance subsidy rate as an instrument, but this time replacing insurance take-up in the second stage by our channel of choice. The regression is therefore performed four separate times. In each regression, we do not control for the other channels, as these could interact with one another, and, as discussed in Section 5.2, these may be bad controls.

This amounts to assuming that, if a farmer changes their (e.g.) pesticide usage following an increase in the insurance subsidy rate, we can attribute this change to the insurance decision. Following this, the second stage estimate therefore gives us the impact of changing pesticide usage, following insurance take-up, on revenues. In other words, this is a LATE on farmers who react to the instrument. The results of this exercise can be found in Table 27. All the coefficients are significant and have a positive sign: this means that when farmers insure, they change their behaviors in a way that increases their revenues.

Because the second-stage coefficients showcase the impact of increasing the channel by 1% at the mean following the insurance take-up, some extra computations need to be performed for a proper interpretation. We need to assess the impact of insurance take-up on the channel by using the coefficients from the first stage, and use these results to scale the coefficients from the second-stage.

Formally, let $\mathbb{E}_0$ denote the mean of the channel over the full sample of N farmers. Let θ_1 denote the first-stage coefficient, θ_2 the 2nd stage coefficient, and β_{11} the first-stage coefficient for the LATE of subsidies on insurance subscription (Equation 19). We know that an increase of 1 percentage point in the subsidy rate increases the average channel by $\theta_1\%$ over the whole sample, so that the new mean $\mathbb{E}_1$ becomes $(1 + \theta_1/100)\mathbb{E}_0)$. We also know that the increase is actually concentrated over those who switched into an insurance contract following the increased subsidies. From the first stage, we know the number of these farmers is $n = \beta_{11} \cdot N$. The new mean $\mathbb{E}_L$ (L because this is a LATE) of the farmers who actually changed their practices is therefore:

$$\mathbb{E}_1 = \frac{(N-n)\mathbb{E}_0 + n\mathbb{E}_L}{N}. \tag{43}$$

Rearranging to isolate $\mathbb{E}_L$:

$$\mathbb{E}_L = \frac{N\mathbb{E}_1 - (N-n)\mathbb{E}_0}{n}. \tag{44}$$

From there, we just compute the variation of $\mathbb{E}_L$ compared to $\mathbb{E}_0$ and multiply by θ_2 to get the treatment effect (TE) on revenue for the switchers, and we multiply by 100 to have a value expressed

2nd stage: θ_2 Effect of channel	Crop protection	Surface	Fertilizers	Specialization
Expenditures for crop protection product per ha (log)	0.48068^{**} (0.23771)			
Total surface of the farm (log)		2.32944^{***} (0.81321)		
Expenditures for fertilizers per ha (log)			0.13157^{***} (0.02871)	
Specialization index (log)				0.40308^{***} (0.05437)
1st stage: θ_1 Effect of subsidy rate (year, crop)	0.00228^{**} (0.00111)	0.00049^{***} (0.00017)	0.00835^{***} (0.00164)	0.00275^{***} (0.00026)
Observations	51,660	51,660	51,660	51,660
Weak Ident.	4.237	8.445	25.976	110.137
Hansen J	0.000	0.000	0.000	0.000
AR F-Test	103.889	111.991	103.191	106.738
Farmer FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Instrument	Yes	Yes	Yes	Yes

*: 90% significance, **: 95% significance, ***: 99% significance. Robust standard errors in parenthesis.

Table 27: Channels IV log estimations on revenue

	n	N	θ_1	θ_2	β_{11}	$\mathbb{E}_0$	$\mathbb{E}_1$	$\mathbb{E}_L$	$100\,\frac{\mathbb{E}_L-\mathbb{E}_0}{\mathbb{E}_0}$	**TE-PP**
Crop protection	205	51,142	0.00228	0.48068	0.004	100.77727	100.77957	101.35170	0.57	**0.27**
Surface	205	51,142	0.00049	2.32944	0.004	104.21	104.21051	104.33765	0.1225	**0.29**
Fertilizer	205	51,142	0.00835	0.13157	0.004	118.14605	118.15591	120.61235	2.0875	**0.27**
Specialization	205	51,142	0.00275	0.40308	0.004	0.48	0.48001	0.4833	0.6875	**0.28**

Table 28: Parameters for the channels computation

in percentage points (PP):

$$\text{TE-PP} = 100\,\frac{\mathbb{E}_L - \mathbb{E}_0}{\mathbb{E}_0}\,\theta_2. \tag{45}$$

The results of this calculation can be found in Table 28. All four of our channels have similar impacts (around 0.27), meaning that they explain each about 2% of the global insurance effect, or at most 8% when combined. Because there may be some interactions between these practices, the combined figure may be lower, assuming everything is positively correlated. This may seem low, but those channels are only those we can observe, and it is likely that the main behavior changes go through other, unobserved channels, such as crop management, land use changes, etc.