Index-Based Yield Protection for Smallholder Farmers

Kehan Lu Jing-Sheng Song Can Zhang
Fuqua School of Business, Duke University, {kehan.lu, jssong, c.zhang}@duke.edu

Government subsidies are common in the agricultural sector to protect farmers from unexpected losses. Two major forms of agricultural subsidies are price protection, under which farmers are subsidized when the market price is low, and yield protection, under which farmers are subsidized when the crop yield is low. While price protection is popular in both developed and emerging economies, implementing yield protection in emerging economies is challenging due to the high costs of yield assessment for small farms. This research examines the design of a recently emerged index-based yield protection policy, which triggers subsidies when a pre-determined index, such as rainfall, predicts a low yield, thereby avoiding costly yield assessment. Our analysis generates several intriguing findings. First, we show that while an increase in index-based subsidy can increase farmers' expected income, it can also increase their income variance due to imperfect yield prediction of the index. Based on this result, we uncover a non-monotonic relationship between the optimal subsidy amount and the accuracy of the index. Second, although price and yield protection are often viewed as strategic substitutes since both can incentivize more planting, we show that they act as strategic complements when the index accuracy is low. Finally, when the government can exert a costly effort to improve index accuracy, contrary to expectations, we find that a tighter budget can lead to a higher optimal investment in index accuracy. Collectively, these insights contribute to a more nuanced understanding of index-based yield protection policies, aiding in developing effective agricultural subsidies.

Key words: smallholder agriculture, yield uncertainty, risk aversion, index-based yield protection
History: January 24, 2024

1. Introduction

Smallholder farming is recognized as the backbone of the agricultural industry, particularly in emerging economies (Herger 2020). According to the United Nations' Food and Agriculture Organization (FAO), smallholder farmers who farm on less than two hectares of land account for approximately 84% of the world's 570 million farmers and produce more than a third of the world's food (Lowder et al. 2021). Despite their sizable population and critical role in the world's food supply, smallholder farmers are disproportionately exposed to risks like price fluctuations and severe weather. For instance, the recent cocoa price depression due to oversupply has posed significant livelihood challenges for smallholder cocoa farmers in West Africa (Taylor 2021). Similarly, in Central America's dry Pacific region, a drastic drought from 2014 to 2016 left 1.6 million farmers grappling with hunger and reliant on humanitarian aid due to low crop yields (Harvey et al. 2018). Further, global food security can also be threatened by these risks if they push farmers to abandon farming. Given these challenges, external support is essential for ensuring farmer welfare and food security and facilitating the sustainable development of smallholder agriculture.

Government subsidies play an instrumental role in protecting farmers against unforeseen losses arising from low prices or low crop yields. Two major types of agricultural subsidies are price protection, which subsidizes farmers when the market price of the crop is low, and yield protection, which subsidizes farmers when the crop yield is low. Both types of policies are widely implemented in many developed countries, such as the United States. However, in emerging economies, while price protection policies are commonly adopted, yield protection policies are much less prevalent (Raithatha and Priebe 2020). This is mainly because subsidizing farmers based on their crop yields requires yield assessment of individual farmers, which can be prohibitively costly in emerging economies given the large number of smallholder farmers (Greatrex et al. 2015, Carter et al. 2017).

To overcome this challenge, an innovative index-based approach to yield protection has been gaining popularity in emerging economies. Leveraging the fact that crop yields are significantly affected by exogenous and regional factors, such as weather conditions, this approach triggers payment to all farmers in a region when a pre-determined index, closely linked to the actual yields in the region, predicts a low yield. For instance, in the past decade, the Indian government has utilized weather-based index policies to subsidize more than 12 million farmers; one of the policy schemes uses the rainfall during the planting season as an index, and farmers receive a payment if the rainfall falls below a certain threshold (CCAFS 2013). Similarly, the Indonesian government launched an area-yield index policy in 2022, where farmers are subsidized if an estimation of the average yield in an area falls short of a specific benchmark (JICA 2022). As of 2017, over 20 emerging economies have started providing index-based yield protection to their farmers (Carter et al. 2017). By using an index as a proxy, this approach significantly reduces administrative costs by eliminating the need for individual yield assessments and makes yield protection a viable option for smallholder farmers in emerging economies. Another advantage of this approach is that it minimizes moral hazard because the payment to farmers is determined based on an index that typically cannot be influenced by individual farmers.

Despite these merits, determining the appropriate subsidy level of an index-based policy remains challenging. To see this, note that any agricultural subsidy policy must strike a balance between providing adequate support to farmers and avoiding the provision of excessive subsidies that lead to an oversupply of crops. This is a complex task in the presence of yield uncertainty even when yield can be directly assessed (Alizamir et al. 2019). On top of this, an index-based policy adds an additional layer of complexity because the subsidy payment is triggered based on an index rather than the actual yield of farmers. Due to the inevitable imperfect correlation between the index and the actual yield, farmers may suffer from low actual yield without receiving a payment. Similarly, farmers may receive a payment even when their actual yield is high. In this setting, the accuracy of the index, namely, its ability to accurately predict the actual yield, plays a critical role in determining the effectiveness of the policy. However, in practice, a popular approach is to determine the payment to farmers based on the predicted yield as indicated by the index value, while the accuracy of the index is typically not explicitly taken into consideration (see, e.g., Kenduiywo et al. 2021, for commonly used index policies in practice). Moreover, in the literature, while several recent studies have examined the optimal design of agricultural subsidies that are based on the actual price or actual yield (e.g., Alizamir et al. 2019, Chintapalli and Tang 2021, Guda et al. 2021), an index-based subsidy has not been formally analyzed. Therefore, our first research question is: How should a government optimally determine the subsidy amount under an index-based yield protection policy, and how does this optimal amount vary with the index accuracy?

As mentioned earlier, price protection policies are common in emerging economies. When both price and index-based yield protection policies are provided, the interplay between these two policies can be intriguing. On the one hand, because price and yield protection policies are designed to safeguard farmers from low prices and low crop yields, respectively, combining both types of policies may be more effective in protecting farmers than either policy alone. On the other hand, since both policies can potentially incentivize more farmers to plant, excessive protection levels under both policies can lead to oversupply and thus a drop in market price, which could be detrimental to farmers. Given these trade-offs, it is crucial for policymakers to understand how price and yield protection policies interact with each other, especially when yield protection is implemented through an index-based approach. This leads to our second research question: How does the presence of price protection affect the optimal design and value of index-based yield protection, and under what conditions are these two policies strategic complements or substitutes?

Lastly, under some index-based yield protection policies, the government may be able to exert a costly effort to enhance the accuracy of the index. For instance, under an area-yield index policy discussed earlier, an estimation of the average yield in an area is used as the index to determine whether to subsidize farmers in the area. To improve the accuracy of such an index, the government can, for example, increase the sample size for yield estimation or divide the area into more sub-areas so that the estimation of yield in each (sub-)area can better reflect the actual yields of individual farmers. Greater index accuracy can increase farmers' chances of receiving subsidies when their actual yield is low. However, enhancing the index accuracy may incur substantial costs. Under tight budget constraints, an increase in investment toward improving index accuracy can result in fewer funds available for subsidy payments, which may compromise the policy's impact on farmer welfare. Therefore, a careful allocation of the budget between subsidy payment and index accuracy is critical in ensuring the effectiveness of an index-based yield protection policy. This motivates our third research question: When the government can invest in improving the index accuracy, how should it optimally allocate its budget between the subsidy amount and index accuracy?

To address these research questions, we develop a game-theoretic model that consists of a unit mass of farmers and a local government. Farmers make planting decisions for a single crop and face uncertain crop yield. The uncertainty in crop yield further leads to uncertainty in the market-clearing price. In this uncertain environment, farmers exhibit risk aversion when deciding whether to plant the crop, and we capture such risk aversion by considering a mean-variance utility. The government offers an index-based yield protection subsidy policy, where the index is partially correlated with the actual yield, and farmers who plant receive a subsidy if the realized index indicates a low yield. To capture the interactions between the government and farmers, we formulate a Stackelberg game where the government first announces the subsidy policy, and then farmers determine whether to plant the crop. The government determines the subsidy amount to maximize the net benefit, defined as the total farmer surplus minus the government expenditure.

Our analysis unveils several insights for policymakers in designing index-based yield protection subsidies. First, we find that while a higher subsidy can increase farmers' expected income, it can also increase their income variance, especially when the index accuracy is low. This leads us to identify a non-monotonic relationship between the optimal subsidy amount and index accuracy: the optimal subsidy amount initially increases and then decreases in index accuracy. This result suggests that commonly used policies that base payments on the predicted yield without explicitly considering index accuracy can be suboptimal, and underscores that it is crucial for policymakers to carefully integrate the role of index accuracy into the design of index-based policies. Moreover, this inverted-U-shaped relationship between the optimal subsidy amount and index accuracy provides guidance for policymakers to tailor subsidy amounts based on the accuracy of available indices, while cautioning against the risk of over-subsidization at both lower and higher ends of index accuracy. Additionally, this result also enables us to characterize the effect of yield variability and farmer risk aversion on the optimal subsidy amount at different levels of index accuracy.

Second, with both price and index-based yield protection policies in place, we characterize the critical role of the index accuracy in determining the interplay between these two policies. Specifically, we first show that if the index accuracy is sufficiently high, price and index-based yield protection are generally strategic substitutes. This result parallels the finding by Alizamir et al. (2019), who show that in a context where farmers are subsidized based on the actual price and yield, protecting both price and yield may not offer added value over price protection alone. However, we also offer a contrasting perspective by showing that if the index accuracy is low, price and index-based yield protection can work as strategic complements, indicating that if a government must employ a low-accuracy index for yield protection (possibly due to data limitations), price protection can augment the value of index-based yield protection for farmers.

Finally, when the government can invest in improving index accuracy, we find that such investment is crucial when the crop yield variability is relatively high. Moreover, contrary to expectations, it can be optimal for the government to allocate more budget to enhance index accuracy under a tighter budget. This result underscores that a restricted budget does not automatically imply a diminished investment in index accuracy, and it holds particular relevance given the increasing budgetary constraints many governments face for crop subsidies as they also strive to address other urgent needs.

The remainder of the paper is structured as follows. In §2, we present a review of related literature. In §3, we introduce our model setup. In §4, we analyze the optimal subsidy amount under an index-based yield protection policy. In §5, we study how the presence of price protection affects the design and value of index-based yield protection. In §6, we study the joint design of the subsidy amount and index accuracy. In §7, we present a calibrated numerical study based on real data for corn production to illustrate how our insights can map to practice. In §8, we present several model extensions and show the robustness of our insights under alternative model setups. Finally, in §9, we conclude with a summary of managerial insights and directions for future research.

2. Literature Review

Our paper contributes to two emerging streams of literature on sustainable and socially responsible operations. First, we contribute to the growing literature on agricultural operations and supply chains. In this research area, a set of papers examines farmers' optimal decisions in farming activities, such as planting and farmland allocation, in the presence of yield uncertainty (Kazaz 2004, Boyabatlı et al. 2019, Zhang and Swaminathan 2020). Given the vulnerability of farmers, especially those in emerging economies, another set of papers analyzes the design of external interventions for improving the welfare of farmers, such as providing information (Liao et al. 2019, Zhou et al. 2021, Shi et al. 2023), developing online platforms for trading or equipment sharing (Ferreira et al. 2017, Levi et al. 2020, Adebola et al. 2022), implementing contract farming (Federgruen et al. 2019, de Zegher et al. 2019, Hu et al. 2019), offering government loans (Pay et al. 2022), and enabling consumers to tip farmers (Alizamir et al. 2022). Like the latter set of papers, we also study the design of interventions to improve farmer welfare, and we contribute to this literature by analyzing a novel form of government support for farmers, index-based yield protection subsidies.

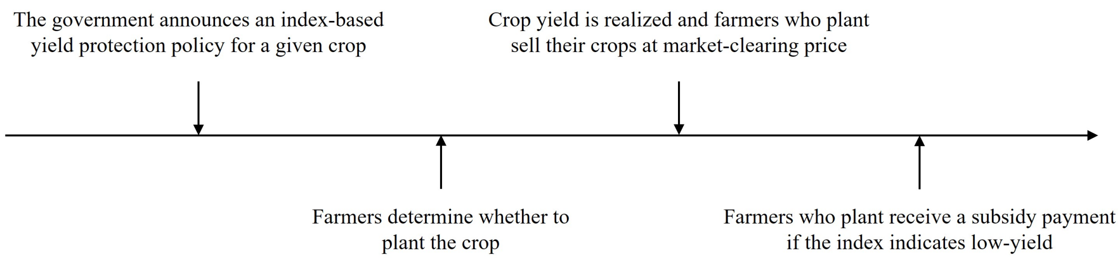
Second, our paper contributes to the operations literature that studies the optimal design of government subsidies. Within this literature, several recent papers examine agricultural subsidies.

For example, Alizamir et al. (2019) compare two crop subsidy policies in the U.S. that protect farmers from low market price and low revenue (due to either low price or low yield), respectively; Akkaya et al. (2021) study government subsidies to promote agricultural innovation; Ramaswami et al. (2018), Guda et al. (2021) and Chintapalli and Tang (2021) examine a price protection policy named minimum support price; and Fan et al. (2023) and Tang et al. (2023) analyze and compare input- and output-based subsidy policies. Among these papers, the only work that studies subsidy policies that subsidize farmers in low-yield scenarios in the presence of yield uncertainty is by Alizamir et al. (2019). Nevertheless, their focus is on policies in the U.S. that subsidize farmers based on actual crop yields, which are typically not suitable for emerging economies with small farm sizes. In contrast, we study an index-based approach for yield protection for smallholder farmers.

Several recent papers study government subsidies in other settings, including the use of subsidies for promoting green energy and technology (e.g., Cohen et al. 2016, Babich et al. 2020) or improving the availability and affordability of essential medicines (e.g., Taylor and Xiao 2014, Levi et al. 2019, Arifoğlu and Tang 2022). While these studies consider subsidies offered to consumers or large producers, we focus on subsidies offered to smallholder farmers. Moreover, we study an index-based approach under which the subsidy payment is determined based on an imperfect predictor of actual outcomes, while in all aforementioned studies, the subsidy payment is determined based on actual outcomes.

Finally, our work relates to the agricultural economics literature that studies interventions for protecting smallholder farmers from low crop yields. Within this literature, a few recent papers empirically examine the benefits of offering index-based yield protection subsidies in emerging economies (e.g., Freudenreich and Mußhoff 2018, Cai et al. 2020). In addition, this literature also explores the construction of indices using various data sources (e.g., Flatnes and Carter 2016, Kenduiywo et al. 2021, Chen et al. 2023), and examines farmers' demand and willingness to pay for index-based insurance products for which farmers need to pay a premium to enroll (e.g., Clarke 2016, Takahashi et al. 2016, Ghosh et al. 2021). Our work complements these studies by analytically studying the optimal design of government subsidies for index-based yield protection. Specifically, we build an analytical model that endogenizes farmers' planting decisions and thus the crop's total supply and its market-clearing price. This enables us to capture the broad implications of index-based yield protection subsidies regarding their effects on the equilibrium planting amount, the surplus of farmers and consumers, and the government expenditure, thereby providing nuanced insights to guide the design of such subsidies.

Figure 1 Sequence of Events



3. Model

We consider a unit mass of farmers and a local government. Farmers make planting decisions for a single crop and face uncertain crop yield. To protect farmers from low crop yield, the government offers an index-based yield protection policy such that farmers who plant will receive a subsidy if the realized index indicates a low yield. To model the government's subsidy choice and the farmers' planting decisions, we formulate a Stackelberg game where the government first announces the subsidy policy, and then farmers determine whether or not to plant the crop (see Figure 1 for an illustration of the sequence of events). Overall, our model captures three features that are salient in the context of index-based yield protection: (1) uncertain crop yield; (2) planting decisions of heterogeneous and risk-averse farmers; and (3) subsidy payment based on an index that is imperfectly correlated to the crop yield. We next describe these key model elements in detail.

Yield uncertainty and market-clearing price. Crop yields are profoundly shaped by various environmental factors like weather conditions. The inherent uncertainties linked to these elements, including unexpected rainfall patterns, varying temperatures, or extreme weather events, can cause drastic variations in crop yields. To capture uncertain crop yields, we represent the yield level of each farmer by a random variable Y. For analytical simplicity, we assume that Y can take on two different outcomes: a low yield, $\mu - \sigma$, and a high yield, $\mu + \sigma$, where $\mu > 0$ and $\sigma \in (0, \mu]$. Both possibilities occur with equal probability. Therefore, the mean (i.e., expected value) of the crop yield is $\mathbb{E}[Y] = \mu$, and the variance is $\mathbb{V}\text{ar}[Y] = \sigma^2$. This two-point distribution is commonly employed by the existing literature to model yield uncertainty (Boyabatlı et al. 2017, Agrawal et al. 2022). Moreover, this assumption does not drive our main findings; we later show that our insights continue to hold under more realistic yield distributions (see §7). To facilitate analytical tractability, we adopt an assumption from existing studies on agricultural supply chains by considering that the yield is perfectly correlated among all farmers (Hu et al. 2019, Alizamir et al. 2019). This assumption aligns with the scenarios where exogenous and regional factors, such as weather conditions, primarily drive yield variability. Furthermore, we relax this assumption in §7, demonstrating that our main insights remain intact even when the yields of individual farmers are different from each other.

Given the uncertain nature of yields, the overall supply of the crop, and consequently, its market price, are subject to fluctuation. Suppose a total of $x \in [0,1]$ farmers choose to plant. Then, the total supply of the crop, given the yield level Y, equates to xY. In line with several existing studies on agricultural supply chains (e.g., Hu et al. 2019, Alizamir et al. 2019), we characterize the market-clearing price of the crop as $a - bxY$, a decreasing function in the total supply xY, where $a > 0$ denotes the maximum possible price and $b > 0$ the sensitivity of the market price to the supply of the crop. To ensure a positive price for any $x \in [0,1]$, we assume $a - b(\mu + \sigma) > 0$.

Risk-averse farmers. Smallholder farmers are particularly vulnerable to income variability, and empirical studies have revealed that they are typically risk-averse in their farming decisions (see, e.g., Chavas et al. 2010). Accordingly, we consider that farmers are risk-averse when deciding whether or not to plant the crop, and we capture such risk aversion through a mean-variance utility. When a total of x farmers choose to plant, given the market price $a - bxY$, each farmer who plants will earn a revenue $(a - bxY)Y$ from selling the crop. To account for the diverse production efficiencies among farmers, we consider a heterogeneous production cost h that is uniformly distributed over $[0, c]$. Consequently, when government subsidy is not provided and if a total of x farmers plant, the net income of the farmer with production cost h from planting is

$$\pi(h|x) = (a - bxY)Y - h.$$

Then, her mean-variance utility is

$$u(h|x) = \mathbb{E}[\pi(h|x)] - \lambda \mathbb{V}\mathrm{ar}[\pi(h|x)] = a\mu - bx(\mu^2 + \sigma^2) - h - \lambda \left(a\sigma - 2bx\mu\sigma\right)^2, \tag{1}$$

where $\lambda > 0$ denotes the risk aversion coefficient.

Each farmer chooses to plant if and only if her utility defined by Equation (1) is positive. We call the number of farmers who choose to plant in equilibrium the *equilibrium planting amount*. For ease of exposition, we assume that farmers' maximum production cost (i.e., c) is higher than a threshold c_1, which ensures that there exists a unique equilibrium planting amount that is strictly less than one without subsidy.[1] Similarly, we assume $\lambda < \frac{\mu+\sigma}{2a\sigma^2}$, which avoids another extreme case where no one plants. Under these conditions, the following lemma characterizes the equilibrium planting amount when no government subsidy is provided, which we denote as x_0^*.

Lemma 1 (Equilibrium planting amount without government subsidy)**.** *The function $u(xc|x)$ is quadratic concave and strictly decreasing in x for $x \in [0,1]$. The equilibrium planting amount without government subsidy, x_0^*, is the unique positive solution to $u(xc|x) = 0$, given by*

$$x_0^* = \frac{-\beta_0 + \sqrt{\beta_0^2 - 4\alpha_0\gamma_0}}{2\alpha_0} \in (0,1), \tag{2}$$

[1] Specifically, we assume $c > c_1 := \max\{a\mu - b(\mu^2+\sigma^2) - \lambda(a\sigma - 2b\mu\sigma)^2, 4\lambda ab\mu\sigma^2 - b(\mu^2+\sigma^2), 2\lambda ab(\mu-\sigma)^2 \frac{\mu\sigma}{\mu+\sigma} - \frac{1}{2}b(\mu-\sigma)^2\}$.

where $\alpha_0 = 4\lambda b^2\mu^2\sigma^2$, $\beta_0 = b(\mu^2+\sigma^2)+c-4\lambda ab\mu\sigma^2$, *and* $\gamma_0 = \lambda a^2\sigma^2 - a\mu$.

For any given planting amount x, we denote each farmer's revenue in the low- and high-yield scenarios as $R_L(x) = [a-bx(\mu-\sigma)](\mu-\sigma)$ and $R_H(x) = [a-bx(\mu+\sigma)](\mu+\sigma)$, respectively. For ease of exposition, we assume $R_L(x_0^*) < R_H(x_0^*)$. This assumption represents a prominent scenario in which yield protection is most crucial, and it aligns with practical situations where farmers typically suffer from low crop yields (Mgbenka et al. 2016, Fan and Rue 2020).

Index-based yield protection by the government. To safeguard smallholder farmers from the risks associated with yield uncertainty, the government provides an index-based yield protection policy. As discussed in §1, an index-based approach avoids the costly yield assessment of individual farmers,[2] making yield protection a viable option for smallholder farmers in emerging economies. The index can take various forms, such as a weather-based index determined by the rainfall level and temperature. We let a random variable I denote the value of the index. In line with the yield variable Y, we assume that I can also take two values. Specifically, $I=1$ indicates a prediction of low yield, while $I=0$ indicates a prediction of high yield. Hence, under the index-based yield protection policy, farmers who plant will receive a subsidy $s \geq 0$ if $I=1$.

For ease of exposition, we assume that the index variable I has the same marginal distribution as the yield variable Y. That is, $I=1$ with probability $\frac{1}{2}$ and $I=0$ otherwise. Later on, we will relax this assumption and explore situations where the index I and the yield level Y have different marginal distributions (see §8). Further, as discussed in §1, the index is generally an imperfect prediction of the actual yield level. To capture such an imperfect correlation between the index I and the yield level Y, we define r as the conditional probability $\mathbb{P}(I=1|Y=\mu-\sigma)$. In other words, r indicates the likelihood that the index accurately predicts a low yield when the actual yield level is low. This probability represents the *accuracy* of the index in yield prediction. Without loss of generality, we assume that the index is at least as accurate as a random guess, that is, $r \geq \frac{1}{2}$. This condition is equivalent to $\mathbb{P}(I=1|Y=\mu-\sigma) \geq \mathbb{P}(I=1)$.

Given the index I and the subsidy amount s, if a total of x farmers plant, the net income of the farmer with production cost h from planting is adjusted as follows:

$$\pi(h|x,s) = (a-bxY)Y + sI - h.$$

Her mean-variance utility is now given by:

$$\begin{aligned} u(h|x,s) &= \mathbb{E}[\pi(h|x,s)] - \lambda \mathbb{V}\mathrm{ar}[\pi(h|x,s)] \\ &= a\mu - bx(\mu^2+\sigma^2) - h + \frac{s}{2} - \lambda\left[(a\sigma - 2bx\mu\sigma)^2 + \frac{s^2}{4} + s\left(r-\frac{1}{2}\right)(2a\sigma - 4bx\mu\sigma)\right]. \end{aligned} \tag{3}$$

[2] In our base model, it may seem easy to assess farmers' yields given that they are assumed to be perfectly correlated. However, we note that we make the assumption of perfect correlation only for analytical tractability, and we later show that our key insights remain intact when farmers' yields are no longer perfectly correlated (see §7).

Same as before, we consider that each farmer chooses to plant if and only if her utility is positive. Let $x^*(s)$ denote equilibrium planting amount when the subsidy amount is equal to s.

To focus our analysis on a realistic range of subsidy amounts, we assume that farmers' revenue in the low-yield scenario plus the subsidy payment does not exceed their revenue in the high-yield scenario. Specifically, let $\bar{s}$ represent the lowest subsidy amount s at which $R_L(x^*(s)) + s \geq R_H(x^*(s))$, that is,

$$\bar{s} = \inf\{s \geq 0 : R_L(x^*(s)) + s \geq R_H(x^*(s))\},$$

and whenever we discuss the subsidy amount s in the remainder of our analysis, we focus on $s \in [0, \bar{s}]$. We note that this is without loss of generality because if the subsidy amount s has already exceeded $\bar{s}$, it is straightforward to show that farmers will be better off if the government starts subsidizing them when the index indicates a high yield rather than a low yield. Moreover, our key insights remain valid even if we consider any $s \geq 0$.

Optimal subsidy design. In practice, a primary goal of agricultural subsidies is to improve the welfare of farmers (Hemming et al. 2018, Chintapalli and Tang 2021). Therefore, similar to Chintapalli and Tang (2021), who study price protection subsidies, we consider that the government chooses the subsidy $s \in [0, \bar{s}]$ to maximize the net benefit, which is defined as the expected total farmer surplus (i.e., the aggregation of all farmers' income) minus the expected government expenditure. Specifically, given the subsidy amount s and the corresponding equilibrium planting amount $x^*(s)$, the expected total farmer surplus is $\int_0^{x^*(s)} \mathbb{E}[(a - bx^*(s)Y)Y + sI - xc]\,\mathrm{d}x$, and the expected government expenditure is $x^*(s)\mathbb{E}[sI]$. Therefore, the net benefit is as follows:

$$\begin{aligned} v(s) &= \int_0^{x^*(s)} \mathbb{E}[(a - bx^*(s)Y)Y + sI - xc]\,\mathrm{d}x - x^*(s)\mathbb{E}[sI] \\ &= a\mu x^*(s) - \left(b(\mu^2 + \sigma^2) + \frac{1}{2}c\right)(x^*(s))^2 \end{aligned} \tag{4}$$

From Equation (4), we observe that the net benefit $v(s)$ is a quadratic concave function of the equilibrium planting amount $x^*(s)$. Let x^{opt} denote the planting amount that maximizes the net benefit. Then, $x^{opt} = \min\{\frac{a\mu}{2b(\mu^2+\sigma^2)+c}, 1\}$, and the optimal subsidy amount, which we denote as s^*,[3] must satisfy that, the corresponding equilibrium planting amount $x^*(s^*)$ is equal to or as close to x^{opt} as possible. To avoid an uninteresting case where there is already too much planting even without subsidy, we focus on scenarios with $x_0^* \leq x^{opt}$ in the remainder of this paper. This assumption can be re-written as that farmers' risk aversion coefficient λ is not too low.

We remark that, alongside enhancing farmer surplus, another possible function of agricultural subsidies is to bolster consumer surplus by ensuring a sufficient supply of crops (Alizamir et al. 2019,

[3] If there are more than one s that maximizes $v(s)$, we define s^* as the smallest one.

Akkaya et al. 2021). We later extend our model to consider an alternate governmental objective that incorporates consumer surplus and demonstrate that our key findings remain robust (§8). Also, in our base model, we consider a scenario wherein the government works with a given index and optimizes the subsidy amount s. We later develop our model to include a scenario where the accuracy of the index can be improved by exerting a costly effort, thereby necessitating the government to make joint decisions regarding the subsidy amount and the index accuracy (§6).

Table 1 below summarizes the notation of the basic model parameters defined in this section.

Table 1 Notation

$Y =$	crop yield, where $Y = \mu - \sigma$ with probability $\frac{1}{2}$ and $Y = \mu + \sigma$ with probability $\frac{1}{2}$
$a =$	maximum possible price
$b =$	sensitivity of market price to the supply of the crop
$h =$	farmers' production cost, which is uniformly distributed over $[0, c]$
$\lambda =$	farmers' risk aversion coefficient
$I =$	index, where $I = 1$ (low yield) with probability $\frac{1}{2}$ and $I = 0$ (high yield) with probability $\frac{1}{2}$
$r =$	index accuracy, which is defined as $\mathbb{P}(I = 1 \mid Y = \mu - \sigma)$
$s =$	subsidy amount to farmers when the index indicates the yield level is low (i.e., when $I = 1$)
$x^*(s) =$	equilibrium planting amount under subsidy amount s
$v(s) =$	net benefit under subsidy amount s

4. Optimal Subsidy of Index-Based Yield Protection

In this section, we first examine farmers' equilibrium planting amount under any given subsidy amount for index-based yield protection (§4.1), and then analyze the structure of the optimal subsidy amount and discuss its managerial implications (§4.2).

4.1. Equilibrium Planting Amount under Given Subsidy

Given farmers' utility defined in Equation (3), we are able to derive a closed-form expression of the equilibrium planting amount under given subsidy amounts, as shown in the following lemma.

Lemma 2 (Equilibrium planting amount with government subsidy)**.** *For any given government subsidy amount* $s \in [0, \bar{s}]$*, the function* $u(xc|x, s)$ *is quadratic concave and strictly decreasing in* x *for* $x \in [0, 1]$*. Moreover, for any given* $s \in [0, \bar{s}]$*, let* $\hat{x}(s)$ *denote the unique positive solution to* $u(xc|x, s) = 0$*. Then, the equilibrium planting amount is* $x^*(s) = \min\{\hat{x}(s), 1\} \in (0, 1]$*, with*

$$\hat{x}(s) = \frac{-\beta(s) + \sqrt{\beta^2(s) - 4\alpha\gamma(s)}}{2\alpha}, \tag{5}$$

where $\alpha = \alpha_0 = 4\lambda b^2\mu^2\sigma^2$, $\beta(s) = b(\mu^2+\sigma^2) + c - 4\lambda ab\mu\sigma^2 + 4\lambda b(r-\frac{1}{2})\mu\sigma s$, *and* $\gamma(s) = \lambda a^2\sigma^2 - ay - 2\lambda a(r-\frac{1}{2})\sigma s - \frac{1}{2}s + \frac{1}{4}\lambda s^2$. *Further, there exists a threshold* c_2 *such that if* $c \geq c_2$, *then the equilibrium planting amount under given subsidy amount* $s \in [0, \bar{s}]$ *is* $x^*(s) = \hat{x}(s) \in (0,1)$.

Lemma 2 characterizes the equilibrium planting amount under given subsidy amounts. Moreover, the lemma demonstrates that if farmers' production cost is sufficiently high (i.e., $c \geq c_2$), then the equilibrium planting amount is strictly less than one. For ease of exposition, we focus on the case where $c \geq c_2$ in the remainder of our analysis, which allows us to restrict our attention to the interior solution of the equilibrium amount and avoid unnecessary technical complications. We also note that our key insights remain valid even if this assumption is relaxed. Building on Lemma 2, we next explore how the subsidy amount s affects the equilibrium planting amount.

Proposition 1 (Effect of subsidy amount on equilibrium planting amount)**.** *The equilibrium planting amount* $x^*(s)$ *is a concave function of* s. *Moreover, there exists a threshold* $r_1 \in [\frac{1}{2}, 1)$ *such that if* $r < r_1$, *then* $x^*(s)$ *increases in* s *if and only if* $s \leq s_1$, *where* $s_1 := \arg\max_{s\in[0,\bar{s}]} x^*(s) \in (0, \bar{s})$ *and* s_1 *increases in* r. *If* $r \geq r_1$, *then* $x^*(s)$ *increases in* s.

Proposition 1 demonstrates that the effect of an increase in the subsidy amount s can vary significantly across different levels of index accuracy. In particular, if the index accuracy is higher than a certain threshold (i.e., $r \geq r_1$), a larger subsidy amount always leads to a higher equilibrium planting amount, although with a diminishing return. However, if the index accuracy is low (i.e., $r < r_1$), a larger subsidy amount encourages more planning only up to a certain point (i.e., s_1). When the subsidy amount exceeds this threshold, a further increase in subsidy no longer incentivizes more planting. Moreover, the threshold s_1 is increasing in the index accuracy r, implying that the effectiveness of an index-based subsidy in incentivizing farmers to plant is enhanced when the index accuracy r is higher. To gain intuition behind these results, the next proposition studies how an increase in the subsidy amount affects the variance of farmer income.[4]

Proposition 2 (Effect of subsidy amount on farmer income variance)**.** *The variance of farmer income* $\mathbb{V}ar[\pi(h|x^*(s), s)]$ *is a convex function of* s. *Moreover, if* $r < 1$, *then* $\mathbb{V}ar[\pi(h|x^*(s), s)]$ *decreases in* s *if and only if* $s \leq s_2$, *where* $s_2 := \arg\min_{s\in[0,\bar{s}]} \mathbb{V}ar[\pi(h|x^*(s), s)] \in (0, \bar{s})$ *and* s_2 *is lower than the threshold* s_1 *in Proposition 1. If* $r = 1$, *then* $\mathbb{V}ar[\pi(h|x^*(s), s)]$ *decreases in* s.

Proposition 2 reveals that, if the index accuracy is one (i.e., subsidy is provided based on actual crop yield), the variance of farmer income decreases in the subsidy amount. Such an effect is

[4] For ease of exposition, we include the details about how the subsidy amount affects the expected farmer income (i.e., $\mathbb{E}[(a - bx^*(s)Y)Y + sI - h])$ within the proof of Proposition 2. Specifically, an increase in the subsidy leads to an increase in the expected farmer income as long as the price sensitivity b is not too high.

expected because a yield-protection subsidy is intended to improve farmers' income for low-yield scenarios.

However, with imperfect index accuracy ($r < 1$), the variance of farmer income is non-monotonic in the subsidy amount. When the subsidy increases from a small level, the variance of farmer income first decreases, resulting in an increase in the equilibrium planting amount. Such an increase in planting amount, in turn, can lead to a decrease in the market price. Moreover, when the index accuracy is low, there is a significant probability that farmers suffer from low actual yield without receiving a subsidy. The combination of a reduced price, low yield, and no subsidy payment leads to a scenario with an even lower income than when no subsidy is provided. Similarly, due to the imperfect index, it is also possible that farmers receive a subsidy when the actual yield is high, resulting in a higher income than when no subsidy is provided. Consequently, after the subsidy amount exceeds a certain threshold (i.e., s_2), a further increase in subsidy leads to an increase in the variance of farmer income. This makes the subsidy less effective in increasing farmers' utility and incentivizing them to plant.

Collectively, Propositions 1 and 2 provide important insights regarding how an index-based yield protection policy with a naturally imperfect index (i.e., $r < 1$) may have distinct implications from a policy that subsidizes farmers based on actual yields (i.e., $r = 1$). In the following corollary, we further characterize how the equilibrium planting amount and the variance of farmer income vary with index accuracy.

Corollary 1 (Effect of index accuracy on equilibrium planting amount and farmer income variance)**.** *Consider a fixed subsidy amount s. Then, the equilibrium planting amount increases in the index accuracy r, and the variance of farmer income decreases in the index accuracy r.*

Corollary 1 shows that under a given subsidy, if the index accuracy is higher, the variance of farmer income will be lower, and more farmers will be incentivized to plant. Thus, this result reiterates that a higher index accuracy enhances the effectiveness of the subsidy in incentivizing farmers to plant.

4.2. Optimal Subsidy Amount

From our analysis in §4.1, it is evident that the index accuracy plays a critical role in determining the effectiveness of an index-based yield protection subsidy. Building on these results, we next characterize how the optimal subsidy amount s^* varies with the index accuracy.

Theorem 1 (Effect of index accuracy on optimal subsidy)**.** *There exists a threshold r_2, where $\frac{1}{2} \leq r_2 \leq r_1$, such that if the index accuracy is low ($r < r_2$), the optimal subsidy amount s^* increases in the index accuracy r; otherwise ($r \geq r_2$), it decreases in the index accuracy r.*

Theorem 1 uncovers a non-monotonic relationship between the optimal subsidy amount and the index accuracy, showing that the optimal subsidy should first increase and then decrease in index accuracy. The key reasoning behind this result is as follows: When the index accuracy is low, an increase in the subsidy amount can increase the variance of farmer income, and therefore it is less effective in improving farmers' utility and incentivizing them to plant (as shown in Proposition 1). Thus, a low subsidy amount maximizes the net benefit. When the index accuracy is higher, the policy can more effectively protect farmers, justifying a higher subsidy. When the index accuracy further increases, however, a high subsidy amount can lead to oversupply and thus a low market-clearing price, which can be detrimental to farmers. In this case, a low subsidy is also preferred.

As mentioned earlier, the existing literature on the optimal design of agricultural subsidies has predominately focused on policies that subsidize farmers based on the actual price and yield. To the best of our knowledge, we are the first to observe this nuanced relationship between the optimal subsidy and index accuracy under an index-based yield protection policy.

Theorem 1 offers critical insights for the design of index-based yield protection subsidies. First, in current index-based yield protection policies, the payment to farmers is often determined based on the predicted yield as indicated by the index value, whereas the index accuracy is often not explicitly taken into consideration (see, e.g., Kenduiywo et al. 2021, for commonly used index policies in practice). Our analysis suggests that such an approach can be suboptimal and highlights that it is crucial for policymakers to carefully choose the subsidy amount by explicitly taking into account the role of index accuracy. Second, by uncovering a non-monotonic, inverted-U-shaped relationship between the optimal subsidy amount and index accuracy, our findings provide insights for policymakers to tailor subsidy amounts based on the accuracy of available indices, ensuring effective farmer support while avoiding over-subsidizing. Finally, with governments employing increasingly accurate indices due to technological progress and new data sources (Benami et al. 2021), our findings also offer timely guidance for effectively adapting subsidy policies to such increases in index accuracy.

With Theorem 1, we are able to further characterize how two critical environmental and economic factors, namely the variability in yield (i.e., σ) and farmers' risk-aversion level (i.e., λ), affect the optimal subsidy amount. Additionally, we also explore the role of index accuracy in mediating these results.

Proposition 3 (Effect of yield variability and risk aversion on optimal subsidy)**.** *(i) The optimal subsidy amount s^* increases in the yield variability σ.*
(ii) There exists a threshold $\lambda_1 > 0$ such that if farmers' risk aversion level is low ($\lambda < \lambda_1$), the optimal subsidy amount s^* increases in the risk aversion level λ; otherwise ($\lambda \geq \lambda_1$), it decreases in λ. Moreover, the threshold λ_1 increases in the index accuracy r.

Proposition 3 (i) suggests that, as expected, the optimal subsidy amount should increase when yield becomes more variable. On the other hand, while one may expect that a higher subsidy is needed when farmers are more risk-averse, Proposition 3 (ii) reveals that this is only true when farmers' risk aversion level is low (i.e., $\lambda < \lambda_1$), and the opposite is true otherwise; moreover, a lower index accuracy results in a lower threshold of risk aversion level (λ_1) at which the optimal subsidy begins to decrease. This result can be understood as follows: As discussed earlier, when the index accuracy is low, there is a significant probability that an increase in subsidy will decrease the lowest possible income for farmers (due to a decrease in price and a high probability of not receiving the subsidy when the actual yield is low), resulting in an increase in their income variance (Proposition 2). Therefore, with a low index accuracy and if farmers are highly risk averse, the value of index-based yield protection is limited, and thus a low subsidy amount is optimal.

5. Interplay between Yield and Price Protection

As discussed in §1, price protection policies that protect farmers from low market prices have been widely implemented in emerging economies (Chintapalli and Tang 2021). In particular, a popular price protection policy is called minimum support price (MSP), under which the government pays farmers the difference between the market price and a pre-specified floor price if the market price is below the floor price (Antonaci et al. 2014). In this section, we study how the presence of such an MSP policy affects the optimal design of an index-based yield protection policy.

We consider an MSP policy with floor price m and an index-based yield protection policy with index I and subsidy amount s. Then, if a total of x farmers plant, the net income of the farmer with production cost h from planting is given by $\pi_m(h|x,s) = \max\{(a - bxY), m\}Y + sI - h$. Therefore, her mean-variance utility is given by $u_m(h|x,s) = \mathbb{E}[\pi_m(h|x,s)] - \lambda\mathbb{V}\text{ar}[\pi_m(h|x,s)]$, where $\lambda > 0$ denotes the risk aversion coefficient. Similar to before, we consider that each farmer chooses to plant if her mean-variance utility is positive, and let $x_m^*(s)$ denote the equilibrium planting amount.

In order to shed light on how the presence of price protection affects the optimal design of index-based yield protection, we consider a fixed floor price m and that the government chooses the yield protection subsidy amount s to maximize the net benefit, which, same as in §3, is defined as the expected total farmer surplus minus the expected government expenditure. Specifically, given the subsidy amount s, the floor price m, and the corresponding equilibrium planting amount $x_m^*(s)$, the expected total farmer surplus is $\int_0^{x_m^*(s)} \mathbb{E}[\max\{(a - bx_m^*(s)Y), m\} \times Y + sI - xc] \, \mathrm{d}x$, and the expected government expenditure is the sum of two components: $x_m^*(s)\mathbb{E}[sI]$ (for yield protection) and $x_m^*(s)\mathbb{E}[\max\{m - (a - bx_m^*(s)Y), 0\} \times Y$ (for price protection). Therefore, the net benefit is

$$v_m(s) := \int_0^{x_m^*(s)} \mathbb{E}[\max\{(a - bx_m^*(s)Y), m\} \times Y + sI - xc] \, \mathrm{d}x - x_m^*(s)\mathbb{E}[sI]$$

$$- x_m^*(s)\mathbb{E}[\max\{m-(a-bx_m^*(s)Y),0\}\times Y]$$
$$= a\mu x_m^*(s) - \left(b(\mu^2+\sigma^2)+\frac{1}{2}c\right)(x_m^*(s))^2. \quad (6)$$

Equation (6) has a similar structure as Equation (4). Therefore, x^{opt} defined in §3 remains the planting amount that maximizes the net benefit. Further, the optimal subsidy amount, which we now denote as s_m^*,[5] must satisfy that the corresponding equilibrium planting amount $x_m^*(s_m^*)$ is equal to or as close to x^{opt} as possible. Before proceeding to our analysis, we note that a higher floor price m may not always benefit farmers because a larger m may lead to more planting and thus a price drop in high-price scenarios (Chintapalli and Tang 2021). In order to focus on a reasonable range of the floor price, in the remainder of our analysis, we consider $m\in[0,\bar{m}]$, where $\bar{m}$ denotes the smallest m such that either $x_m^*(0)\geq x^{opt}$ or $\partial x_m^*(0)/\partial m\leq 0$. We remark that this is reasonable because if $x_m^*(0)>x^{opt}$ or $\partial x_m^*(0)/\partial m<0$, then it is straightforward to show that reducing the floor price can increase the net benefit. In addition, to focus on a practical range of yield protection subsidy s, similar to before, we consider that for a fixed m, farmers' revenue in the low-yield scenario plus the subsidy s does not exceed their revenue in the high-yield scenario. We slightly abuse our notation and continue to let $\bar{s}$ denote this upper bound for the yield protection subsidy s.

We first show how the index-based yield protection subsidy s affects the equilibrium planting amount with the presence of a floor price m.

Proposition 4 (Effect of yield protection subsidy on equilibrium planting amount)**.** *Consider a fixed floor price m. There exist two thresholds $\frac{1}{2}\leq r_{m,1}\leq r_{m,2}\leq 1$ such that if the index accuracy is low ($r<r_{m,1}$), $x_m^*(s)$ increases in s if and only if $s\leq s_{m,1}$, where $s_{m,1}:=\arg\max_{s\in[0,\bar{s}]} x_m^*(s)\in[0,\bar{s})$ and $s_{m,1}$ increases in both r and m; if the index accuracy is high ($r\geq r_{m,2}$), $x^*(s)$ increases in s.*

Proposition 4 characterizes the impact of the index-based yield protection subsidy s on the equilibrium planting amount, examining how the index accuracy r and the floor price m may influence this effect. First, consistent with Proposition 1, Proposition 4 shows that, in the presence of a price protection policy, a higher index-based yield protection subsidy can stimulate more planting up to a threshold $s_{m,1}$, and the threshold increases in the index accuracy. This underlines the increased effectiveness of a subsidy in motivating planting with greater index accuracy r.

Second, Proposition 4 reveals that the threshold $s_{m,1}$ also increases with the floor price m, indicating that a larger floor price can boost the utility improvement offered by an index-based yield protection subsidy. The rationale behind this result can be explained as follows: Refer back

[5] If there are more than one s that maximizes $v_m(s)$, we define s_m^* as the smallest one.

to Proposition 1, which state that when the subsidy amount increases from a small value, more farmers are incentivized to plant, which results in a lower market-clearing price and potentially a higher variance of farmer income. As a result, a further increase in subsidy becomes less effective in increasing farmers' utility. However, a floor price m ensures farmers a minimum selling price, mitigating the aforementioned concerns about low market price and enhancing the benefit of a subsidy increase.

Based on these results, we next study how the floor price m affects the optimal subsidy under index-based yield protection. Since both price and yield protection can incentivize more farmers to plant, it might seem intuitive that these two policies would function as strategic substitutes; that is, the value of yield protection will be smaller in the presence of price protection. Consistent with this intuition, Alizamir et al. (2019) show that in a context where farmers are subsidized based on the actual price and yield, safeguarding both price and yield may not offer added value over price protection alone. However, our following proposition presents a contrasting perspective, demonstrating that this intuition does not necessarily hold under index-based yield protection.

Theorem 2 (Effect of floor price on optimal subsidy)**.** *Let $r_{m,1}, r_{m,2}$ be the two thresholds defined in Proposition 4. Then, if the index accuracy is low ($r < r_{m,1}$), the optimal subsidy amount s_m^* increases in the floor price m; if the index accuracy is high ($r \geq r_{m,2}$), the optimal subsidy amount s_m^* decreases in the floor price m.*

Theorem 2 presents a nuanced perspective on the relationship between price protection and index-based yield protection. When the index accuracy is high (i.e., $r > r_{m,2}$), we show that, as expected, a higher floor price does indeed lead to a lower optimal yield protection subsidy. This implies that under high index accuracy, price protection and index-based yield protection act as strategic substitutes.

However, the proposition also shows that this is no longer true when the index accuracy is low ($r \leq r_{m,1}$). In this case, a higher floor price results in a higher optimal yield protection subsidy, implying that price protection and index-based yield protection function as strategic complements. The logic behind this counter-intuitive outcome relates back to Theorem 1, which states that when index accuracy is low and in the absence of price protection, the optimal yield protection subsidy should be kept low, because an increased subsidy can increase farmers' income variance due to a drop in market price and the high possibility of not receiving the subsidy in the low-yield scenario. However, as Proposition 4 reveals, price protection can mitigate the concern about low market price, making a higher yield protection subsidy more appealing. Thus, with low index accuracy, the presence of price protection can justify a higher optimal subsidy for index-based yield protection compared to scenarios without price protection.

Theorem 2 offers important insights for the design of an index-based yield protection policy along with a price protection policy. On the one hand, it establishes that the two policies act as strategic complements when the index accuracy is low. Practically, this suggests that if a government is constrained to employ a low-accuracy index for yield protection (possibly due to data limitations), and if the current floor price is low or non-existent, it may be advantageous to raise both the floor price and the yield protection subsidy. This approach leverages the complementary nature of these two policies under conditions of low index accuracy. On the other hand, the proposition also demonstrates that price protection and index-based yield protection act as strategic substitutes when index accuracy is high. This finding indicates that if the index is fairly accurate, an increase in one type of subsidy should be accompanied by a reduction in the other. Overall, we believe these insights contribute to a more thorough understanding of the interplay between price and index-based yield protection policies. By recognizing the strategic complementarity and substitution effects between these two policies, policymakers can better tailor their interventions to fully exploit the synergy of these policies yet avoid over-subsidizing.

6. Joint Design of Subsidy Amount and Index Accuracy

In this section, we consider a scenario where the government not only chooses the subsidy amount but also can exert a costly effort to improve the index accuracy. As discussed in §1, a practical example for this scenario is the design of an area-yield index that uses a sample average estimate of the area-level yield to determine whether to subsidize farmers in the area. Under such an index, the government can achieve a higher index accuracy by increasing the sample size of the area or dividing the area into more sub-areas so that the estimation of yield within each sub-area is more accurate.

A higher index accuracy can intuitively benefit farmers by lessening the chance that they miss out on a subsidy despite suffering a low yield. However, enhancing index accuracy often comes with costs (e.g., enlarging sample sizes or considering smaller sub-areas under an area-yield index could lead to higher administrative expenses). When operating under budgetary constraints, the government may find that investing more in improving index accuracy results in reduced funds available for subsidy payouts. This presents a delicate balancing act between offering higher subsidies to farmers and pursuing greater index accuracy. We delve into this tension, investigating how the government can optimally navigate the dual tasks of determining subsidy amounts and index accuracy.

Let $r_0 \in [\frac{1}{2}, 1]$ denote the status quo index accuracy, i.e., the accuracy that is readily achievable before exerting a costly effort. We consider a quadratic accuracy improvement cost $\phi(r) := \kappa(r - r_0)^2$, reflecting the increasing marginal cost of improving the index accuracy. Under any given government's decision of subsidy amount s and index accuracy r, let $x^*(s, r)$ denote the equilibrium planting amount and $\psi(s, r) := x^*(s, r)\mathbb{E}[sI]$ the expected total subsidy payment to farmers.

Further, let $\bar{s}$ denote the upper bound on the subsidy amount defined in §3 and let $B > 0$ denote the government's budget level. Then, the government's problem is to choose $r \in [r_0, 1]$ and $s \in [0, \bar{s}]$ that satisfy the budget constraint $\psi(s,r) + \phi(r) \leq B$ to maximize the net benefit (i.e., the expected total farmer surplus minus the expected government expenditure) defined as follows:

$$v(s,r) = \int_0^{x^*(s,r)} \mathbb{E}[(a - bx^*(s,r)Y)Y + sI - xc]\, \mathrm{d}x - \psi(s,r) - \phi(r). \tag{7}$$

Let s^* and r^* denote the jointly optimal subsidy amount and index accuracy, respectively.[6] We first study how a key factor in the context of yield protection, the yield variability (i.e., σ), affects the government's decisions in the following proposition.

Proposition 5 (Effect of yield variability on jointly optimal subsidy amount and index accuracy)**.** *For any given budget level $B > 0$, there exist two thresholds σ_1 and σ_2, where $0 \leq \sigma_1 < \sigma_2 \leq \mu$, such that if $\sigma \leq \sigma_1$, $r^* = r_0$ and s^* increases in σ; if $\sigma_1 \leq \sigma < \sigma_2$, r^* increases in σ.*

Proposition 5 suggests that when the yield variability is low (i.e., $\sigma < \sigma_1$), it is less important to invest in index accuracy, and it is sufficient to use the budget to subsidize farmers. It is because farmers' income for the low-yield case is not too low when the yield variability is low. In this case, *how much* subsidy to receive is more critical than *when* to receive it for farmers. However, farmers' income in the low-yield case can be very low when the yield variability is higher. In this case, farmers can be significantly better off if there is a higher chance of receiving the subsidy when the actual yield is low. That is, index accuracy becomes more important for farmers. Accordingly, as shown in Proposition 5, the optimal index accuracy increases in the yield variability when the yield variability exceeds a certain threshold (i.e., $\sigma \geq \sigma_1$).

Our findings in Proposition 5 are particularly relevant in the era of climate change, a phenomenon leading to increased crop yield fluctuations due to rapid shifts in weather patterns (Khanal et al. 2018). As yield variability increases, Proposition 5 underscores the crucial need to prioritize investments in index accuracy, even at the expense of the budget allocated for subsidy payments.

Next, we explore how the government's budget constraint affects decision-making. Often, when a decision maker grapples with the dilemma of investing to find a better way to serve beneficiaries (e.g., enhancing index accuracy in our context) and allocating resources to directly serve beneficiaries (e.g., subsidizing farmers based on an existing index), it may seem intuitive that lower budgets should prompt less investment in the former to ensure sufficient resources are available for providing services. This logic has been supported in several other contexts, such as patient triage

[6] If there are multiple jointly optimal solutions to the optimization problem, r^* denote the lower joint optimal index accuracy and s^* the lowest corresponding jointly optimal subsidy amount.

and treatment decisions (Sun et al. 2018) and beneficiary advisory and service delivery (Arora et al. 2022). However, in the following proposition, we show that this intuition may not hold in our setting. To explicitly capture the dependence of the government's optimal decisions on the budget level, we let $s^*(B)$ and $r^*(B)$ denote the optimal subsidy amount and index accuracy under budget level B.

Theorem 3 (Effect of budget on jointly optimal subsidy and index accuracy)**.** *For any given budget levels B_1 and B_2 such that $0 < B_1 < B_2$, there exists a closed interval $I_\sigma \subseteq [0, \mu]$ such that if $\sigma \in I_\sigma$, then the low budget level B_1 leads to a higher optimal index accuracy (i.e., $r^*(B_1) \geq r^*(B_2)$) and a lower optimal subsidy (i.e., $s^*(B_1) \leq s^*(B_2)$) compared to the high budget level B_2.*

Theorem 3 uncovers an intriguing characteristic of the optimal balance between subsidy and index accuracy: Within a medium yield variability range, it is optimal for the government to invest more in enhancing index accuracy when faced with a tighter budget constraint. The rationale behind this result is as follows: with sufficient budget and moderate yield variability, the government must strike a balance between the subsidy amount and index accuracy to incentivize an adequate number of farmers to plant (so that farmers can earn revenue from selling the crop), while also avoiding oversupply (so that the market price is not too low). In this case, although both increasing the subsidy and improving the index accuracy can incentivize farmers to plant, the preferred approach is to increase the subsidy. This stems from the fact that subsidy payment is an internal transfer between the government and farmers, while improving index accuracy incurs additional costs. Conversely, when the budget is tight, oversupply is less of a concern, and thus it is essential for the government to identify the most effective approach to incentivize more farmers to plant. In this case, improving the index accuracy is crucial as it ensures that farmers receive the subsidy when most needed, thereby enhancing the subsidy's effectiveness in increasing farmers' utility.

These results underscore a key insight: a reduced budget does not automatically imply diminished investment in index accuracy. From a practical perspective, this insight holds particular relevance given the increasing budgetary constraints many governments face for crop subsidies as they also strive to address other urgent needs (Ye et al. 2021). While it may seem intuitive for governments to reduce spending on index accuracy in times of budget cuts, our findings suggest that the opposite can be true. That is, it could be optimal for governments to increase investment in index accuracy, enabling the limited subsidy budget to be deployed more effectively for farmer support.

7. Calibrated Numerical Study

In this section, we present a calibrated numerical study using real data on corn production to demonstrate that the main insights obtained from our base model remain intact under alternative contexts. In doing so, we also provide a concrete yield index example.

7.1. Data and Experiment Setup

In this section, we introduce our experiment setup and describe how we parameterize our model based on available data for corn production in Indonesia. For the purpose of illustration, we simulate $n = 10,000$ farmers and consider that each farmer has a land size of one hectare (Nazmi et al. 2021). First, we estimate the farmers' household yields. Instead of assuming all farmers' yields are binary and perfectly correlated as in our base model, we use a more realistic continuous distribution, namely, a normal distribution, to model crop yield (Alizamir et al. 2019). Moreover, we allow farmers' yields to be only partially correlated (Ramirez et al. 2003). Specifically, let y_i denote the yield of farmer i, where $y_i \sim \mathcal{N}(\mu, \sigma^2)$, $i = 1, ..., n$, and let $\rho \in [0, 1]$ denote the correlation among farmers' yields, i.e., $\rho = \text{corr}(y_i, y_j)$ for $i \neq j$. Based on the annual yield data from 2011 to 2021 for corn production in Indonesia (USDA 2021), we estimate that $\mu = 3.029$ tons per hectare, $\sigma = 0.418$ tons per hectare, and $\rho = 0.6$ (see Appendix B for details).

Next, we estimate the market-clearing price parameters a and b and the production cost parameter c. To estimate a and b, we fit a linear regression model between the annual total corn supply and the inflation-adjusted corn prices in Indonesia, which are both available in USDA's Grain and Feed Annual Reports (USDA 2021), and we obtain $a =$\$467.32 per ton and $b =$\$0.00189 per ton-squared (see more details about the data in Appendix B). For production costs, we keep the assumption of uniform distribution and consider that farmer i has a production cost $\frac{i}{n}c$. Based on available data from the USDA's website (USDA 2023), we estimate the average corn production cost to be \$875.1 per hectare. By equating $\frac{c}{2} =$ to 875.1, we obtain $c = \$1750.2$ per hectare.

Finally, to showcase the relevance of our results in practical settings, we construct an area-yield index defined as the sample average estimate of the area-level yield. As discussed earlier, such an index has been implemented by several governments, including the Indonesian government. Suppose d samples are used to construct the index, and let $k_1, ..., k_d$ denote d farmers that are randomly selected. Then, the sample average $\hat{Y}(d) = \frac{1}{d}\sum_{i=1}^{d} y_{k_i}$ is used to determine whether to trigger a payment. Specifically, we consider that a subsidy s will be paid to all farmers who plant if the value of $\hat{Y}(d)$ is lower than a threshold. For the purpose of illustration, we use μ as the threshold. Let $I = \mathbf{1}\{\hat{Y}(d) < \mu\}$ be an indicator function taking 1 if the sample average $\hat{Y}(d)$ is lower than μ and 0 otherwise. Suppose k farmers plant. Then, the utility of farmer k is

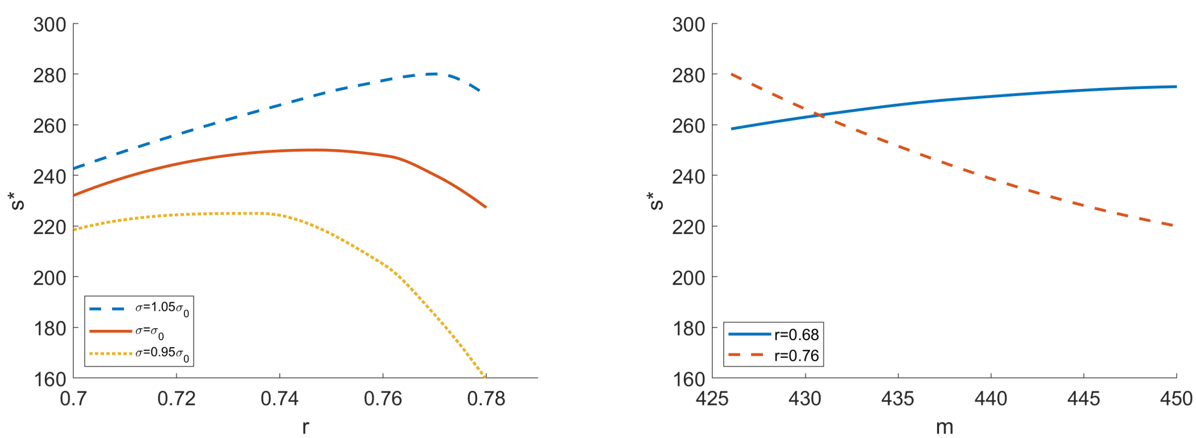
$$\mathbb{E}\left[(a - b\sum_{i=1}^{k} y_i)y_k - \frac{k}{n}c + sI\right] - \lambda\mathbb{V}\text{ar}\left[(a - b\sum_{i=1}^{k} y_i)y_k - \frac{k}{n}c + sI\right].$$

Farmer k will choose to plant corn if their utility is non-negative. The risk aversion coefficient is set to be $\lambda = 0.01$, which satisfies the assumptions stated in §3 to ensure that without subsidy, a positive number of but not all farmers choose to plant (see Appendix B for details).

7.2. Numerical Experiments and Results

Using the above setup, we conduct three groups of numerical experiments to demonstrate the robustness of our key insights stated in Propositions 3, 5, and 7, respectively. First, we explore how the optimal subsidy amount varies with index accuracy, i.e., the probability of receiving a payment conditioned on that the actual yield of a farmer is below μ. For the area-yield index described in §7.1, the index accuracy is determined by the sample size d. Consequently, in the experiments, we use different sample sizes to generate the area-yield index. Then, under each sample size, which corresponds to a unique index accuracy, we numerically search for the optimal subsidy amount. Figure 2(a) presents the optimal subsidy under varying index accuracy, showing that the optimal subsidy amount initially grows and then drops with the index accuracy, corroborating our findings in Theorem 1. In addition, a comparison of the three plots in Figure 2(a) also demonstrates the insight in Proposition 3, which states that the optimal subsidy increases in yield variability.

Figure 2 Optimal Subsidy Amount under Different Index Accuracies and Floor Prices

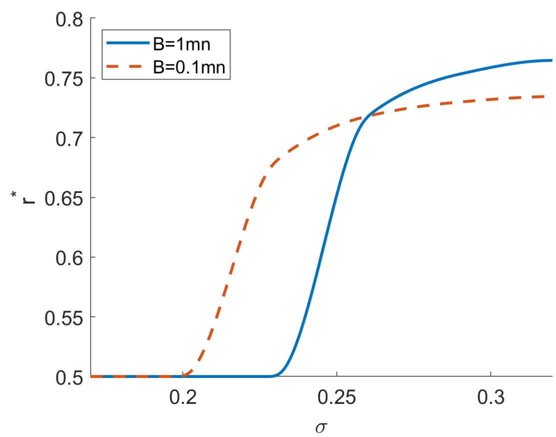


(a) Effect of Index Accuracy (b) Effect of Floor Price

Our second group of experiments explores the interplay between price protection and index-based yield protection. Here, in addition to the area-yield index policy, we also consider an MSP policy with a floor price m. That is, we replace the price $a - b\sum_{i=1}^{k} y_i$ in Equation (7.1) by $\max\left\{\left(a - b\sum_{i=1}^{k} y_i\right), m\right\}$. We consider two different values of index accuracy (i.e., sample size) r. For each fixed r, we search for the optimal subsidy amount under different floor prices m. As shown in Figure 2(b), when the index accuracy is low, the optimal subsidy under index-based yield protection increases as the floor price increases. Conversely, when the index accuracy is high, the optimal subsidy decreases in the floor price. The results resonate with the key insights from Theorem 2.

Our final set of experiments demonstrates the effect of budget when the government can exert a costly effort to improve the index accuracy. Instead of assuming a quadratic cost for improving index accuracy as in our analytical model, we now consider that the government incurs \$10 for each increase in sample size. Due to the diminishing return from a larger sample size, this corresponds to a convex cost in improving index accuracy. Under each given budget, we search for the jointly optimal subsidy and index accuracy (i.e., sample size). By comparing the optimal index accuracy under two different budget levels as shown in Figure 3, we observe that in an intermediate range of yield variability, the optimal index accuracy under the lower budget $B = 0.1$ (million \$) is strictly higher than that under the higher budget $B = 1$ (million \$), demonstrating the insights shown in Theorem 3.

Figure 3 Jointly Optimal Index Accuracy under Two Budgets



8. Extensions

In this section, we extend our model in several aspects to showcase the robustness of our insights. For each extension, we outline the model setup and summarize the main findings below. Detailed analysis and proofs are presented in Appendix C.

8.1. Conditional Value at Risk (CVaR) for Farmer Risk Aversion

In our base model, we capture farmers' risk aversion through a mean-variance utility approach. In this section, we consider an alternative risk measure, the conditional value at risk (CVaR), to model farmers' risk aversion and show that our key insights continue to hold. In particular, CVaR describes the conditional expected payoff in the worst α fraction of cases, where α measures the risk aversion level: the lower the value of α, the higher the level of risk aversion.

Recall that given the index I and the subsidy amount s, if x farmers plant, the net income of a farmer with production cost h from planting is $\pi(h|x,s) = (a - bxY)Y + sI - h$. Then, the CVaR at α level is given by

$$CVaR_\alpha(h|x,s) = \frac{1}{\alpha}\left(\mathbb{E}[\pi(h|x,s)\mathbf{1}_{\{\pi(h|x,s)\le\pi_\alpha\}}] + \pi_\alpha(\alpha - \mathbb{P}(\pi(h|x,s)\le\pi_\alpha))\right),$$

where $\pi_\alpha = \inf\{z \in \mathbb{R} : \mathbb{P}(\pi(h|x,s) \leq z) \geq \alpha\}$ denotes the lower α quantile of $\pi(h|x,s)$. Similar to before, we consider that each farmer chooses to plant if and only if $CVaR_\alpha(h|x,s)$ is positive. In practice, smallholder farmers often exhibit a high risk-averse level, and 5% or 10% CVaR has been commonly used to model farmer risk aversion in the literature (Liu et al. 2006, Pagnoncelli and Piazza 2012). Thus, to model farmers' risk aversion and for ease of exposition, we consider $\alpha \in (0, \frac{1}{4})$. Then, if $r \leq 1-2\alpha$ (which is equivalently to $\frac{1-r}{2} \geq \alpha$), the expression of $CVaR_\alpha(h|x,s)$ can be simplified as $CVaR_\alpha(h|x,s) = (a - bx(\mu-\sigma)))(\mu-\sigma) - h$. If $r > 1-2\alpha$ (which is equivalently to $\frac{1-r}{2} < \alpha$), we have $CVaR_\alpha(h|x,s) = (a - bx(\mu-\sigma)))(\mu-\sigma) - h + \frac{\left(\alpha - \frac{1-r}{2}\right)s}{\alpha}$.

In the CVaR model, the optimal subsidy amount may not be continuous in the index accuracy (due to a discrete distribution for crop yield), which complicates the analysis. Nevertheless, we are able to prove that the optimal subsidy amount remains non-monotonic in index accuracy (see Appendix C.1 for a formal characterization). This is because if the index accuracy is rather low, increasing the subsidy amount cannot increase farmers' lower tail conditional expected payoff. Thus, the optimal subsidy that maximizes the net benefit is low. When the index accuracy increases, the policy becomes more effective in increasing farmers' lower tail payoff, justifying a higher subsidy. Finally, when the index accuracy is sufficiently high, a high subsidy results in oversupply, and thus the optimal subsidy amount starts to decrease in index accuracy. With these results, we further numerically verify that our other insights also remain intact under the CVaR model.

8.2. Consumer Surplus

In addition to improving farmer welfare, agricultural subsidies can also be used to improve consumer surplus by ensuring a sufficient supply of crops. In this section, we extend our model to consider an alternative government objective that incorporates consumer surplus and show that our key insights remain intact. Following the existing literature on agricultural supply chains and subsidies (Alizamir et al. 2019, Chintapalli and Tang 2021), we define the consumer surplus by integrating the utility of consumers who purchase the product. Specifically, given the market price $a - bx^*(s)Y$, all consumers whose willingness to pay is higher than the market price will purchase the product. Then, the total utility of these consumers can be obtained by $\int_0^{x^*(s)} \mathbb{E}[((a-bxY) - (a - bx^*(s)Y))Y]\,\mathrm{d}x$. Let $v_c(s)$ denote the alternative objective function, comprising the total farmer surplus and the total consumer surplus, minus the government expenditure. That is,

$$v_c(s) := \int_0^{x^*(s)} \mathbb{E}[(a - bx^*(s)Y)Y + sI - xc]dx + \int_0^{x^*(s)} \mathbb{E}[((a-bxY) - (a - bx^*(s)Y)Y]\,\mathrm{d}x$$
$$- x^*(s)\mathbb{E}[sI].$$

By following a similar analysis as in §4-6, we are able to analytically characterize the optimal design of index-based yield protection policy and show that our key insights continue to hold (see

Appendix C.2 for details). Moreover, we find that when taking into account consumer surplus, the government should offer greater subsidies to farmers than in the base model, especially when the index accuracy is high. The reason is that, when the index accuracy is high, an increase in the subsidy amount can effectively incentivize more farmers to plant (as shown in Proposition 1), resulting in more supply and thus benefiting consumers. Therefore, in this case, the optimal subsidy should increase as compared to the base model. When the index accuracy is lower, however, an increase in subsidy is less effective in incentivizing farmers to plant, resulting in less benefit for consumers. Therefore, in this case, the benefit of increasing the subsidy is lower.

8.3. Different Distributions for Yield and Index

In our base model, we consider that the yield level Y and the index I have the same marginal distribution, which allows us to simplify the exposition and highlight our key insights more clearly. In this section, we extend our model to consider different distributions for Y and I and demonstrate that our key insights continue to hold. Specifically, while the probability for the actual yield Y to be low is $\frac{1}{2}$, we now consider that the government can design an index such that the probability of $I = 1$ (i.e., low-yield prediction) can be any number $p \in (0, 1]$. That is, p denotes the subsidy payment probability. In practice, this probability can be adjusted by appropriately choosing the index. For instance, for a rainfall index that triggers payments if the rainfall level is below a certain threshold, the government can increase the payment probability by setting a higher threshold.

Given the payment probability p, we consider a joint distribution of Y and I as shown in Table 2. In line with our base model, the parameter r represents the probability that the actual yield is low conditioned on $I = 1$. When $p = \frac{1}{2}$, the joint distribution becomes the same as in our base model. To ensure that all probabilities are nonnegative, we assume that $rp \le \frac{1}{2}$.

Table 2 New Joint Probability Table for Index-based Yield Protection

	$I = 0$	$I = 1$
$Y = \mu + \sigma$	$\frac{1}{2} - (1 - r)p$	$(1 - r)p$
$Y = \mu - \sigma$	$\frac{1}{2} - rp$	rp

With this more general joint distribution of Y and I, our model remains analytically tractable, and we show that our key insights hold true for any value of p (see Appendix C.3). Moreover, we compare the optimal subsidy amount under different payment probability p. In particular, as the probability of yield being low is $\frac{1}{2}$, we focus on the higher range of payment probability $p \in [\frac{1}{2}, 1]$. Through extensive numerical experiments, we discover an additional insight: When both the index accuracy r and the payment probability p are low, an increase in the the payment probability leads to a greater optimal subsidy amount; when either the index accuracy or the payment probability is high, a greater payment probability results in a lower optimal subsidy amount.

This result provides a managerial takeaway. When the government faces a low index accuracy (i.e., low r) as well as a low payment probability (i.e., low p), the chance for the farmers to receive the subsidy at low yield is low and thus the subsidy is not very effective in benefiting the farmers. Therefore, in this case, the government should consider moderately increasing the payment probability to boost the value of the subsidy payment. However, if the index accuracy or the payment probability is high, a high subsidy amount in addition to the high probability of receiving the subsidy would lead to oversupply. Thus, in this case, the government should offer a lower subsidy amount.

8.4. Effect of Premium

In our base model, we consider scenarios where farmers can access yield protection subsidies from the government at no cost. Examples of such scenarios include several weather and satellite index policies in Kenya and Sri Lanka and an area-yield index policy in Peru, which are fully funded by the government (Hazell et al. 2017, Aheeyar et al. 2021). On the other hand, there are also yield-protection policies where farmers must pay a premium to access the benefits. For instance, to enroll in an area-yield index policy in Indonesia, farmers must pay a premium to cover about 20% of the policy expenses, while the government covers the remaining 80% (JICA 2022). To capture such scenarios, we extend our model to consider that, to enroll in the index-based yield protection policy, farmers are required to pay a premium. This premium is set such that the total premium paid by the farmers equals a fraction $\xi \in [0,1]$ of the expected total payment they receive. This is equivalent to each farmer paying a premium of $\xi\mathbb{E}[sI]$, where we continue to let s denote the payment a farmer receives when the index predicts a low-yield situation (i.e., when $I=1$). Then, the mean-variance utility of a farmer with the production cost h is given by

$$\mathbb{E}[(a-bxY)Y+sI-h-\xi\mathbb{E}[sI]]-\lambda\mathbb{V}\mathrm{ar}[(a-bxY)Y+sI-h-\xi\mathbb{E}[sI]].$$

Under this setup, our model remains analytically tractable, and we show that our key insights continue to hold for any value of ξ (see Appendix C.4). Moreover, through extensive numerical analysis, we find that when the index accuracy is low, a higher value of ξ (i.e., farmers paying a higher fraction of the expenses) leads to a lower optimal payment s; however, when the index accuracy is high, a higher value of ξ leads to a higher optimal payment. The intuition for this result can be explained as follows: Recall that if the index accuracy is low, an increase in the payment s can increase farmers' income variance. When ξ increases, this concern becomes even more prominent because farmers' lowest possible income will be even lower (due to the premium payment). As a result, the policy becomes even less effective, leading to a lower optimal payment. Conversely, if the index accuracy is high, a high payment s can potentially lead to oversupply, while an increase in ξ helps mitigate such concerns, making a higher payment optimal.

9. Conclusion

Farmers, particularly those in emerging economies, are vulnerable to risks associated with variable weather conditions and uncertain crop yields. However, protecting smallholder farmers from low crop yields is generally challenging due to the high cost of yield assessment. In this paper, we study a recently emerged index-based approach for yield protection, under which the government subsidizes farmers when a pre-determined index indicates that the crop yield is low.

In order to shed light on the optimal design of an index-based yield protection policy, we build an analytical model that captures three salient features of our problem: (1) uncertain crop yield; (2) planting decisions of heterogeneous and risk-averse farmers; and (3) an index that is imperfectly correlated to the crop yield. Our model, while incorporating these essential features, retains tractability, enabling us to analytically study the government's optimal subsidy design.

Our analysis reveals several insights with practical implications. First, by characterizing how the subsidy payment affects individual farmers' income variance and their incentives to plant under different index accuracies, we uncover a non-monotonic relationship between the optimal subsidy and the index accuracy. This result provides guidance for governments to appropriately choose the subsidy amounts under different index accuracies; it also implies that the prevalent approach for determining payment to farmers, which often ignores the index accuracy, is suboptimal.

Second, we identify the critical role of index accuracy in the interplay between price protection and index-based yield protection policies. Specifically, since both price and yield protection can incentivize more farmers to plant, it might seem intuitive that these two policies would function as strategic substitutes. However, we show that this is only true if the index accuracy is sufficiently high. Otherwise, when index accuracy is low, these two policies work as strategic complements. This finding suggests that if a government is constrained to employ a low-accuracy index, the existence of price protection can make index-based yield protection more valuable.

Lastly, we examine a scenario where index accuracy can be enhanced through a costly effort, and the government allocates a limited budget between increasing the index accuracy and subsidizing farmers. We underscore the importance of index accuracy under high yield variability and when the government faces a stringent budget constraint on total expenditure. Notably, we demonstrate that under a more constrained budget, it can be optimal for the government to spend more in improving the index accuracy, even if it necessitates a reduction in subsidies to farmers.

Collectively, we believe these insights contribute to a more nuanced understanding of index-based yield protection policies and provide practical guidance for governments in designing agricultural subsidies. Moreover, through a combination of analytical and numerical studies, we further validate the robustness of our insights under various practical settings, including a more realistic continuous distribution for yield and partially correlated yield among farmers, consideration of consumer surplus in the government's objective, an alternative model for farmer risk aversion, different marginal distributions of the crop yield and the index, and premiums paid by farmers.

As an early work that studies the optimal design of index-based yield protection policies, we conclude by discussing a few directions for future research. While this paper primarily focuses on characterizing the role of index accuracy in determining the optimal subsidy design and the allocation of budget between index accuracy and subsidy payment, it would be interesting to study how to optimally construct an index from existing data and then map the index to the subsidy payment. Furthermore, in addition to government-launched yield protection policies, insurance companies have started to provide commercial index-based yield insurance products to farmers, but with low take-up rates (Annan and Datta 2022). Therefore, it would also be interesting to study how to design such commercial products to increase the take-up rates while maintaining the profitability of insurance companies, and how agri-business firms, such as food processing companies, can leverage such products to protect their suppliers and manage their own risks.

References

Adebola, O., P. Arora, C. Zhang. 2022. Farm equipment sharing in emerging economies. *Working Paper, Available at SSRN 4190725.*

Agrawal, V., G. Esenduran, Ş. Yücel. 2022. Delivery terms for voluntary carbon offsets. *Working Paper, Available at SSRN.*

Aheeyar, M., K. Samarakoon, S. de Silva. 2021. *Bundled weather index insurance pilot for drought-affected areas in Sri Lanka: Reaching marginal farmers*, vol. 201. International Water Management Institute (IWMI).

Akkaya, D., K. Bimpikis, H. Lee. 2021. Government interventions to promote agricultural innovation. *Manufacturing & Service Operations Management* **23**(2) 437–452.

Alizamir, S., F. Iravani, B. Kalkanci. 2022. Tip your farmer? Implications of tipping in agriculture on sustainability and financial inclusion. *Working Paper, Available at SSRN 3780216.*

Alizamir, S., F. Iravani, H. Mamani. 2019. An analysis of price vs. revenue protection: Government subsidies in the agriculture industry. *Management Science* **65**(1) 32–49.

Annan, F., B. Datta. 2022. Risk, informal institutions, and index insurance. *Journal of Environmental Economics and Management* **113** 102635.

Antonaci, L., M. Demeke, A. Vezzani. 2014. The challenges of managing agricultural price and production risks in sub-Saharan Africa.

Arifoğlu, K., C.S. Tang. 2022. A two-sided incentive program for coordinating the influenza vaccine supply chain. *Manufacturing & Service Operations Management* **24**(1) 235–255.

Arora, P., M. Rahmani, K. Ramachandran. 2022. Doing less to do more? Optimal service portfolio of non-profits that serve distressed individuals. *Manufacturing & Service Operations Management* **24**(2) 883–901.

Babich, V., R. Lobel, Ş. Yücel. 2020. Promoting solar panel investments: Feed-in-tariff vs. tax-rebate policies. *Manufacturing & Service Operations Management* **22**(6) 1148–1164.

Benami, E., Z. Jin, M.R. Carter, A. Ghosh, R.J. Hijmans, A. Hobbs, B. Kenduiywo, D.B. Lobell. 2021. Uniting remote sensing, crop modelling and economics for agricultural risk management. *Nature Reviews Earth & Environment* **2**(2) 140–159.

Boyabatlı, O., J. Nasiry, Y. Zhou. 2019. Crop planning in sustainable agriculture: Dynamic farmland allocation in the presence of crop rotation benefits. *Management Science* **65**(5) 2060–2076.

Boyabatlı, O., J. Nguyen, T. Wang. 2017. Capacity management in agricultural commodity processing and application in the palm industry. *Manufacturing & Service Operations Management* **19**(4) 551–567.

Cai, J., A. De Janvry, E. Sadoulet. 2020. Subsidy policies and insurance demand. *American Economic Review* **110**(8) 2422–2453.

Carter, M., A. de Janvry, E. Sadoulet, A. Sarris. 2017. Index insurance for developing country agriculture: A reassessment. *Annual Review of Resource Economics* **9** 421–438.

CCAFS. 2013. Improved rainfall thresholds for index insurance in India. *CCAFS Report*.

Chavas, J.P., R.G. Chambers, R.D. Pope. 2010. Production economics and farm management: a century of contributions. *American Journal of Agricultural Economics* **92**(2) 356–375.

Chen, Z., Y. Lu, J. Zhang, W. Zhu. 2023. Managing weather risk with a neural network-based index insurance. *Management Science*.

Chintapalli, P., C.S. Tang. 2021. The value and cost of crop minimum support price: Farmer and consumer welfare and implementation cost. *Management Science* **67**(11) 6839–6861.

Clarke, D.J. 2016. A theory of rational demand for index insurance. *American Economic Journal: Microeconomics* **8**(1) 283–306.

Cohen, M.C., R. Lobel, G. Perakis. 2016. The impact of demand uncertainty on consumer subsidies for green technology adoption. *Management Science* **62**(5) 1235–1258.

de Zegher, J.F., D.A. Iancu, H.L. Lee. 2019. Designing contracts and sourcing channels to create shared value. *Manufacturing & Service Operations Management* **21**(2) 271–289.

Fan, S., C. Rue. 2020. The role of smallholder farms in a changing world. *The role of smallholder farms in food and nutrition security*. Springer, Cham, 13–28.

Fan, T., Q. Feng, Y. Li, J.G. Shanthikumar, Y. Wu. 2023. Output-oriented agricultural subsidy design. *Management Science*.

Federgruen, A., U. Lall, A.S. Şimşek. 2019. Supply chain analysis of contract farming. *Manufacturing & Service Operations Management* **21**(2) 361–378.

Ferreira, K., J. Goh, E. Valavi. 2017. Intermediation in the supply of agricultural products in developing economies. *Harvard Business School Research Paper Series*.

Flatnes, J.E., M.R. Carter. 2016. Fail-safe index insurance without the cost: A satellite based conditional audit approach.

Freudenreich, H., O. Mußhoff. 2018. Insurance for technology adoption: An experimental evaluation of schemes and subsidies with maize farmers in Mexico. *Journal of Agricultural Economics* **69**(1) 96–120.

Ghosh, R.K., S. Gupta, V. Singh, P.S. Ward. 2021. Demand for crop insurance in developing countries: New evidence from India. *Journal of agricultural economics* **72**(1) 293–320.

Greatrex, H., J. Hansen, S. Garvin, R. Diro, M. Le Guen, S. Blakeley, K. Rao, D. Osgood. 2015. Scaling up index insurance for smallholder farmers: Recent evidence and insights. *CCAFS Report*.

Guda, H., M. Dawande, G. Janakiraman, T. Rajapakshe. 2021. An economic analysis of agricultural support prices in developing economies. *Production and Operations Management* **30**(9) 3036–3053.

Harvey, C.A., M. Saborio-Rodríguez, M.R. Martinez-Rodríguez, B. Viguera, A. Chain-Guadarrama, R. Vignola, F. Alpizar. 2018. Climate change impacts and adaptation among smallholder farmers in Central America. *Agriculture & Food Security* **7**(1) 1–20.

Hazell, P., R. Sberro-Kessler, P. Varangis. 2017. *When and how should agricultural insurance be subsidized? Issues and good practices*. World Bank.

Hemming, D.J., E.W. Chirwa, A. Dorward, H.J. Ruffhead, R. Hill, J. Osborn, L. Langer, L. Harman, H. Asaoka, C. Coffey, et al. 2018. Agricultural input subsidies for improving productivity, farm income, consumer welfare and wider growth in low-and lower-middle-income countries: A systematic review. *Campbell Systematic Reviews* **14**(1) 1–153.

Herger, M. 2020. What is the role of agriculture for sustainable development? Lessons from India.

Hu, M., Y. Liu, W. Wang. 2019. Socially beneficial rationality: The value of strategic farmers, social entrepreneurs, and for-profit firms in crop planting decisions. *Management Science* **65**(8) 3654–3672.

JICA. 2022. Launching ceremony for AYII (area yield index insurance) pilot project for paddy in Karawang, West Java, https://www.jica.go.jp/Resource/indonesia/english/office/others/press220309.html, last accessed on August 7, 2023.

Kazaz, B. 2004. Production planning under yield and demand uncertainty with yield-dependent cost and price. *Manufacturing & Service Operations Management* **6**(3) 209–224.

Kenduiywo, B.K., M.R. Carter, A. Ghosh, R.J. Hijmans. 2021. Evaluating the quality of remote sensing products for agricultural index insurance. *PloS One* **16**(10) e0258215.

Khanal, U., C. Wilson, V. Hoang, B. Lee. 2018. Farmers' adaptation to climate change, its determinants and impacts on rice yield in Nepal. *Ecological Economics* **144** 139–147.

Levi, R., G. Perakis, G. Romero. 2019. Near-optimality of uniform copayments for subsidies and taxes allocation problems. *Operations Research* **67**(2) 548–561.

Levi, R., M. Rajan, S. Singhvi, Y. Zheng. 2020. Improving farmers' income on online agri-platforms: Theory and field implementation of a two-stage auction. *Available at SSRN 3486623*.

Liao, C., Y. Chen, C.S. Tang. 2019. Information provision policies for improving farmer welfare in developing countries: Heterogeneous farmers and market selection. *Manufacturing & Service Operations Management* **21**(2) 254–270.

Liu, J., C. Men, V.E. Cabrera, S.P. Uryasev, C. Fraisse. 2006. Cvar model for optimizing crop insurance under climate variability. *Available at SSRN 878214*.

Lowder, S.K., M.V. Sánchez, R. Bertini. 2021. Which farms feed the world and has farmland become more concentrated? *World Development* **142** 105455.

Mgbenka, R.N., E.N. Mbah, C.I. Ezeano. 2016. A review of smallholder farming in Nigeria: Need for transformation. *International Journal of Agricultural Extension and Rural Development Studies* **3**(2) 43–54.

Nazmi, M.S., R.A. Dardak, R.A. Rani, M.R. Rabu. 2021. Benchmarking Indonesia for the development of grain corn industry in Malaysia. *Food and Fertilizer Technology Center for the Asian and Pacific Region Platform (FFTC-AP Platform)*.

Pagnoncelli, B.K., A. Piazza. 2012. The optimal harvesting problem under risk aversion. *Available at www. optimization-online. org*.

Pay, K., S. Singhvi, Y. Zheng. 2022. Improving cash-constrained smallholder farmers' revenue: The role of government loans. *Working Ppaper, Available at SSRN 4135868*.

Raithatha, R., J. Priebe. 2020. Agricultural insurance for smallholder farmers - Digital innovations for scale.

Ramaswami, B, S Seshadri, KV Subramanian. 2018. The welfare economics of storage-based price supports. *Working Paper, Indian Statistical Institute, Delhi, India* .

Ramirez, O.A., S. Misra, J. Field. 2003. Crop-yield distributions revisited. *American Journal of Agricultural Economics* **85**(1) 108–120.

Shi, Y., J. de Zegher, I. Lo. 2023. Two-sided benefits of price transparency in smallholder supply chains. *Working Paper, Available at SSRN 4052928*.

Sun, Z., N.T. Argon, S. Ziya. 2018. Patient triage and prioritization under austere conditions. *Management Science* **64**(10) 4471–4489.

Takahashi, K., M. Ikegami, M. Sheahan, C.B. Barrett. 2016. Experimental evidence on the drivers of index-based livestock insurance demand in Southern Ethiopia. *World Development* **78** 324–340.

Tang, C.S., Y. Wang, M. Zhao. 2023. The impact of input and output farm subsidies on farmer welfare, income disparity, and consumer surplus. *Management Science*.

Taylor, S. 2021. "It's difficult to feed our families": Volatile cocoa prices are pushing West African farmers further into poverty. *CNBC News*.

Taylor, T.A., W. Xiao. 2014. Subsidizing the distribution channel: Donor funding to improve the availability of malaria drugs. *Management Science* **60**(10) 2461–2477.

USDA. 2021. Grain and feed annual. https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Grain%20and%20Feed%20Annual_Jakarta_Indonesia_03-27-2021, last accessed on July 30, 2023.

USDA, Economic Research Service. 2023. Recent cost and returns - Corn. https://www.ers.usda.gov/data-products/commodity-costs-and-returns/commodity-costs-and-returns/#Historical%20Costs%20and%20Returns:%20Hogs,%20all, last accessed on August 06, 2023.

Wang, H.H., S.D. Hanson, R.J. Myers, J.R. Black. 1998. The effects of crop yield insurance designs on farmer participation and welfare. *American Journal of Agricultural Economics* **80**(4) 806–820.

Ye, F., Z. Cai, Y. Chen, Y. Li, G. Hou. 2021. Subsidize farmers or bioenergy producer? The design of a government subsidy program for a bioenergy supply chain. *Naval Research Logistics (NRL)* **68**(8) 1082–1097.

Zhang, Y., J.M. Swaminathan. 2020. Improved crop productivity through optimized planting schedules. *Manufacturing & Service Operations Management* **22**(6) 1165–1180.

Zhou, J., X. Fan, Y. Chen, C.S. Tang. 2021. Information provision and farmer welfare in developing economies. *Manufacturing & Service Operations Management* **23**(1) 230–245.

Online Supplement for "Index-Based Yield Protection for Smallholder Farmers"

Appendix A Proofs of Analytical Results

Appendix A.1 Proofs of Analytical Results in §4

Proof of Lemma 1: According to the formulation of farmers' utility $u(h|x)$ defined in Equation (1), for any given planting amount x, the production cost h is the only parameter that differentiates the utility of the farmers who plant, where a higher h leads to a lower utility. Therefore, in equilibrium, it must be farmers with lower h who choose to plant, and those with higher h who choose to not plant. Then, if $x = 0$ is an equilibrium planting amount, we must have $u(0|0) \leq 0$ (i.e., the utility of the lowest cost farmer is non-positive when no one plants); if $x = 1$ is an equilibrium planting amount, we must have $u(c|1) \geq 0$ (i.e., the utility of the highest cost farmer is non-negative when everyone plants); if $x \in (0,1)$ is an equilibrium planting amount, we must have $u(xc|x) = 0$ (i.e., the farmer whose planting cost is xc is indifferent between planting and not planting). We derive the equilibrium planting amount and its closed-form expression in the following two steps:

Step 1: We show that there is a unique positive solution to $u(xc|x) = 0$, which lies in $(0,1)$. To do so, it is sufficient to prove the following three parts: (i) $u(xc|x)$ strictly decreases in x for $x \geq 0$, (ii) $u(0|0) > 0$, and (iii) $u(c|1) < 0$.

First, for part (i), recall that Y takes values $\mu + \sigma$ and $\mu - \sigma$ with probability $\frac{1}{2}$ respectively. Besides, recall that $R_H(x) = [a - bx(\mu + \sigma)](\mu + \sigma)$ (revenue from selling the crop with high yield) and $R_L(x) = [a - bx(\mu - \sigma)](\mu - \sigma)$ (revenue from selling the crop with low yield). Then, we have

$$\mathbb{E}[\pi(xc|x)] = \frac{1}{2}R_H(x) + \frac{1}{2}R_L(x) - xc = a\mu - bx(\mu^2 + \sigma^2) - xc$$

$$\mathbb{V}\text{ar}[\pi(xc|x)] = \frac{1}{4}(R_H(x) - R_L(x))^2 = (a\sigma - 2bx\mu\sigma)^2. \tag{A.1}$$

Plugging these into the expression of $u(xc|x)$ from Equation (1), we have

$$\begin{aligned} u(xc|x) &= \mathbb{E}[\pi(xc|x)] - \lambda\mathbb{V}\text{ar}[\pi(xc|x)] \\ &= a\mu - bx(\mu^2 + \sigma^2) - xc - \lambda(a\sigma - 2bx\mu\sigma)^2 \\ &= -\alpha_0 x^2 - \beta_0 x - \gamma_0, \end{aligned} \tag{A.2}$$

where $\alpha_0 = 4\lambda b^2\mu^2\sigma^2$, $\beta_0 = b(\mu^2 + \sigma^2) + c - 4\lambda ab\mu\sigma^2$, and $\gamma_0 = \lambda a^2\sigma^2 - a\mu$.

Clearly, we have $\alpha_0 > 0$. Thus, from Equation (A.2), $u(xc|x)$ is a quadratic concave function in x. Moreover, since $c > c_1 \geq 4\lambda ab\sigma^2\mu - b(\mu^2 + \sigma^2)$ as discussed in §3, we have $\beta_0 > 0$. Hence, $u(xc|x)$ strictly decreases in x for $x \geq 0$.

Second, for part (ii), since $\lambda < \frac{\mu+\sigma}{2a\sigma^2} \leq \frac{\mu}{a\sigma^2}$ as discussed in §3, we have $u(0|0) = -\gamma_0 > 0$.

Finally, for part (iii), since $c > c_1 \geq a\mu - b(\mu^2+\sigma^2) - \lambda(a\sigma - 2b\mu\sigma)^2$ as discussed in §3, we have $u(c|1) < 0$.

Collectively, we conclude that there must exist a unique positive solution to $u(xc|x) = 0$ in $(0,1)$, which we denote as $\hat{x}_0$. From Equation (A.2), we have[7]

$$\hat{x}_0 = \frac{-\beta_0 + \sqrt{\beta_0^2 - 4\alpha_0\gamma_0}}{2\alpha_0} \in (0,1).$$

Step 2: We prove that there is a unique equilibrium planting amount $x_0^* = \hat{x}_0 \in (0,1)$, where $\hat{x}_0$ is the unique positive solution to $u(xc|x) = 0$. Since we have shown in Step 1 that $u(0|0) > 0$ and $u(c|1) < 0$, based on the discussion at the beginning of this proof, both $x = 0$ and $x = 1$ are not equilibrium planting amount. Moreover, we have shown in Step 1 that $\hat{x}_0 \in (0,1)$ is the unique positive solution to $u(xc|x) = 0$. Therefore, it remains to prove that $\hat{x}_0$ is an equilibrium planting amount.

When the planting amount is given by $x = \hat{x}_0$, no farmer whose production cost is lower than $\hat{x}_0 c$ would deviate from planting because $u(\hat{x}_0 c|\hat{x}_0) = 0$ and thus they are having positive utility from planting; no farmer whose production cost is higher than or equal to $\hat{x}_0 c$ would deviate from not planting because if they choose to plant instead, then, as we have that $u(xc|x)$ strictly decreases in x, they would have negative utility. Hence, $\hat{x}_0$ is an equilibrium planting amount, and thus $x_0^* = \hat{x}_0$ is a unique equilibrium planting amount.

Combining Step 1 and Step 2, we have $x_0^* = \hat{x}_0 = \frac{-\beta_0 + \sqrt{\beta_0^2 - 4\alpha_0\gamma_0}}{2\alpha_0} \in (0,1)$. □

Proof of Lemma 2: Following the same logic as in the proof of Lemma 1, given $s \in [0, \bar{s}]$, if $x = 0$, $x = 1$, or $x \in (0,1)$ is an equilibrium planting amount, we must have $u(0|0,s) \leq 0$, $u(c|1,s) \geq 0$ or $u(xc|x,s) = 0$ respectively. We derive the equilibrium planting amount and its closed-form expression in the following two steps:

Step 1: We show that, given $s \in [0, \bar{s}]$, there is a unique positive solution to $u(xc|x,s) = 0$. To do so, it is sufficient to prove the following three parts for any given $s \in [0, 2a\sigma]$ (we prove this step for any $s \in [0, 2a\sigma]$, and at the end of this step, we show $\bar{s} \leq 2a\sigma$): (i) $u(xc|x,s)$ strictly decreases in x for $x \geq 0$, (ii) $u(0|0,s) > 0$ and (iii) $u(xc|x,s) \leq 0$ for some $x > 0$.

First, for part (i), recall that $R_H(x) = [a - bx(\mu+\sigma)](\mu+\sigma)$ (revenue from selling the crop with high yield) and $R_L(x) = [a - bx(\mu-\sigma)](\mu-\sigma)$ (revenue from selling the crop with low yield). Then, based on the expression of $\pi(h|x,s)$ in §3, we have

$$\mathbb{E}[\pi(xc|x,s)] = \frac{1}{2}R_H(x) + \frac{1}{2}R_L(x) - xc + \frac{1}{2}s = a\mu - bx(\mu^2+\sigma^2) - xc + \frac{1}{2}s$$

[7] The other solution is negative and cannot be $\hat{x}_0$.

$$\begin{aligned}\mathbb{V}\text{ar}[\pi(h|x,s)] &= \mathbb{V}\text{ar}[(a-bxY)Y] + \mathbb{V}\text{ar}[sI] + 2\text{Cov}((a-bxY)Y, sI) \\ &= \frac{1}{4}(R_H(x)-R_L(x))^2 + \frac{1}{4}s^2 - s\left(r-\frac{1}{2}\right)(R_H(x)-R_L(x)) \qquad \text{(A.3)} \\ &= (a\sigma - 2bx\mu\sigma)^2 + \frac{1}{4}s^2 - s\left(r-\frac{1}{2}\right)(2a\sigma - 4bx\mu\sigma).\end{aligned}$$

Plugging these into the expression of $u(xc|x,s)$ in §3, we have

$$\begin{aligned}u(xc|x,s) &= \mathbb{E}[\pi(xc|x,s)] - \lambda\mathbb{V}\text{ar}[\pi(xc|x,s)] \\ &= a\mu - bx(\mu^2+\sigma^2) + \frac{1}{2}s - xc - \lambda\left[(a\sigma-2bx\mu\sigma)^2 + \frac{1}{4}s^2 - s\left(r-\frac{1}{2}\right)(2a\sigma-4bx\mu\sigma)\right] \\ &= -\alpha x^2 - \beta(s)x - \gamma(s), \qquad \text{(A.4)}\end{aligned}$$

where $\alpha = 4\lambda b^2\mu^2\sigma^2$, $\beta(s) = b(\mu^2+\sigma^2) + c - 4\lambda ab\mu\sigma^2 + 4\lambda b(r-\frac{1}{2})\mu\sigma s$, and $\gamma(s) = \lambda a^2\sigma^2 - a\mu - 2\lambda a(r-\frac{1}{2})\sigma s - \frac{1}{2}s + \frac{1}{4}\lambda s^2$.

Clearly, we have $\alpha = \alpha_0 > 0$. Thus, from Equation (A.4), $u(xc|x,s)$ is a quadratic concave function in x. Moreover, we have $\beta(s) = \beta_0 + 4\lambda b(r-\frac{1}{2})\mu\sigma s > 0$, where the inequality holds because we assumed $r \geq \frac{1}{2}$ in §3 and we showed $\beta_0 > 0$ in the proof of Lemma 1. Therefore, we must have $u(xc|x,s)$ strictly decreases in x for $x \geq 0$ for any given $s \in [0, 2a\sigma]$.

Second, for part (ii), we have $u(0|0,s) = -\gamma(s)$ from Equation (A.4). Thus, for any given $s \in [0, 2a\sigma]$, in order to prove that $u(0|0,s) > 0$, it is equivalent to prove that $\gamma(s) < 0$. Further, since we have that $\gamma(s)$ is a quadratic convex function of s and we also have that $\gamma(0) = \gamma_0 < 0$ from the proof of Lemma 1, in order to prove that $u(0|0,s) > 0$ for any given $s \in [0, 2a\sigma]$, it remains to prove that $\gamma(2a\sigma) < 0$. From the expression of $\gamma(s)$, We have

$$\begin{aligned}\gamma(2a\sigma) &= \lambda a^2\sigma^2 - a\mu - 2\lambda a(r-\frac{1}{2})\sigma(2a\sigma) - \frac{1}{2}(2a\sigma) + \frac{1}{4}\lambda(2a\sigma)^2 \\ &\leq 2\lambda a^2\sigma^2 - a\mu - a\sigma \\ &< 0,\end{aligned}$$

where the first inequality holds because we have $r \geq \frac{1}{2}$, and the second inequality holds because we have $\lambda < \frac{\mu+\sigma}{2a\sigma^2}$. Therefore, for any given $s \in [0, 2a\sigma]$, we have $u(0|0,s) > 0$.

Finally, for part (iii), from Equation (A.4), we have that $u(xc|x,s)$ is a quadratic concave function. Thus, there must exist an $x > 0$ such that $u(xc|x,s) \leq 0$.

Collectively, we conclude that given $s \in [0, 2a\sigma]$, there must exist a unique positive solution to $u(xc|x,s) = 0$, which we denote as $\hat{x}(s)$. From Equation (A.4), we have[8]

$$\hat{x}(s) = \frac{-\beta(s) + \sqrt{\beta^2(s) - 4\alpha\gamma(s)}}{2\alpha}. \qquad \text{(A.5)}$$

[8] The other solution is negative and cannot be $\hat{x}(s)$.

Before we proceed to Step 2, we prove that $\bar{s} \le 2a\sigma$. In §3, $\bar{s}$ is defined as the lowest subsidy amount that satisfies $R_L(x^*(s)) + s \ge R_H(x^*(s))$. Thus, for any $s \in [0, \bar{s}]$, we have $R_L(x^*(s)) + s \le R_H(x^*(s))$, hence

$$s \le R_H(x^*(s)) - R_L(x^*(s)) = 2a\sigma - 4b\mu\sigma x^*(s) \le 2a\sigma,$$

where the last inequality holds because $x^*(s) \in [0,1]$. Hence, we have $\bar{s} \le 2a\sigma$.

Step 2: We show that given $s \in [0, \bar{s}]$, there is a unique equilibrium planting amount $x^*(s) = \min\{\hat{x}(s), 1\} \in (0,1]$, where $\hat{x}(s)$ is the unique positive solution to $u(xc|x,s) = 0$. We have shown in Step 1 that $u(0|0,s) > 0$ for any given $s \in [0, \bar{s}]$, so, based on the discussion at the beginning of this proof, $x = 0$ cannot be equilibrium planting amount. Then, it remains to check $x = 1$ and $x = \hat{x}(s)$ if $\hat{x}(s) \in (0,1)$.

Consider any $s \in [0, \bar{s}]$. If $\hat{x}(s) \ge 1$, then, as we have shown in Step 1 that $u(xc|x,s)$ strictly decreases in x, we have $u(c|1,s) \ge 0$. In this case, we have the equilibrium planting amount $x^*(s) = 1$ because all the farmers would have positive utility from planting and would not deviate. On the other hand, if $\hat{x}(s) \in (0,1)$, then, by the same logic as in Step 2 of the proof of Lemma 1, we have the equilibrium planting amount $x^*(s) = \hat{x}(s)$. Collectively, for any $s \in [0, \bar{s}]$, there is a unique equilibrium planting amount given by $x^*(s) = \min\{\hat{x}(s), 1\}$.

Combining Step 1 and Step 2, for any given $s \in [0, \bar{s}]$, we have $x^*(s) = \min\{\hat{x}(s), 1\} = \min\left\{\frac{-\beta(s)+\sqrt{\beta^2(s)-4\alpha\gamma(s)}}{2\alpha}, 1\right\} \in (0,1]$.

Finally, we prove that there exists a threshold c_2 such that if $c \ge c_2$, then the equilibrium planting amount under given subsidy amount $s \in [0, \bar{s}]$ is $x^*(s) = \hat{x}(s) \in (0,1)$. To do so, based on the discussion in Step 2, it is sufficient to prove that there exists a threshold c_2 such that if $c \ge c_2$, then $u(c|1,s) < 1$ for any $s \in [0, \bar{s}]$. Consider a fixed $s \in [0, \bar{s}]$. We have

$$u(c|1,s) = a\mu - b(\mu^2 + \sigma^2) + \frac{1}{2}s - c - \lambda\left[(a\sigma - 2b\mu\sigma)^2 + \frac{1}{4}s^2 - s\left(r - \frac{1}{2}\right)(2a\sigma - 4b\mu\sigma)\right].$$

From its expression, we have that $u(c|1,s)$ linearly decreases in c and that $u(c|1,s)$ is continuous in s. Moreover, we have shown in Step 1 that $\bar{s} \le 2a\sigma$. Therefore, there must exist a threshold c_2 such that if $c \ge c_2$, then $u(c|1,s) < 1$ for any $s \in [0, \bar{s}]$. □

Before we move on to the proof of Proposition 1, we prove two auxiliary lemmas which serve as important building blocks for proving the rest of our results. Lemma A.1 introduces a connection between the first derivative of the equilibrium planting amount $x^*(s)$ and the first derivative of the farmer's utility $u(xc|x,s)$ at $x = x^*(s)$, with respect to any parameter in our model. Lemma A.2 characterizes the concavity of the equilibrium planting amount $x^*(s)$ in the subsidy s.

Lemma A.1. *Let η be one of the parameters in the model (η can represent s, r, σ, λ etc.). For any $s \in [0, \bar{s}]$, if $x^*(s) = \hat{x}(s)$, then $\frac{\partial}{\partial\eta}x^*(s)$ has the same sign as $\frac{\partial}{\partial\eta}u(xc|x,s)|_{x=x^*(s)}$.*

Proof of Lemma A.1: Consider a fixed $s \in [0, \bar{s}]$. From Lemma 2, we know that when $x^*(s) = \hat{x}(s)$, $x^*(s)$ satisfies the equation $u(x^*(s)c|x^*(s), s) = 0$. By taking the derivative of both sides of this equation with respect to η, we have

$$\frac{\partial}{\partial x}u(xc|x,s)\Big|_{x=x^*(s)} \times \frac{\partial}{\partial \eta}x^*(s) + \frac{\partial}{\partial \eta}u(xc|x,s)\Big|_{x=x^*(s)} = 0. \tag{A.6}$$

From Equation (A.4), we know that $u(xc|x,s)$ is a quadratic concave function in x. Since $x^*(s)$ is the larger root of the equation $u(x^*(s)c|x^*(s), s) = 0$ by Lemma 2, we have $\frac{\partial}{\partial x}u(xc|x,s)|_{x=x^*(s)} < 0$. By plugging this into Equation (A.6), we conclude that $\frac{\partial}{\partial \eta}u(xc|x,s)|_{x=x^*(s)}$ has the same sign as $\frac{\partial}{\partial \eta}x^*(s)$. □

Lemma A.2. *For $s \in [0, \bar{s}]$, if $x^*(s) = \hat{x}(s)$, then $\frac{\partial^2}{\partial s^2}x^*(s) < 0$. Moreover, $\frac{\partial}{\partial s}x^*(s)|_{s=0} > 0$.*

Proof of Lemma A.2: First, we prove that $\hat{x}(s)$ is a strictly concave function in s. By Lemma 2, we have

$$\hat{x}(s) = \frac{-\beta(s) + \sqrt{\beta^2(s) - 4\alpha\gamma(s)}}{2\alpha}.$$

Based on the expressions of α, $\beta(s)$ and $\gamma(s)$ in Lemma 2, we can get that $\beta(s)$ is linear in s and $\beta^2(s) - 4\alpha\gamma(s)$ is a quadratic concave function in s as we consider $r \geq \frac{1}{2}$. Since the square root function (i.e., $\sqrt{\cdot}$) is a concave and increasing function, $\hat{x}(s)$ is a strictly concave function in s. Thus, if $x^*(s) = \hat{x}(s)$, then $\frac{\partial^2}{\partial s^2}x^*(s) < 0$.

Next, we prove that $\frac{\partial}{\partial s}x^*(s)|_{s=0} > 0$. By Lemma 1, we have $x^*(0) = x_0^* = \hat{x}(0)$. Then, according to Lemma A.1, $\frac{\partial}{\partial s}x^*(s)|_{s=0}$ has the same sign as $\frac{\partial}{\partial s}u(xc|x,s)|_{x=x_0^*, s=0}$. We have

$$\frac{\partial}{\partial s}u(xc|x,s)\Big|_{x=x_0^*, s=0} = \frac{1}{2} + \lambda\left(r - \frac{1}{2}\right)(R_H(x_0^*) - R_L(x_0^*)).$$

Since we consider $r \geq \frac{1}{2}$ and $R_H(x_0^*) > R_L(x_0^*)$ (i.e., high-yield revenue is not lower than low-yield revenue without any subsidy), we have $\frac{\partial}{\partial s}u(xc|x,s)|_{x=x_0^*, s=0} > 0$. Thus, $\frac{\partial}{\partial s}x^*(s)|_{s=0} > 0$. □

Proof of Proposition 1: Given $r \in [\frac{1}{2}, 1]$, since we consider $x^*(s) = \hat{x}(s)$ for any $s \in [0, \bar{s}]$, according to Lemma A.2, $x^*(s)$ is strictly concave in s, and we define $s_1 := \arg\max_{s \in [0,\bar{s}]} x^*(s)$. Then, by the definition of s_1, we have $x^*(s)$ increases in s if and only if $s \leq s_1$. Moreover, we have $s_1 > 0$ because we showed in Lemma A.2 that $\frac{\partial}{\partial s}x^*(s)|_{s=0} > 0$.

Next, we prove that there exists a threshold r_1 such that if $r < r_1$, then we have $s_1 < \bar{s}$ and s_1 increases in r; if $r \geq r_1$, then we have $s_1 = \bar{s}$ (i.e., $x^*(s)$ increases in s for all $s \in [0, \bar{s}]$). We prove these results in the following three steps:

Step 1: We prove that for any given $r \in [\frac{1}{2}, 1]$, if $s_1 < \bar{s}$, then s_1 strictly increases in r. According to Lemma A.2, $x^*(s)$ is strictly concave in s. Thus, given $s_1 < \bar{s}$, we must have $\frac{\partial}{\partial s}x^*(s)|_{s=s_1} = 0$. Then, in order to prove that s_1 increases in the r, we would like to prove that $\frac{\partial^2}{\partial s \partial r}x^*(s)|_{s=s_1} > 0$.

Since $x^*(s_1) = \hat{x}(s_1)$, Equation (A.6) holds at $s = s_1$. By taking the derivative with respect to r of both sides of Equation (A.6) and substituting the subsidy amount s for η, we have

$$\left(\frac{\partial^2}{\partial x^2}u(xc|x,s)\Big|_{x=x^*(s_1),s=s_1}\frac{\partial}{\partial r}x^*(s)\Big|_{s=s_1}+\frac{\partial^2}{\partial x\partial r}u(xc|x,s)\Big|_{x=x^*(s_1),s=s_1}\right)\frac{\partial}{\partial s}x^*(s)\Big|_{s=s_1}$$
$$+\frac{\partial}{\partial x}u(xc|x,s)\Big|_{x=x^*(s_1),s=s_1}\frac{\partial^2}{\partial s\partial r}x^*(s)\Big|_{s=s_1}+\frac{\partial^2}{\partial x\partial s}u(xc|x,s)\Big|_{x=x^*(s_1),s=s_1}\frac{\partial}{\partial r}x^*(s)\Big|_{s=s_1}$$
$$+\frac{\partial^2}{\partial s\partial r}u(xc|x,s)\Big|_{x=x^*(s_1),s=s_1}=0.$$

Because we have $\frac{\partial}{\partial s}x^*(s)|_{s=s_1}=0$, the first term vanishes and the equation above can be reduced to

$$\frac{\partial}{\partial x}u(xc|x,s)\Big|_{x=x^*(s_1),s=s_1}\frac{\partial^2}{\partial s\partial r}x^*(s)\Big|_{s=s_1}+\frac{\partial^2}{\partial x\partial s}u(xc|x,s)\Big|_{x=x^*(s_1),s=s_1}\frac{\partial}{\partial r}x^*(s)\Big|_{s=s_1}$$
$$+\frac{\partial^2}{\partial s\partial r}u(xc|x,s)\Big|_{x=x^*(s_1),s=s_1}=0. \tag{A.7}$$

In addition, because discussed in the proof of Lemma A.1 that $\frac{\partial}{\partial x}u(xc|x,s)|_{x=x^*(s)}<0$, in order to prove that $\frac{\partial^2}{\partial s\partial r}x^*(s)|_{s=s_1}>0$, we only need to prove that $\frac{\partial^2}{\partial x\partial s}u(xc|x,s)|_{x=x^*(s_1),s=s_1}\frac{\partial}{\partial r}x^*(s)|_{s=s_1}+\frac{\partial^2}{\partial s\partial r}u(xc|x,s)|_{x=x^*(s_1),s=s_1}>0$.

From Equation (A.4), we can get $\frac{\partial^2}{\partial x\partial s}u(xc|x,s)=-4\lambda b(r-\frac{1}{2})\sigma\mu$ and $\frac{\partial^2}{\partial s\partial r}u(xc|x,s)=\lambda(R_H(x)-R_L(x))$. From Equation (A.6), we have

$$\begin{aligned}\frac{\partial}{\partial r}x^*(s)&=-\frac{\partial u(xc|x,s)/\partial r}{\partial u(xc|x,s)/\partial x}\Big|_{x=x^*(s)}\\&=\frac{\lambda s[R_H(x^*(s))-R_L(x^*(s))]}{2\alpha x^*(s)+\beta(s)}\\&=\frac{\lambda s[R_H(x^*(s))-R_L(x^*(s))]}{\sqrt{\beta^2(s)-4\alpha\gamma(s)}}.\end{aligned}$$

The third equality is achieved by substituting the expression in Equation (A.5) for $x^*(s)$. Then, we can write

$$\frac{\partial^2}{\partial x\partial s}u(xc|x,s)\Big|_{x=x^*(s_1),s=s_1}\frac{\partial}{\partial r}x^*(s)\Big|_{s=s_1}+\frac{\partial^2}{\partial s\partial r}u(xc|x,s)\Big|_{x=x^*(s_1),s=s_1}$$
$$=\lambda[R_H(x^*(s_1))-R_L(x^*(s_1))]\left[1-\frac{4\lambda b(r-\frac{1}{2})\sigma\mu s_1}{\sqrt{\beta^2(s_1)-4\alpha\gamma(s_1)}}\right] \tag{A.8}$$

Since we consider $s\in[0,\bar{s}]$ and $\bar{s}$ is the lowest subsidy amount that satisfies $R_L(x^*(s))+s\geq R_H(x^*(s))$, we have $R_H(x^*(s))-R_L(x^*(s))\geq s\geq 0$ for all $s\in[0,\bar{s}]$. Then, the right hand side of Equation (A.8) is positive if and if only $\sqrt{\beta^2(s_1)-4\alpha\gamma(s_1)}>4\lambda b(r-\frac{1}{2})\sigma\mu s_1$.

By the definition of s_1, we have

$$\frac{\partial}{\partial s}x^*(s)\Big|_{s=s_1}=\frac{1}{2\alpha}\left(-4\lambda b(r-\frac{1}{2})\sigma\mu+\frac{\partial(\beta^2(s)-4\alpha\gamma(s))/\partial s}{\sqrt{\beta^2(s)-4\alpha\gamma(s)}}\Big|_{s=s_1}\right)=0$$

and thus

$$4\lambda b(r-\frac{1}{2})\sigma\mu = \frac{\partial(\beta^2(s)-4\alpha\gamma(s))/\partial s}{\sqrt{\beta^2(s)-4\alpha\gamma(s)}}\Big|_{s=s_1}.$$

Multiplying both sides of the equation above by s_1, we can have

$$4\lambda b(r-\frac{1}{2})\sigma\mu s_1 = \frac{\partial(\beta^2(s)-4\alpha\gamma(s))/\partial s}{\sqrt{\beta^2(s)-4\alpha\gamma(s)}}\Big|_{s=s_1} \times s_1. \tag{A.9}$$

Further, as shown in the proof of Lemma A.2, $\beta^2(s)-4\alpha\gamma(s)$ is concave in s and increasing in s at $s=0$, which implies that

$$[\beta^2(s_1)-4\alpha\gamma(s_1)]-[\beta^2(0)-4\alpha\gamma(0)] > \frac{\partial}{\partial s}\left(\beta^2(s)-4\alpha\gamma(s)\right)\Big|_{s=s_1} \times s_1.$$

We have that $\beta^2(0)-4\alpha\gamma(0)=\beta_0^2-4\alpha_0\gamma_0>0$ as implied by Lemma 1. So, we have

$$\sqrt{\beta^2(s_1)-4\alpha\gamma(s_1)} > \frac{\partial(\beta^2(s)-4\alpha\gamma(s))/\partial s}{\sqrt{\beta^2(s)-4\alpha\gamma(s)}}\Big|_{s=s_1} \times s_1. \tag{A.10}$$

Hence, by combining (A.9) and (A.10), we have $\sqrt{\beta^2(s_1)-4\alpha\gamma(s_1)} > 4\lambda b(r-\frac{1}{2})\mu\sigma s_1$, and thus $\frac{\partial^2}{\partial s\partial r}x^*(s)|_{s=s_1}>0$, which implies that s_1 strictly increases in r.

Step 2: We prove that if $s_1=\bar{s}$ under some $r\in[\frac{1}{2},1]$, then we have $s_1=\bar{s}$ under any higher index accuracy. Specifically, let $r_1\in[\frac{1}{2},1]$ denote the lowest index accuracy under which $s_1=\bar{s}$ (we prove the existence of such r_1 in Step 3). Then, it is sufficient to prove that we have $s_1=\bar{s}$ for any $r\geq r_1$.

Recall that $\hat{x}(s)=\frac{-\beta(s)+\sqrt{\beta^2(s)-4\alpha\gamma(s)}}{2\alpha}$ by Lemma 2. As shown in the proof of Lemma A.2, we have that $\beta(s)$ is linearly increasing in s, $\beta^2(s)-4\alpha\gamma(s)$ is a quadratic concave function in s, and $\hat{x}(s)$ is increasing in s when $s=0$. Therefore, there must exist a threshold, denoted as $\hat{s}$, such that $\frac{\partial}{\partial s}\hat{x}(s)|_{s=\hat{s}}=0$ and $\hat{x}(s)$ increases in s if and only if $s\leq\hat{s}$. Then, by definition of s_1, we have $s_1=\hat{s}<\bar{s}$ when $r<r_1$.

Then, to prove the goal of this step, it is sufficient to prove that if $r\geq r_1$, then $\hat{s}\geq\bar{s}$. Therefore, it is sufficient to prove that for any given r such that $\hat{s}=\bar{s}$, we have $\frac{\partial}{\partial r}\hat{s}>0$ and $\frac{\partial}{\partial r}\bar{s}\leq 0$.

Following the analysis in Step 1, we have that, whenever $\hat{s}\leq\bar{s}$, $\hat{s}$ strictly increases in r. Thus, it remains to study how $\bar{s}$ changes in r when $\hat{s}=\bar{s}$. Since $x^*(s)$ is continuous in s, $R_H(x^*(s))-R_L(x^*(s))$ is also continuous in s. Moreover, when $s=0$, we assumed $R_H(x^*(s))-R_L(x^*(s))>s$ (i.e., $R_H(x^*(0))-R_L(x^*(0))>0$) and we have shown in the proof of Lemma 2 that when $s=2a\sigma$, $R_H(x^*(s))-R_L(x^*(s))\leq s$ (i.e., $R_H(x^*(2a\sigma))-R_L(x^*(2a\sigma))\leq 2a\sigma$). Recall that $\bar{s}$ is the lowest subsidy amount that satisfies $R_L(x^*(s))+s\geq R_H(x^*(s))$. Thus, we have

$$\bar{s}=R_H(x^*(\bar{s}))-R_L(x^*(\bar{s}))=2a\sigma-4b\mu\sigma x^*(\bar{s}). \tag{A.11}$$

Then, we take derivative of both sides of Equation (A.11) with respect to r, and have

$$\frac{\partial}{\partial r}\bar{s} = -4b\mu\sigma\left(\frac{\partial}{\partial r}x^*(s)\Big|_{s=\bar{s}} + \frac{\partial}{\partial s}x^*(s)\Big|_{s=\bar{s}}\frac{\partial}{\partial r}\bar{s}\right). \tag{A.12}$$

Note that when $\hat{s} = \bar{s}$, $\frac{\partial}{\partial s}x^*(s)|_{s=\bar{s}} = 0$. Thus, in order to prove that $\frac{\partial}{\partial r}\bar{s} \le 0$ when $\hat{s} = \bar{s}$, it remains to prove that $\frac{\partial}{\partial r}x^*(s)|_{s=\bar{s}} \ge 0$.

For any $s \in [0, \bar{s}]$, since $x^*(s) = \hat{x}(s)$, by Lemma A.1, $\frac{\partial}{\partial r}x^*(s)$ and $\frac{\partial}{\partial r}u(xc|x,s)|_{x=x^*(s)}$ has the same sign. We have

$$\frac{\partial}{\partial r}u(xc|x,s)\Big|_{x=x^*(s)} = \lambda s[R_H(x^*(s)) - R_L(x^*(s))].$$

Since we consider $s \in [0, \bar{s}]$ and $\bar{s}$ is defined in §3 as the lowest subsidy amount that satisfies $R_L(x^*(s)) + s \ge R_H(x^*(s))$, we have $R_H(x^*(s)) - R_L(x^*(s)) \ge s \ge 0$. Thus, $\frac{\partial}{\partial r}u(xc|x,s)|_{x=x^*(s)} \ge 0$ and we have, for any $s \in [0, \bar{s}]$,

$$\frac{\partial}{\partial r}x^*(s) \ge 0. \tag{A.13}$$

So, $\frac{\partial}{\partial r}x^*(s)|_{s=\bar{s}} \ge 0$ and thus by Equation (A.12), $\frac{\partial}{\partial r}\bar{s} \le 0$ when $\hat{s} = \bar{s}$.

Step 3: We prove the existence of $r_1 \in [\frac{1}{2}, 1]$, where r_1 is defined in Step 2 as the lowest index accuracy under which $s_1 = \bar{s}$, and we also prove that $r_1 < 1$. We first prove that such $r_1 \in [\frac{1}{2}, 1]$ exists. To do so, it is sufficient to prove that when $r = 1$, we have $s_1 = \bar{s}$. We prove this result by contradiction. Suppose $s_1 < \bar{s}$ for all $r \in [\frac{1}{2}, 1]$. Then, based on the definition of $\bar{s}$, we have $s_1 < R_H(x^*(s_1)) - R_L(x^*(s_1))$ for all $r \in [\frac{1}{2}, 1]$. On the other hand, recall that $s_1 = \arg\max_{s\in[0,\bar{s}]} x^*(s)$. For any given $r \in [\frac{1}{2}, 1]$, since $s_1 < \bar{s}$, for any $s \in [s_1, \bar{s}]$, we must have $\frac{\partial}{\partial s}x^*(s) \le 0$. Then, according to Lemma A.1, we have $\frac{\partial}{\partial s}u(xc|x,s)|_{x=x^*(s_1),s=s_1} \le 0$, which is equivalent to

$$s_1 \ge \frac{1}{\lambda} + 2(r - \frac{1}{2})[R_H(x^*(s_1)) - R_L(x^*(s_1))].$$

Plugging $r = 1$ into the above inequality, we have $s_1 \ge \frac{1}{\lambda} + R_H(x^*(s_1)) - R_L(x^*(s_1)) > R_H(x^*(s_1)) - R_L(x^*(s_1))$, which leads to a contradiction. Thus, we have $s_1 = \bar{s}$ when $r = 1$. This implies the existence of r_1.

Next, we prove that $r_1 < 1$ by contradiction. Suppose $r_1 = 1$. Then, as discussed in Step 2, we have $s_1 = \hat{s} = \bar{s}$ when $r = r_1 = 1$. By the definition of $\hat{s}$ in Step 2, it implies that $\frac{\partial}{\partial s}x^*(s)|_{s=s_1} = 0$. Thus, according to Lemma A.1, we have $\frac{\partial}{\partial s}u(xc|x,s)|_{x=x^*(s_1),s=s_1} = 0$, which implies

$$s_1 = \frac{1}{\lambda} + 2(r - \frac{1}{2})[R_H(x^*(s_1)) - R_L(x^*(s_1))]. \tag{A.14}$$

By Equation (A.14), when $r = r_1 = 1$, we have $s_1 > [R_H(x^*(s_1)) - R_L(x^*(s_1))]$. It contradicts with the earlier conclusion that $s_1 = \bar{s}$ when $r = r_1 = 1$. Hence, we conclude that $r_1 < 1$.

Collectively, putting together all three steps, we have that, there exists a threshold $r_1 \in [\frac{1}{2}, 1)$ such that when $r < r_1$, we have $s_1 < \bar{s}$ and s_1 increases in r; when $r \geq r_1$, we have $s_1 = \bar{s}$. The proof of this part of the proposition is completed. □

Proof of Proposition 2: To find how the income variance changes in s, we start off by showing that the income variance is strictly convex in s. From Equation (A.3), the variance of farmer income is independent of the production cost h and thus is the same for all the farmers who plant. Thus, we only need to consider the farmer with production cost $x^*(s)c$ and show that $\mathbb{V}\text{ar}[\pi(x^*(s)c|x^*(s), s)]$ is strictly convex in s.

According to Lemma 2, since we consider $x^*(s) = \hat{x}(s) \in (0, 1)$, we have $u(x^*(s)c|x^*(s), s) = 0$. By taking the second derivative of both sides of $u(x^*(s)c|x^*(s), s) = 0$ with respect to the subsidy amount s, we have

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathbb{E}[\pi(x^*(s)c|x^*(s), s)] - \lambda\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathbb{V}\text{ar}[\pi(x^*(s)c|x^*(s), s)] = 0,$$

which implies that $\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathbb{V}\text{ar}[\pi(x^*(s)c|x^*(s), s)]$ and $\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathbb{E}[\pi(x^*(s)c|x^*(s), s)]$ have the same sign. We also have

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathbb{E}[\pi(x^*(s)c|x^*(s), s)] = -[b(\mu^2 + \sigma^2) + c]\frac{\partial^2}{\partial s^2}x^*(s) > 0$$

where the inequality holds because, as we mentioned in Lemma A.2, $x^*(s) = \hat{x}(s)$ is strictly concave in s. So, the variance of the farmer income $\mathbb{V}\text{ar}[\pi(h|x^*(s), s)]$ is strictly convex in s.

Then, building on the convexity of the variance in s, in order to justify the statement in the proposition, it is sufficient to show that the variance strictly decreases at $s = 0$ and increases at some $s \in (0, \bar{s}]$. Further, we have, from Equation (A.3),

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{V}\text{ar}[\pi(h|x^*(s), s)] =& \frac{1}{2}[R_H(x^*(s)) - R_L(x^*(s))]\frac{\partial}{\partial s}[R_H(x^*(s)) - R_L(x^*(s))] + \frac{1}{2}s \\ &- \left(r - \frac{1}{2}\right)[R_H(x^*(s)) - R_L(x^*(s))] - s\left(r - \frac{1}{2}\right)\frac{\partial}{\partial s}[R_H(x^*(s)) - R_L(x^*(s))] \\ =& -2b\mu\sigma[R_H(x^*(s)) - R_L(x^*(s))]\frac{\partial}{\partial s}x^*(s) + \frac{1}{2}s \\ &- \left(r - \frac{1}{2}\right)[R_H(x^*(s)) - R_L(x^*(s))] + 4\left(r - \frac{1}{2}\right)b\mu\sigma s\frac{\partial}{\partial s}x^*(s). \end{aligned} \quad \text{(A.15)}$$

From Equation (A.15), at $s = 0$, we have

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{V}\text{ar}[\pi(h|x^*(s), s)]\Big|_{s=0} = -2b\mu\sigma[R_H(x_0^*) - R_L(x_0^*)]\frac{\partial}{\partial s}x^*(s)\Big|_{s=0} - \left(r - \frac{1}{2}\right)[R_H(x_0^*) - R_L(x_0^*)].$$

By Lemma A.2, we know $\frac{\partial}{\partial s}x^*(s)|_{s=0} > 0$. Besides, we assumed $r \in [\frac{1}{2}, 1]$ and $R_H(x_0^*) > R_L(x_0^*)$ in §3. Therefore, the income variance strictly decreases in s at $s = 0$. It remains to show that the income variance increases at some $s \in (0, \bar{s}]$.

Next, we prove that the income variance increases at some $s \in (0, \bar{s}]$. Based on the result of Proposition 1, we separate the analysis into two cases: In Case 1, $\frac{\partial}{\partial s}x^*(s) \geq 0$ for all $s \in [0, \bar{s}]$ (i.e., $r \geq r_1$); and in Case 2, there exists a threshold $s_1 < \bar{s}$ such that $\frac{\partial}{\partial s}x^*(s)|_{s=s_1} = 0$ (i.e., $r < r_1$).

Case 1: We consider the scenario where $\frac{\partial}{\partial s}x^*(s) \geq 0$ for all $s \in [0, \bar{s}]$. We would like to prove that the income variance increases in s at $s = \bar{s}$. By Equation (A.11), we have $\bar{s} = R_H(x^*(\bar{s})) - R_L(x^*(\bar{s}))$. Plugging it into Equation (A.15), we have

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{V}\mathrm{ar}[\pi(h|x^*(s), s)]\Big|_{s=\bar{s}} = -2b\mu\sigma\bar{s}\frac{\partial}{\partial s}x^*(s)\Big|_{s=\bar{s}} + \frac{1}{2}\bar{s} - \left(r - \frac{1}{2}\right)\bar{s} + 4\left(r - \frac{1}{2}\right)b\mu\sigma\bar{s}\frac{\partial}{\partial s}x^*(s)\Big|_{s=\bar{s}}$$
$$= \left(-4b\mu\sigma\frac{\partial}{\partial s}x^*(s)\Big|_{s=\bar{s}} + 1\right)(1 - r)\bar{s} \tag{A.16}$$

Since $r \leq 1$ and $\bar{s} > 0$, in order to prove that $\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{V}\mathrm{ar}[\pi(h|x^*(s), s)]|_{s=\bar{s}} \geq 0$, we only need to prove that $-4b\mu\sigma\frac{\partial}{\partial s}x^*(s)|_{s=\bar{s}} + 1 > 0$.

From Equation (A.6), we have

$$\begin{aligned}
4b\mu\sigma\frac{\partial}{\partial s}x^*(s)\Big|_{s=\bar{s}} &= -4b\mu\sigma\frac{\partial u(xc|x,s)/\partial s}{\partial u(xc|x,s)/\partial x}\Big|_{s=\bar{s},x=x^*(\bar{s})} \\
&= \frac{2b\mu\sigma - 2\lambda b\mu\sigma\bar{s} + 4\lambda b\left(r - \frac{1}{2}\right)\mu\sigma\bar{s}}{2\alpha x^*(\bar{s}) + \beta(\bar{s})} \\
&= \frac{2b\mu\sigma - 2\lambda b\mu\sigma(2a\sigma - 4b\mu\sigma x^*(\bar{s})) + 4\lambda b\left(r - \frac{1}{2}\right)\mu\sigma\bar{s}}{2\alpha x^*(\bar{s}) + \beta(\bar{s})} \\
&= \frac{2b\mu\sigma - 4\lambda ab\mu\sigma^2 + 8\lambda b^2\mu^2\sigma^2 x^*(\bar{s}) + 4\lambda b\left(r - \frac{1}{2}\right)\mu\sigma\bar{s}}{2\alpha x^*(\bar{s}) + \beta(\bar{s})} \\
&= \frac{2\alpha x^*(\bar{s}) + \left(2b\mu\sigma - 4\lambda ab\mu\sigma^2 + 4\lambda b\left(r - \frac{1}{2}\right)\mu\sigma\bar{s}\right)}{2\alpha x^*(\bar{s}) + \beta(\bar{s})} \\
&< \frac{2\alpha x^*(\bar{s}) + \beta(\bar{s})}{2\alpha x^*(\bar{s}) + \beta(\bar{s})} \\
&= 1,
\end{aligned}$$

where the inequality holds because $b(\mu^2 + \sigma^2) \geq 2b\mu\sigma$ and $c > 0$. Therefore, $-4b\mu\sigma\frac{\partial}{\partial s}x^*(s)|_{s=\bar{s}} + 1 > 0$ and $\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{V}\mathrm{ar}[\pi(h|x^*(s), s)]|_{s=\bar{s}} \geq 0$.

Moreover, from Equation (A.16), we have $\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{V}\mathrm{ar}[\pi(h|x^*(s), s)]|_{s=\bar{s}} = 0$ if and only if $r = 1$, which indicates that the income variance decreases in s for $s \in [0, \bar{s}]$ when $r = 1$.

Case 2: We consider the scenario where there exists a threshold $s_1 < \bar{s}$ such that $\frac{\partial}{\partial s}x^*(s)|_{s=s_1} = 0$. We would like to prove that the income variance increases in s at $s = s_1$.

By plugging $s = s_1$ into Equation (A.15), we have

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{V}\mathrm{ar}[\pi(h|x^*(s), s)]\Big|_{s=s_1} = \frac{1}{2}s_1 - \left(r - \frac{1}{2}\right)[R_H(x^*(s_1)) - R_L(x^*(s_1))].$$

Since $\frac{\partial}{\partial s}x^*(s)|_{s=s_1} = 0$, by Equation (A.14), we have $s_1 = \frac{1}{\lambda} + 2(r - \frac{1}{2})[R_H(x^*(s_1)) - R_L(x^*(s_1))]$. Therefore, $\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{V}\mathrm{ar}[\pi(h|x^*(s), s)]|_{s=s_1} = \frac{1}{2}s_1 - (r - \frac{1}{2})[R_H(x^*(s)) - R_L(x^*(s))] = \frac{1}{2\lambda} > 0$.

Note that in this case, we must have $r<1$. Otherwise, if $r=1$, then $s_1=\frac{1}{\lambda}+[R_H(x^*(s_1))-R_L(x^*(s_1))]>R_H(x^*(s_1))-R_L(x^*(s_1))$, which contradicts with the assumption that $s_1<\bar{s}$.

Combining two cases, we proved that, when $r\in[\frac{1}{2},1)$, $\mathbb{V}\mathrm{ar}[\pi(h|x^*(s),s)]$ first decreases and then increases in the subsidy amount s; when $r=1$, $\mathbb{V}\mathrm{ar}[\pi(h|x^*(s),s)]$ only decreases in s.

At the end, we prove that $s_1\geq s_2$. Since we know that $s_1,s_2\in[0,\bar{s}]$ and we have shown in Proposition 1 that $s_1=\bar{s}$ when $r\geq r_1$, we only need to prove $s_1\geq s_2$ when $r<r_1$. In Case 2, we have that when $r<r_1$, the income variance increases in s at $s=s_1$. Therefore, by the definition of s_2, we have $s_1\geq s_2$.

Remark that we present how the expected farmer income changes in s in Proposition A.1. □

Proposition A.1. *There exists a threshold b_1 such that if $b\leq b_1$, then the expected farmer income $\mathbb{E}[\pi(h|x^*(s),s)]$ increases in s.*

Proof of Proposition A.1: First, we show that the expected farmer income is convex in s. In equilibrium, given a subsidy amount s, the expected income of the farmer with production cost h is

$$\mathbb{E}[\pi(h|x^*(s),s)]=a\mu-bx^*(s)\times(\mu^2+\sigma^2)+\frac{1}{2}s-h.$$

By taking the second derivative of $\mathbb{E}[\pi(h|x^*(s),s)]$ with respect to s, we have

$$\frac{\partial^2}{\partial s^2}\mathbb{E}[\pi(h|x^*(s),s)]=-b(\mu^2+\sigma^2)\frac{\partial^2}{\partial s^2}x^*(s)>0,$$

where the inequality holds because we have shown in Lemma A.2 that $\frac{\partial^2}{\partial s^2}x^*(s)<0$. Thus, the expected farmer income is convex in s.

Then, due to the convexity, in order to study how the expected income changes in s, we focus on how it changes in s at $s=0$. We take the first derivative of $\mathbb{E}[\pi(h|x^*(s),s)]$ with respect to s at $s=0$ and have

$$\frac{\partial}{\partial s}\mathbb{E}[\pi(g|x^*(s),s)]\Big|_{s=0}=-b(\mu^2+\sigma^2)\frac{\partial}{\partial s}x^*(s)\Big|_{s=0}+\frac{1}{2}.$$

We have shown in Lemma A.2 that $\frac{\partial}{\partial s}x^*(s)|_{s=0}>0$. Therefore, there must exist a $b_1>0$ such that for any $b\leq b_1$, $\frac{\partial}{\partial s}\mathbb{E}[\pi(g|x^*(s),s)]|_{s=0}\geq 0$. Since the expected farmer income is convex in s, if $b\leq b_1$, then the expected farmer income increases in s. □

Proof of Corollary 1: By Equation (A.13), we have shown that $x^*(s)$ increases in r. It remains to prove that the variance of farmer income $\mathbb{V}\mathrm{ar}[\pi(h|x^*(s),s)]$ decreases in r.

From Equation (A.3), we have

$$\frac{\mathrm{d}}{\mathrm{d}r}\mathbb{V}\mathrm{ar}[\pi(h|x^*(s),s)]=\frac{1}{2}[R_H(x^*(s))-R_L(x^*(s))]\times\frac{\partial}{\partial r}[R_H(x^*(s))-R_L(x^*(s))]$$

$$
\begin{aligned}
&- s[R_H(x^*(s)) - R_L(x^*(s))] - s\left(r - \frac{1}{2}\right) \times \frac{\partial}{\partial r}[R_H(x^*(s)) - R_L(x^*(s))] \\
=&\frac{1}{2}[R_H(x^*(s)) - R_L(x^*(s))] \times \left(-4b\mu\sigma \frac{\partial}{\partial r} x^*(s)\right) \\
&- s[R_H(x^*(s)) - R_L(x^*(s))] - s\left(r - \frac{1}{2}\right) \times \left(-4b\mu\sigma \frac{\partial}{\partial r} x^*(s)\right) \\
=&-2b\mu\sigma \left(\frac{\partial}{\partial r} x^*(s)\right) \left([R_H(x^*(s)) - R_L(x^*(s))] - 2\left(r - \frac{1}{2}\right) s\right) \\
&- s[R_H(x^*(s)) - R_L(x^*(s))] \\
\leq&0,
\end{aligned}
$$

where the inequality holds because we consider $r \in [\frac{1}{2}, 1]$ and we have shown in the proof of Proposition 1 that $\frac{\partial}{\partial r} x^*(s) \geq 0$ and $R_H(x^*(s)) - R_L(x^*(s)) \geq s \geq 0$. Thus, we have $\mathbb{V}\text{ar}[\pi(h|x^*(s), s)]$ decreases in r.

□

Proof of Theorem 1: Recall the formulation of $v(s)$ in Equation (4):

$$
\begin{aligned}
v(s) &= \int_0^{x^*(s)} \mathbb{E}[(a - bx^*(s)Y)Y - xc + sI] \, \mathrm{d}x - x^*(s)\mathbb{E}[sI] \\
&= v_1 \times x^*(s) - v_2 \times (x^*(s))^2.
\end{aligned}
$$

where $v_1 = a\mu$ and $v_2 = b(\mu^2 + \sigma^2) + \frac{1}{2}c$. Since $v_1 > 0$ and $v_2 > 0$, $v(s)$ is a quadratic concave function in $x^*(s)$. Then, the value of $x^*(s)$ that maximizes $v(s)$ is defined as follows:

$$
x^{opt} = \min\left\{\frac{v_1}{2v_2}, 1\right\} = \min\left\{\frac{a\mu}{2b(\mu^2 + \sigma^2) + c}, 1\right\}. \tag{A.17}
$$

As the government's objective is to find the subsidy amount s to maximize $v(s)$, the optimal s^* must make $x^*(s)$ as close to x^{opt} as possible. In addition, we assumed $x_0^* \leq x^{opt}$ in §3. Therefore, as we discussed in §3, s^* must satisfy either (1) $x^*(s^*) < x^{opt}$ and $s^* = \inf\{\arg\max_{s \in [0,\bar{s}]} x^*(s)\}$, or (2) $s^* = \inf\{s : x^*(s) = x^{opt}\}$.

We define $r_c \geq \frac{1}{2}$ as the smallest index accuracy under which $x^*(s) = x^{opt}$ for some $s \in [0, \bar{s}]$. (We prove the existence of such r_c at the end.) Then, we separate the rest of the proof into two cases: In Case 1, $r < r_c$ (this case does not exist if $r_c = \frac{1}{2}$); and in Case 2, $r \geq r_c$.

Case 1: We consider the scenario where $r < r_c$. First, we characterize s^*. By definition of r_c, in this case, for any $s \in [0, \bar{s}]$, we have $x^*(s) < x^{opt}$. Thus, as discussed earlier, s^* must satisfy $x^*(s^*) < x^{opt}$ and $s^* = \inf\{\arg\max_{s \in [0,\bar{s}]} x^*(s)\}$. Moreover, from Lemma A.2, we have that $x^*(s)$ is a strictly concave function in s. Therefore, $x^*(s)$ has a unique maximizer and $s^* = \arg\max_{s \in [0,\bar{s}]} x^*(s)$.

Next, we study how s^* changes in r. Recall that r_1 is a threshold defined in Proposition 1 such that if $r < r_1$, $x^*(s)$ increases in s if and only if $s \leq s_1$; and if $r \geq r_1$, $x^*(s)$ increases in s. We further separate this case into two sub-cases: (a) $r < r_1$ and (b) $r \in [r_1, r_c)$ (this sub-case does not exist if $r_1 \geq r_c$).

(a): We consider the scenario where $r < r_1$. Then, we have $s^* = \arg\max_{s\in[0,\bar{s}]} x^*(s) = s_1$ by the result of Proposition 1. Moreover, we have shown in Proposition 1 that s_1 increases in r. Thus, s^* increases in r.

(b): We consider the scenario where $r \in [r_1, r_c)$. Then, $s^* = \arg\max_{s\in[0,\bar{s}]} x^*(s) = \bar{s}$ by the result of Proposition 1. Further, by Proposition 1, when $r \geq r_1$, we have $\frac{\partial}{\partial s} x^*(s)|_{s=\bar{s}} \geq 0$, which, by Equation (A.12), implies $\frac{\partial}{\partial r}\bar{s} \leq 0$. Thus, we have s^* decreases in r.

Case 2: We consider the scenario where $r \geq r_c$. First, we characterize s^*. By definition of r_c, when $r = r_c$, we have $x^*(s) = x^{opt}$ for some $s \in [0, \bar{s}]$. Since we consider $x^*(0) \leq x^{opt}$ and we have $x^*(s)$ increases in r by Corollary 1, for any $r \geq r_c$, we must also have $x^*(s) = x^{opt}$ for some $s \in [0, \bar{s}]$. Therefore, in this case, $s^* = \inf\{s : x^*(s) = x^{opt}\}$.

Next, we study how s^* changes in r. Since we consider $x^*(0) \leq x^{opt}$ and s^* is the lowest subsidy amount that satisfies $x^*(s^*) = x^{opt}$, we must have that $x^*(s)$ increases in s at $s = s^*$. Moreover, we showed in Corollary 1 that $x^*(s)$ increases in r. Hence, s^* must decrease in r to maintain $s^* = \inf\{s : x^*(s) = x^{opt}\}$.

We define $r_2 := \min\{r_c, r_1\}$. Then, combining Case 1 and Case 2, we have s^* increases in r if and only if $r < r_2$.

As the last step of this proof, we prove the existence of r_c by contradiction. Suppose there is no such r_c, which implies that for all $r \in [\frac{1}{2}, 1]$ and $s \in [0, \bar{s}]$, $x^*(s) < x^{opt}$. However, according to the proof of Lemma 2, as we plug in $r = 1$ and $s = \bar{s}$ into Equation (A.4) and then solve it, we get

$$x^*(s) = \frac{a(\mu+\sigma)}{b(\mu+\sigma)^2 + c} > \min\left\{\frac{a\mu}{2b(\mu^2+\sigma^2)+c}, 1\right\} = x^{opt}$$

which contradicts with the assumption. Therefore, r_c must exist in $[\frac{1}{2}, 1]$. □

To help prove Proposition 3, we need the following auxiliary lemma to study the relationship between the first derivative of $x^*(s)$ and the first derivative of x^{opt}, both with respect to the yield variability σ.

Lemma A.4. *Let $v_1 = a\mu$ and $v_2 = b(\mu^2+\sigma^2) + \frac{1}{2}c$. If $x^*(s) = \hat{x}(s)$ and $x^*(s) \leq \frac{v_1}{2v_2} \leq 1$, we have $\frac{\partial}{\partial\sigma} x^*(s) < \frac{\partial}{\partial\sigma} x^{opt} < 0$.*

Proof of Lemma A.4: From Equation (A.17), we have that, when $\frac{v_1}{2v_2} \leq 1$, $x^{opt} = \frac{v_1}{2v_2}$ and $\frac{\partial}{\partial\sigma} x^{opt} < 0$. The rest of the proof is to show $\frac{\partial}{\partial\sigma} x^*(s) < \frac{\partial}{\partial\sigma} x^{opt}$.

For ease of exposition, we let $x^*(s, \lambda, b)$ denote the equilibrium planting amount $x^*(s)$ under risk aversion λ and price sensitivity b. By observing Equation (A.4) and the formulation of x^{opt} in

Equation (A.17), we find that $x^{opt} = \frac{v_1}{2v_2} = x^*(0,0,2b)$. So, to prove this lemma, we only need to show for any $s \in [0,\bar{s}]$, $\frac{\partial}{\partial\sigma}x^*(s,\lambda,b) < \frac{\partial}{\partial\sigma}x^*(0,0,2b)$. We prove that in the following three steps.

Step 1: We show $\frac{\partial}{\partial\sigma}x^*(s,\lambda,b) < \frac{\partial}{\partial\sigma}x^*(s,0,b)$. It is sufficient to prove that $\frac{\partial^2}{\partial\sigma\partial\lambda}x^*(s) < 0$. By Lemma 2, as we consider $x^*(s) = \hat{x}(s)$, $x^*(s)$ satisfies $u(x^*(s)c|x^*(s),s) = 0$. By taking second derivative of both sides of $u(x^*(s)c|x^*(s),s) = 0$ with respect to σ and λ, we have,

$$\left(\frac{\partial^2}{\partial x^2}u(xc|x,s)\Big|_{x=x^*(s)}\frac{\partial}{\partial\sigma}x^*(s) + \frac{\partial^2}{\partial x\partial\sigma}u(xc|x,s)\Big|_{x=x^*(s)}\right)\frac{\partial}{\partial\lambda}x^*(s) + \frac{\partial}{\partial x}u(xc|x,s)\Big|_{x=x^*(s)}\frac{\partial^2}{\partial\sigma\partial\lambda}x^*(s)$$
$$+\frac{\partial^2}{\partial x\partial\lambda}u(xc|x,s)\Big|_{x=x^*(s)}\frac{\partial}{\partial\sigma}x^*(s) + \frac{\partial^2}{\partial\sigma\partial\lambda}u(xc|x,s)\Big|_{x=x^*(s)} = 0.$$

In order to figure out the sign of $\frac{\partial^2}{\partial\sigma\partial\lambda}x^*(s)$, we next would like to find the signs of all other terms in the equation. From Equation (A.4), $u(xc|x,s)$ is a concave quadratic function in x, so $\frac{\partial^2}{\partial x^2}u(xc|x,s)|_{x=x^*(s)} < 0$. Since we showed in the proof of Lemma 2 that $x^*(s)$ is the larger root of $u(xc|x,s) = 0$, $\frac{\partial}{\partial x}u(xc|x,s)|_{x=x^*(s)} < 0$.

For the sign of $\frac{\partial}{\partial\sigma}x^*(s)$ and $\frac{\partial}{\partial\lambda}x^*(s)$, by Lemma A.1, we need to find the sign of $\frac{\partial}{\partial\sigma}u(xc|x,s)$ and $\frac{\partial}{\partial\lambda}u(xc|x,s)$ at $x = x^*(s)$. For $\frac{\partial}{\partial\sigma}u(xc|x,s)$ at $x = x^*(s)$, we have

$$\frac{\partial}{\partial\sigma}u(xc|x,s)\Big|_{x=x^*(s)} = -2bx^*(s)\sigma - \lambda(2a - 4bx^*(s)\mu)\left(\frac{1}{2}[R_H(x^*(s)) - R_L(x^*(s))] - s\left(r - \frac{1}{2}\right)\right). \tag{A.18}$$

As we consider $s \le \bar{s}$, $R_H(x^*(s)) - R_L(x^*(s)) \ge s \ge 0$. Thus $\frac{1}{2}[R_H(x^*(s)) - R_L(x^*(s))] - s(r - \frac{1}{2}) > 0$. Besides, $2a - 4bx^*(s)\mu = [R_H(x^*(s)) - R_L(x^*(s))]/\sigma \ge 0$. Therefore, $\frac{\partial}{\partial\sigma}u(xc|x,s)|_{x=x^*(s)} < 0$ and thus $\frac{\partial}{\partial\sigma}x^*(s) < 0$. Then, for $\frac{\partial}{\partial\lambda}u(xc|x,s)$ at $x = x^*(s)$, we have

$$\frac{\partial}{\partial\lambda}u(xc|x,s)\Big|_{x=x^*(s)} = -\mathbb{V}\mathrm{ar}[\pi(x^*(s)c|x^*(s),s)] < 0. \tag{A.19}$$

So, $\frac{\partial}{\partial\lambda}x^*(s) < 0$.

Moreover, from Equation (A.18) and Equation (A.19), we have

$$\frac{\partial^2}{\partial\sigma\partial\lambda}u(xc|x,s)\Big|_{x=x^*(s)} = -(2a - 4bx^*(s)\mu)\left[\frac{1}{2}[R_H(x^*(s)) - R_L(x^*(s))] - s\left(r - \frac{1}{2}\right)\right] < 0$$
$$\frac{\partial^2}{\partial\sigma\partial x}u(xc|x,s)\Big|_{x=x^*(s)} = 4\lambda b\sigma\left[\frac{1}{2}[R_H(x^*(s)) - R_L(x^*(s))] - s\left(r - \frac{1}{2}\right)\right] > 0$$
$$\frac{\partial^2}{\partial\lambda\partial x}u(xc|x,s)\Big|_{x=x^*(s)} = 2b\mu\sigma\left[\frac{1}{2}[R_H(x^*(s)) - R_L(x^*(s))] - s\left(r - \frac{1}{2}\right)\right] > 0.$$

Plugging these inequalities back, we can get $\frac{\partial^2}{\partial\sigma\partial\lambda}x^*(s) < 0$. Thus $\frac{\partial}{\partial\sigma}x^*(s,\lambda,b) < \frac{\partial}{\partial\sigma}x^*(s,0,b)$.

Step 2: We have $\frac{\partial}{\partial\sigma}x^*(s,0,b) < \frac{\partial}{\partial\sigma}x^*(s,0,2b)$ because

$$\frac{\partial}{\partial\sigma}x^*(s,0,b) = \frac{\partial}{\partial\sigma}\left(\frac{a\mu}{b(\mu^2+\sigma^2)+c}\right) = \frac{-a\mu(\mu^2+\sigma^2)}{[b(\mu^2+\sigma^2)+c]^2},$$

which increases in b.

Step 3: We have $\frac{\partial}{\partial\sigma}x^*(s,0,2b) = \frac{\partial}{\partial\sigma}x^*(0,0,2b)$ because

$$\frac{\partial}{\partial\sigma}x^*(s,0,2b) = \frac{\partial}{\partial\sigma}\left(\frac{a\mu + \frac{1}{2}s}{2b(\mu^2+\sigma^2)+c}\right) = \frac{\partial}{\partial\sigma}\left(\frac{a\mu}{2b(\mu^2+\sigma^2)+c}\right) = \frac{\partial}{\partial\sigma}x^*(0,0,2b).$$

As we combine the three steps, we have $\frac{\partial}{\partial\sigma}x^*(s,\lambda,b) < \frac{\partial}{\partial\sigma}x^*(0,0,2b)$, which implies $\frac{\partial}{\partial\sigma}x^*(s) < \frac{\partial}{\partial\sigma}x^{opt}$. $\square$

Proof of Proposition 3: (i): Recall that r_1 is a threshold defined in Proposition 1 such that if $r < r_1$, $x^*(s)$ increases in s if and only if $s \le s_1$; and if $r \ge r_1$, $x^*(s)$ increases in s. Also, recall that r_c is a threshold defined in the proof of Theorem 1 as the lowest index accuracy under which $x^*(s) = x^{opt}$ for some $s \in [0,\bar{s}]$. From the proof of Theorem 1, given $r \in [\frac{1}{2},1]$, s^* takes its value from one of the following: (1) $s^* = s_1$ if $r < \min\{r_1,r_c\}$, (2) $s^* = \bar{s}$ if $r \in [r_1,r_c)$ (if $r_1 > r_c$, this scenario does not exist), or (3) $s^* = \inf\{s : x^*(s) = x^{opt}\}$ if $r \ge r_c$. Hence we analyze these three scenarios separately: In Case 1, $r < \min\{r_1,r_c\}$; in Case 2, $r \in [r_1,r_c)$; and in Case 3, $r \ge r_c$.

Case 1: We consider the scenario where $r < \min\{r_1,r_c\}$. In this case, we have $s^* = s_1 < \bar{s}$ and $x^*(s_1) < x^{opt}$. Since $s_1 < \bar{s}$, we have $\frac{\partial}{\partial s}x^*(s)|_{s=s_1} = 0$. Then, Equation (A.14) holds. By taking derivative of both sides of Equation (A.14) with respect to σ, we have

$$\begin{aligned}\frac{\partial}{\partial\sigma}s_1 &= 2\left(r-\frac{1}{2}\right)\left[(2a-4b\mu x^*(s_1)) - 4b\mu\sigma\left(\frac{\partial}{\partial\sigma}x^*(s)\Big|_{s=s_1} + \frac{\partial}{\partial s}x^*(s)\Big|_{s=s_1}\frac{\partial}{\partial\sigma}s_1\right)\right]\\ &= 2\left(r-\frac{1}{2}\right)\left[(2a-4b\mu x^*(s_1)) - 4b\mu\sigma\frac{\partial}{\partial\sigma}x^*(s)\Big|_{s=s_1}\right].\end{aligned}$$

We have $2a - 4b\mu x^*(s_1) = [R_H(x^*(s_1)) - R_L(x^*(s_1))]/\sigma > 0$. Besides, according to Lemma A.4, we have $\frac{\partial}{\partial\sigma}x^*(s)|_{s=s_1} < 0$. Thus, $\frac{\partial}{\partial\sigma}s_1 > 0$, which implies s^* increases in σ in this case.

Case 2: We consider the scenario where $r \in [r_1,r_c)$. In this case, $s^* = \bar{s}$. By taking derivative of both sides of the equation $\bar{s} = 2a\sigma - 4b\mu\sigma x^*(\bar{s})$ (from Equation (A.11)) with respect to σ, we have

$$\frac{\partial}{\partial\sigma}\bar{s} = (2a-4b\mu x^*(\bar{s})) - 4b\mu\sigma\left(\frac{\partial}{\partial\sigma}x^*(s)\Big|_{s=\bar{s}} + \frac{\partial}{\partial s}x^*(s)\Big|_{s=\bar{s}}\frac{\partial}{\partial\sigma}\bar{s}\right),$$

which can be rewritten as

$$\left(1 + 4b\mu\sigma\frac{\partial}{\partial s}x^*(s)\Big|_{s=\bar{s}}\right)\frac{\partial}{\partial\sigma}\bar{s} = (2a-4b\mu x^*(\bar{s})) - 4b\mu\sigma\frac{\partial}{\partial\sigma}x^*(s)\Big|_{s=\bar{s}.}$$

From Proposition 1, we have $\frac{\partial}{\partial s}x^*(s)|_{s=\bar{s}} \ge 0$ when $r \ge r_1$. So $1 + 4b\mu\sigma\frac{\partial}{\partial s}x^*(s)|_{s=\bar{s}}$ is positive. Moreover, through the similar analysis as in Case 1, the right-hand side of the equation can be proven positive. Therefore, $\frac{\partial}{\partial\sigma}\bar{s} > 0$, which implies s^* increases in σ in this case.

Case 3: We consider the scenario where $r \ge r_c$. In this case, $s^* = \inf\{s : x^*(s) = x^{opt}\}$. We first know that $\frac{\partial}{\partial\sigma}x^*(s)|_{s=s^*} < \frac{\partial}{\partial\sigma}x^{opt}$ according to Lemma A.4. Moreover, as discussed in the proof of

Theorem 1, when $r \geq r_c$, $x^*(s)$ increases in s at $s = s^*$. Hence, s^* must increase in σ to maintain $s^* = \inf\{s : x^*(s) = x^{opt}\}$.

Collectively, in all three cases, we have s^* increases in σ. Thus, part (i) of the proposition holds.

(ii): As shown in the proof of Lemma A.4, $x^*(s)$ decreases in λ for any $s \in [0, \bar{s}]$. In addition, x^{opt} is independent of λ. Therefore, if we have $x^*(s^*) = x^{opt}$ under some λ, then we must have $x^*(s^*) = x^{opt}$ for all smaller values of λ. Let λ_c denote the largest λ such that $x^*(s^*) = x^{opt}$ (if we have $x^*(s^*) = x^{opt}$ for all feasible λ that satisfies our model assumptions stated in §3, we set $\lambda_c = \infty$; if we have $x^*(s^*) < x^{opt}$ for all feasible λ that satisfies our model assumptions stated in §3, we set λ_c to be any negative number). Then, we have $x^*(s^*) = x^{opt}$ if $\lambda \leq \lambda_c$ and $x^*(s^*) < x^{opt}$ if $\lambda > \lambda_c$.

We next separate the proof into two cases: In Case 1, $\lambda < \lambda_c$; and in Case 2, $\lambda \geq \lambda_c$.

Case 1: We consider the scenario where $\lambda < \lambda_c$. In this case, s^* must increase in λ to maintain $s^* = \inf\{s : x^*(s) = x^{opt}\}$ because we have shown $x^*(s)$ decreases in λ and $\frac{\partial}{\partial s} x^*(s)|_{s=s^*} \geq 0$ if $x^*(s^*) = x^{opt}$ in the proof of Theorem 1.

Case 2: We consider the scenario where $\lambda \geq \lambda_c$. Recall that $x^*(s) = \hat{x}(s)$ and from the proof of Proposition 1 that $\hat{s}$ is the subsidy amount such that $\frac{\partial}{\partial s}\hat{x}(s)|_{s=\hat{s}} = 0$ and $\hat{x}(s)$ increases in s if and only if $s \leq \hat{s}$. Next, we consider the following two cases:

(a): We consider the scenario where $\hat{s} \leq \bar{s}$ when $\lambda = \lambda_c$. Next, we prove that $\hat{s} \leq \bar{s}$ for any $\lambda \geq \lambda_c$. If this is true, then we must have $s^* = \hat{s}$ for any $\lambda \geq \lambda_c$, which implies that, to analyze how s^* changes in λ, it is sufficient to analyze how $\hat{s}$ changes in λ. Further, to prove $\hat{s} \leq \bar{s}$ for any $\lambda \geq \lambda_c$, because of the continuity of both $\hat{s}$ and $\bar{s}$, it is sufficient to prove that, for any $\lambda > \lambda_c$, whenever $\hat{s} = \bar{s}$, we have $\frac{\partial}{\partial \lambda}\hat{s} < 0$ and $\frac{\partial}{\partial \lambda}\bar{s} > 0$.

To prove $\frac{\partial}{\partial \lambda}\hat{s} < 0$ when $\hat{s} \leq \bar{s}$, it is sufficient to prove that $\frac{\partial^2}{\partial s \partial \lambda} x^*(s)|_{s=\hat{s}} < 0$ when $\hat{s} \leq \bar{s}$. Since $x^*(s) = \hat{x}(s)$, according to Lemma 2, $x^*(s)$ satisfies $u(x^*(s)c|x^*(s), s) = 0$. Then, by taking second order derivative of both sides of $u(x^*(s)c|x^*(s), s) = 0$ with respect to s and λ, we have

$$\left(\frac{\partial^2}{\partial x^2} u(xc|x,s)\Big|_{x=x^*(\hat{s}),s=\hat{s}} \frac{\partial}{\partial \lambda} x^*(s)\Big|_{s=\hat{s}} + \frac{\partial^2}{\partial x \partial \lambda} u(xc|x,s)\Big|_{x=x^*(\hat{s}),s=\hat{s}} \right) \frac{\partial}{\partial s} x^*(s)\Big|_{s=\hat{s}}$$
$$+ \frac{\partial}{\partial x} u(xc|x,s)\Big|_{x=x^*(\hat{s}),s=\hat{s}} \frac{\partial^2}{\partial s \partial \lambda} x^*(s)\Big|_{s=\hat{s}} + \frac{\partial^2}{\partial x \partial \lambda} u(xc|x,s)\Big|_{x=x^*(\hat{s}),s=\hat{s}} \frac{\partial}{\partial \lambda} x^*(s)\Big|_{s=\hat{s}}$$
$$+ \frac{\partial^2}{\partial s \partial \lambda} u(xc|x,s)\Big|_{x=x^*(\hat{s}),s=\hat{s}} = 0.$$

The first term vanishes because $\frac{\partial}{\partial s} x^*(s)|_{s=\hat{s}} = 0$. By the proof of Lemma A.4, we have that $\frac{\partial}{\partial x} u(xc|x,s)|_{x=x^*(\hat{s}),s=\hat{s}} < 0$, $\frac{\partial^2}{\partial x \partial \lambda} u(xc|x,s)|_{x=x^*(\hat{s}),s=\hat{s}} > 0$ and $\frac{\partial}{\partial \lambda} x^*(s)|_{s=\hat{s}} < 0$. We also have

$$\frac{\partial^2}{\partial s \partial \lambda} u(xc|x,s)\Big|_{x=x^*(\hat{s}),s=\hat{s}}$$

$$=\left(r-\frac{1}{2}\right)[R_H(x^*(\hat{s}))-R_L(x^*(\hat{s}))]-\frac{1}{2}\hat{s}$$
$$=\left(r-\frac{1}{2}\right)[R_H(x^*(\hat{s}))-R_L(x^*(\hat{s}))]-\frac{1}{2}\left(\frac{1}{\lambda}+2\left(r-\frac{1}{2}\right)[R_H(x^*(\hat{s}))-R_L(x^*(\hat{s}))]\right)$$
$$=-\frac{1}{2\lambda}<0,$$

where the second equality holds due to Equation (A.14) (by replacing s_1 with $\hat{s}$). As we plug these signs of the terms back, we have $\frac{\partial^2}{\partial s\partial\lambda}x^*(s)|_{s=\hat{s}}<0$ when $\hat{s}\le\bar{s}$. Thus, $\frac{\partial}{\partial\lambda}\hat{s}<0$ when $\hat{s}\le\bar{s}$.

To prove $\frac{\partial}{\partial\lambda}\bar{s}>0$ when $\hat{s}=\bar{s}$, by taking derivative of both sides of the equation $\bar{s}=2a\sigma-4b\mu\sigma x^*(\bar{s})$ (from Equation (A.11)) with respect to λ, we have

$$\frac{\partial}{\partial\lambda}\bar{s}=-4b\mu\sigma\left(\frac{\partial}{\partial\lambda}x^*(s)\Big|_{s=\bar{s}}+\frac{\partial}{\partial s}x^*(s)\Big|_{s=\bar{s}}\frac{\partial}{\partial\lambda}\bar{s}\right). \tag{A.20}$$

We already showed $\frac{\partial}{\partial\lambda}x^*(s)|_{s=\bar{s}}<0$, and $\frac{\partial}{\partial s}x^*(s)|_{s=\bar{s}}=0$ when $\hat{s}=\bar{s}$. Thus, $\frac{\partial}{\partial\lambda}\bar{s}>0$ when $\hat{s}=\bar{s}$.

Given that $\frac{\partial}{\partial\lambda}\hat{s}<0$ and $\frac{\partial}{\partial\lambda}\bar{s}>0$ whenever $\hat{s}=\bar{s}$, we have $\hat{s}\le\bar{s}$ for any $\lambda\ge\lambda_c$. Then, by the definition of s^*, we have $s^*=\hat{s}$ for any $\lambda\ge\lambda_c$. We have just shown that $\hat{s}$ decreases in λ. Therefore, s^* decreases in λ when $\lambda\ge\lambda_c$. In this case, we define $\lambda_1:=\lambda_c$.

(b): We consider the scenario where $\hat{s}>\bar{s}$ when $\lambda=\lambda_c$. Let λ_2 denote the lowest risk aversion level no less than λ_c under which $\hat{s}\le\bar{s}$ (if for all $\lambda\ge\lambda_c$, we have $\hat{s}>\bar{s}$, then we set $\lambda_2=\infty$).

Then, for $\lambda\in[\lambda_c,\lambda_2)$, we have $\hat{s}>\bar{s}$ and thus $s^*=\bar{s}$. Moreover, we have $\frac{\partial}{\partial s}x^*(s)|_{s=\bar{s}}\ge 0$ when $\hat{s}\ge\bar{s}$ and we already showed $\frac{\partial}{\partial\lambda}x^*(s)|_{s=\bar{s}}<0$. Thus, by Equation (A.20), we have s^* increases in λ when $\lambda\in[\lambda_c,\lambda_2)$. On the other hand, for $\lambda\ge\lambda_2$, since we already proved in (a) that, whenever $\hat{s}=\bar{s}$, we have $\frac{\partial}{\partial\lambda}\hat{s}<0$ and $\frac{\partial}{\partial\lambda}\bar{s}>0$, we must have $s^*=\hat{s}\le\bar{s}$ and s^* decreases in λ. In this case, we define $\lambda_1:=\lambda_2$.

Putting together Case 1 and Case 2, we conclude that s^* increases in λ if and only if $\lambda<\lambda_1$.

Finally, we show that λ_1 increases in the index accuracy r. On the one hand, if $\lambda_1=\lambda_c$, which is the largest risk aversion level under which $x^*(s^*)=x^{opt}$, λ_1 must increase in r because $x^*(s)$ increases in r by Corollary 1 and decreases in λ by the proof of Lemma A.4. On the other hand, if $\lambda_1=\lambda_2$, which is the lowest risk aversion level no less than λ_c under which $\hat{s}=\bar{s}$, then, by continuity of $\hat{s}$ and $\bar{s}$ in λ, we must have $\hat{s}=\bar{s}$ when $\lambda=\lambda_1$. In this case, λ_1 must also increase in r because we have shown in the Step 2 of the proof of Proposition 1 that whenever $\hat{s}=\bar{s}$, we have $\frac{\partial}{\partial r}\hat{s}>0$ and $\frac{\partial}{\partial r}\bar{s}\le 0$. □

Appendix A.2 Proofs of Analytical Results in §5

We have the following auxiliary lemma to present the closed-form expression of the equilibrium planting amount under both price and yield protection.

Lemma A.5 (Equilibrium planting amount with price and yield protection)**.** *Consider any given floor price $m \in [0, \bar{m}]$ and index-based yield protection subsidy $s \in [0, \bar{s}]$. Let $\hat{x}_m(s)$ denote the unique positive solution to $u_m(xc|x,s) = 0$. Then, the equilibrium planting amount is $x^*_m(s) = \min\{\hat{x}_m(s), 1\} \in (0,1]$. Moreover, there exist two thresholds m_1 and m_2 (which depend on s but not m) such that $0 < m_1 < m_2 \le \bar{m}$ and*
(i) if $m \in [0, m_1)$, then $\hat{x}_m(s) = \hat{x}(s)$, where $\hat{x}(s)$ is defined in Lemma 2;
(ii) if $m \in [m_1, m_2]$, then

$$\hat{x}_m(s) = \frac{-\beta_m(s) + \sqrt{\beta^2_m(s) - 4\alpha_m\gamma_m(s)}}{2\alpha_m},$$

where $\alpha_m = \frac{1}{4}\lambda b^2(\mu-\sigma)^4$, $\beta_m(s) = \frac{1}{2}b(\mu-\sigma)^2 + c + \frac{1}{2}\lambda b(\mu-\sigma)^2(m(\mu+\sigma) - a(\mu-\sigma)) - \lambda b(r - \frac{1}{2})s(\mu-\sigma)^2$, $\gamma_m(s) = \frac{1}{4}\lambda(m(\mu+\sigma) - a(\mu-\sigma))^2 - \frac{1}{2}m(\mu+\sigma) - \frac{1}{2}a(\mu-\sigma) - \lambda(r-\frac{1}{2})s(m(\mu+\sigma) - a(\mu-\sigma)) - \frac{1}{2}s + \lambda\frac{1}{4}s^2$;
(iii) if $m \in (m_2, \bar{m}]$, then

$$\hat{x}_m(s) = \frac{m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \lambda\frac{1}{4}s^2 + 2\lambda\sigma(r-\frac{1}{2})sm}{c}.$$

Further, there exists a threshold c_3 such that if $c \ge c_3$, then the equilibrium planting amount under given floor price $m \in [0, \bar{m}]$ and subsidy amount $s \in [0, \bar{s}]$ is $x^*_m(s) = \hat{x}_m(s) \in (0,1)$.

Proof of Lemma A.5: Recall that $u_m(h|x,s) = \mathbb{E}[\pi_m(h|x,s)] - \lambda\mathbb{V}\text{ar}[\pi_m(h|x,s)]$ denote the utility of the farmer whose production cost is h, and $\pi_m(h|x,s) = \max\{(a - bxY), m\}Y + sI - h$ denote the farmer's net income. Following the same logic as in the proof of Lemma 1, given $m \in [0, \bar{m}]$ and $s \in [0, \bar{s}]$, if $x = 0$, $x = 1$, or $x \in (0,1)$ is an equilibrium planting amount, then we must have $u_m(0|0,s) \le 0$, $u_m(c|1,s) \ge 0$ or $u_m(xc|x,s) = 0$ respectively.

Recall that a is defined as the maximum possible market-clearing price. We first prove the lemma by focusing on the more practical case with $m \le a$ in the following four steps (we discuss how our analysis can be extended to consider the case with $m > a$ at the end of the proof).

Step 1: We show that there is a unique positive solution to $u_m(xc|x,s) = 0$. To do so, it is sufficient to prove the following three parts for any given $s \in [0, 2a\sigma]$ (we prove this step for any $s \in [0, 2a\sigma]$, and at the end of Step 2, we show $\bar{s} \le 2a\sigma$): (1) $u_m(xc|x,s)$ strictly decreases in x for $x \ge 0$, (2) $u_m(0|0,s) > 0$ and (3) $u_m(xc|x,s) \le 0$ for some $x > 0$.

First, for (1), let $x_1 = \frac{a-m}{b(\mu+\sigma)}$ (i.e., $a - bx_1(\mu+\sigma) = m$) and $x_2 = \frac{a-m}{b(\mu-\sigma)}$ (i.e., $a - bx_2(\mu-\sigma) = m$). We prove that $u_m(xc|x,s)$ strictly decreases in x for $x \ge 0$ separately on three intervals of x: (a) $[0, x_1)$, (b) $[x_1, x_2]$, and (c) (x_2, ∞).

(a): For $x \in [0, x_1)$, by the definition of x_1, we have $a - bx(\mu-\sigma) > a - bx(\mu+\sigma) > m$. Thus, from the expression of $\pi_m(h|x,s)$ and $u_m(xc|x,s)$, we have that $u_m(xc|x,s) = u(xc|x,s)$, which is farmers' utility with only yield protection defined in §3. We already know from the proof of Lemma 2 that $u(xc|x,s)$ strictly decreases in x for $x \geq 0$. Therefore, $u_m(xc|x,s)$ strictly decreases in x for $x \in [0, x_1)$.

(b): For $x \in [x_1, x_2]$, by the definition of x_1 and x_2, we have $a - bx(\mu-\sigma) \geq m \geq a - bx(\mu+\sigma)$. Recall that $R_L(x) = (a - bx(\mu-\sigma))(\mu-\sigma)$ defined in §3 denotes the revenue with low yield. Then, we have

$$\begin{aligned}u_m(xc|x,s) =& \mathbb{E}[\max\{(a-bxY),m\}Y + sI - xc] - \lambda\mathbb{V}\text{ar}[\max\{(a-bxY),m\}Y + sI - xc]\\ =& \frac{1}{2}[m(\mu+\sigma) + R_L(x) + s] - xc\\ & - \lambda\left(\frac{1}{4}(m(\mu+\sigma) - R_L(x))^2 + \frac{1}{4}s^2 - s\left(r - \frac{1}{2}\right)(m(\mu+\sigma) - R_L(x))\right). \qquad \text{(A.21)}\end{aligned}$$

By taking derivative of $u_m(xc|x,s)$ with respect to x, we have, from Equation (A.21),

$$\frac{\partial}{\partial x}u_m(xc|x,s) = -\frac{1}{2}b(\mu-\sigma)^2 - c - \lambda b(\mu-\sigma)^2\left(\frac{1}{2}(m(\mu+\sigma) - R_L(x)) - s\left(r - \frac{1}{2}\right)\right).$$

For the sign of $\frac{\partial}{\partial x}u_m(xc|x,s)$, since we consider $r \leq 1$ and $s \leq 2a\sigma$, we have

$$\begin{aligned}\frac{1}{2}(m(\mu+\sigma) - R_L(x)) - s\left(r - \frac{1}{2}\right) &\geq \frac{1}{2}(m(\mu+\sigma) - R_L(x_1)) - a\sigma\\ &= \frac{1}{2}\left(m(\mu+\sigma) - a(\mu-\sigma) + \frac{(a-m)(\mu-\sigma)^2}{\mu+\sigma}\right) - a\sigma\\ &= (m-a)\frac{2\mu\sigma}{\mu+\sigma}\\ &\geq -a\frac{2\mu\sigma}{\mu+\sigma.}\end{aligned}$$

Moreover, since $c > c_1 \geq 2\lambda ab(\mu-\sigma)^2\frac{\mu\sigma}{\mu+\sigma} - \frac{1}{2}b(\mu-\sigma)^2$ (as discussed in §3), we have $\frac{\partial}{\partial x}u_m(xc|x,s) < 0$ and $u_m(xc|x,s)$ strictly decreases in x for $x \in [x_1, x_2]$.

(c): For $x > x_2$, by the definition of x_2, we have $m > a - bx(\mu-\sigma) > a - bx(\mu+\sigma)$. Then, we have

$$\begin{aligned}u_m(xc|x,s) &= \mathbb{E}[mY + sI - h] - \lambda\mathbb{V}\text{ar}[mY + sI - h]\\ &= m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \frac{1}{4}\lambda s^2 + 2\lambda\sigma\left(r - \frac{1}{2}\right)sm - xc, \qquad \text{(A.22)}\end{aligned}$$

and thus

$$\frac{\partial}{\partial x}u_m(xc|x,s) = -c < 0.$$

So, $u_m(xc|x,s)$ strictly decreases in x for $x > x_2$. Combining all three cases and considering that $u_m(xc|x,s)$ is continuous in x, we conclude that $u_m(xc|x,s)$ strictly decreases in x for all $x \geq 0$.

Second, for (2), we prove that $u_m(0|0,s) > 0$. When no farmer plants (i.e., $x = 0$), the market-clearing price is given by a for both the high- and low-yield cases. As we consider $m \leq a$, we must have $u_m(0|0,s) = u(0|0,s) = -\gamma(s) > 0$, where the last inequality is shown in the proof of Lemma 2.

Finally, for (3), we prove that $u_m(xc|x,s) < 0$ for some x. Since we have shown earlier that $u_m(xc|x,s)$ strictly decreases in x for $x \geq 0$ and is linear in x for $x > x_2$, there must exist an $x \geq 0$ such that $u_m(xc|x,s) < 0$.

Collectively, we conclude that there must exist a unique positive solution to $u_m(xc|x,s) = 0$, which we denote as $\hat{x}_m(s)$.

Step 2: We show that there is a unique equilibrium planting amount $x_m^*(s) = \min\{\hat{x}_m(s), 1\}$, where $\hat{x}_m(s)$ is the unique positive solution to $u_m(xc|x,s) = 0$. Since we have shown that $u_m(0|0,s) > 0$ in Step 1, $x = 0$ cannot be an equilibrium plating amount. Then, based on the discussion at the beginning of this proof, it remains to check $x = 1$ and $x = \hat{x}_m(s)$ if $\hat{x}_m(s) \in (0,1)$.

Consider any $s \in [0, \bar{s}]$. If $\hat{x}_m(s) \in (0,1)$, then, by the same logic as in Step 2 of the proof of Lemma 1, we have the equilibrium planting amount $x_m^*(s) = \hat{x}_m(s)$. If $\hat{x}_m(s) \geq 1$, then, as we just showed that $u_m(xc|x,s)$ strictly decreases in x, we must have $u_m(c|1,s) \geq 0$. In this case, we have the equilibrium planting amount $x_m^*(s) = 1$ because all the farmers would have positive utility from planting and would not deviate. Hence, there is a unique equilibrium planting amount given by $x_m^*(s) = \min\{\hat{x}_m(s), 1\}$.

Before we proceed to Step 3, we prove that $\bar{s} \leq 2a\sigma$. Recall that $\bar{s}$ is the lowest subsidy amount that satisfies

$$\max\{a - bx_m^*(s)(\mu - \sigma), m\}(\mu - \sigma) + s \geq \max\{a - bx_m^*(s)(\mu + \sigma), m\}(\mu + \sigma).$$

Thus, in order to prove $\bar{s} \leq 2a\sigma$, it is sufficient to show that, at $s = 2a\sigma$, we have $\max\{a - bx_m^*(s)(\mu - \sigma), m\}(\mu - \sigma) + s \geq \max\{a - bx_m^*(s)(\mu + \sigma), m\}(\mu + \sigma)$. Consider the following three cases: First, if $a - bx^*(2a\sigma)(\mu - \sigma) > a - bx^*(2a\sigma)(\mu + \sigma) > m$, then we have

$$\begin{aligned}
\max\{a - bx_m^*(2a\sigma)(\mu - \sigma), m\}(\mu - \sigma) + 2a\sigma &= (a - bx_m^*(2a\sigma)(\mu - \sigma))(\mu - \sigma) + 2a\sigma \\
&= a(\mu + \sigma) - bx_m^*(2a\sigma)(\mu - \sigma)^2 \\
&> a(\mu + \sigma) - bx_m^*(2a\sigma)(\mu + \sigma)^2 \\
&= (a - bx_m^*(2a\sigma)(\mu + \sigma))(\mu + \sigma) \\
&= \max\{a - bx_m^*(2a\sigma)(\mu + \sigma), m\}(\mu + \sigma).
\end{aligned}$$

Second, if $a - bx^*(2a\sigma)(\mu - \sigma) \geq m \geq a - bx^*(2a\sigma)(\mu + \sigma)$, then we have

$$\begin{aligned}
\max\{a - bx_m^*(2a\sigma)(\mu - \sigma), m\}(\mu - \sigma) + 2a\sigma &= (a - bx_m^*(2a\sigma)(\mu - \sigma))(\mu - \sigma) + 2a\sigma \\
&> a(\mu + \sigma) \\
&\geq m(\mu + \sigma) \\
&= \max\{a - bx_m^*(2a\sigma)(\mu + \sigma), m\}(\mu + \sigma).
\end{aligned}$$

Third, if $m > a - bx^*(2a\sigma)(\mu-\sigma) > a - bx^*(2a\sigma)(\mu+\sigma)$, then we have

$$\begin{aligned}\max\{a - bx_m^*(2a\sigma)(\mu-\sigma), m\}(\mu-\sigma) + 2a\sigma &= m(\mu-\sigma) + 2a\sigma \\ &\geq m(\mu-\sigma) + 2m\sigma \\ &= m(\mu+\sigma) \\ &= \max\{a - bx_m^*(2a\sigma)(\mu+\sigma), m\}(\mu+\sigma).\end{aligned}$$

Combining all three cases, we conclude that, at $s = 2a\sigma$, we have $\max\{a - bx_m^*(s)(\mu-\sigma), m\}(\mu-\sigma) + s \geq \max\{a - bx_m^*(s)(\mu+\sigma), m\}(\mu+\sigma)$, which implies $\bar{s} \leq 2a\sigma$.

Step 3: We derive the expression of $\hat{x}_m(s)$. Consider the following three cases:

(a): We consider the scenario where $\hat{x}_m(s) < x_1$. Then, as shown earlier, we have $u_m(xc|x,s) = u(xc|x,s)$ at $x = \hat{x}_m(s)$ and thus $\hat{x}_m(s) = \hat{x}(s)$ defined in Lemma 2. Therefore, we conclude that whenever the unique positive solution $\hat{x}_m(s)$ satisfies $\hat{x}_m(s) < x_1$, we have that $\hat{x}_m(s) = \hat{x}(s)$.

(b): We consider the scenario where $\hat{x}_m(s) \in [x_1, x_2]$. Then, based on Equation (A.21), $\hat{x}_m(s)$ must be the solution to the following equation

$$\begin{aligned}u_m(xc|x,s) =& \frac{1}{2}[m(\mu+\sigma) + R_L(x) + s] - xc \\ &- \lambda\left(\frac{1}{4}(m(\mu+\sigma) - R_L(x))^2 + \frac{1}{4}s^2 - s\left(r - \frac{1}{2}\right)(m(\mu+\sigma) - R_L(x))\right) \\ =& -\alpha_m x^2 - \beta_m(s)x - \gamma_m(s) \\ =& 0,\end{aligned}$$

where $\alpha_m = \frac{1}{4}\lambda b^2(\mu-\sigma)^4$, $\beta_m(s) = \frac{1}{2}b(\mu-\sigma)^2 + c + \frac{1}{2}\lambda b(\mu-\sigma)^2(m(\mu+\sigma) - a(\mu-\sigma)) - \lambda b(r - \frac{1}{2})s(\mu-\sigma)^2$ and $\gamma_m(s) = \frac{1}{4}\lambda(m(\mu+\sigma) - a(\mu-\sigma))^2 - \frac{1}{2}m(\mu+\sigma) - \frac{1}{2}a(\mu-\sigma) - \lambda(r-\frac{1}{2})s(m(\mu+\sigma) - a(\mu-\sigma)) - \frac{1}{2}s + \lambda\frac{1}{4}s^2$. Then, we have[9]

$$\hat{x}_m(s) = \frac{-\beta_m(s) + \sqrt{\beta_m^2(s) - 4\alpha_m\gamma_m(s)}}{2\alpha_m}. \tag{A.23}$$

Therefore, we conclude that whenever the unique positive solution $\hat{x}_m(s)$ satisfies $\hat{x}_m(s) \in [x_1, x_2]$, we have that $\hat{x}_m(s) = \frac{-\beta_m(s) + \sqrt{\beta_m^2(s) - 4\alpha_m\gamma_m(s)}}{2\alpha_m}$.

(c): We consider the scenario where $\hat{x}_m(s) > x_2$. Then, based on Equation (A.22), $\hat{x}_m(s)$ must be the solution to the following equation

$$\mathbb{E}[mY + sI - h] - \lambda\mathbb{V}\text{ar}[mY + sI - h] = 0.$$

Then, we have

$$\hat{x}_m(s) = \frac{m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \frac{1}{4}\lambda s^2 + 2\lambda\sigma(r - \frac{1}{2})sm}{c}. \tag{A.24}$$

[9] The other solution cannot be $\hat{x}_m(s)$ because we have shown that $u_m(xc|x,s)$ decreases in x.

Therefore, we conclude that whenever the unique positive solution $\hat{x}_m(s)$ satisfies $\hat{x}_m(s) > x_2$, we have that $\hat{x}_m(s) = (m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \frac{1}{4}\lambda s^2 + 2\lambda\sigma(r - \frac{1}{2})sm)/c$.

Step 4: We prove that for any given $s \in [0, 2a\sigma]$ (recall from our proof in Step 1 that we must have $\bar{s} \le 2a\sigma$), there exist two thresholds $0 < m_1 < m_2 \le a$ such that if $m < m_1$, then $\hat{x}_m(s) < x_1$; if $m_1 \le m \le m_2$, then $x_1 \le \hat{x}_m(s) \le x_2$; and if $m_2 < m \le a$, then $\hat{x}_m(s) > x_2$.

Consider a fixed $s \in [0, 2a\sigma]$. Since $x_1 < x_2$, in order to prove the existence such m_1 and m_2, it is sufficient to prove that (1) if $\hat{x}_m(s) < x_1$ for some floor price m, then $\hat{x}_m(s) < x_1$ for any lower floor price and (2) if $\hat{x}_m(s) > x_2$ for some floor price m, then $\hat{x}_m(s) > x_2$ for any higher floor price.

For (1), first, we prove that $\hat{x}_m(s) < x_1$ if and only if $\hat{x}(s) < x_1$. By Step 3, we have shown that if $\hat{x}_m(s) < x_1$, then $\hat{x}_m(s) = \hat{x}(s)$. Thus, if $\hat{x}_m(s) < x_1$, then we have $\hat{x}(s) = \hat{x}_m(s) < x_1$. On the other hand, if $\hat{x}(s) < x_1$, then we have

$$\begin{aligned}
u_m(\hat{x}(s)c|\hat{x}(s), s) =& \mathbb{E}[\max\{(a - b\hat{x}(s)Y), m\}Y + sI - \hat{x}(s)c] \\
& - \lambda \mathbb{V}\mathrm{ar}[\max\{(a - b\hat{x}(s)Y), m\}Y + sI - \hat{x}(s)c] \\
=& \mathbb{E}[(a - b\hat{x}(s)Y)Y + sI - \hat{x}(s)c] - \lambda \mathbb{V}\mathrm{ar}[(a - b\hat{x}(s)Y)Y + sI - \hat{x}(s)c] \\
=& u(\hat{x}(s)c|\hat{x}(s), s) \\
=& 0,
\end{aligned}$$

where the second equality holds because, by the definition of x_1, we have $a - b\hat{x}(s)(\mu - \sigma) > a - b\hat{x}(s)(\mu + \sigma) > m$ and the third and the fourth equality hold because of the definitions of $u(xc|x, s)$ and $\hat{x}(s)$ in Lemma 2. Thus, $\hat{x}(s)$ is a solution to the equation $u_m(xc|x, s) = 0$. Moreover, we have shown in Step 1 that $\hat{x}_m(s)$ is the unique positive solution to $u_m(xc|x, s) = 0$. Hence, we have $\hat{x}_m(s) = \hat{x}(s)$, which implies $\hat{x}_m(s) < x_1$. Collectively, we have $\hat{x}_m(s) < x_1$ if and only if $\hat{x}(s) < x_1$.

Then, in order to prove (1), it is sufficient to prove that if $\hat{x}(s) < x_1$ for some floor price m, then $\hat{x}(s) < x_1$ for any lower floor price. It can be proven since we already know that x_1 strictly decreases in m and that $\hat{x}(s)$ is independent of m.

For (2), let $x'_m(s) = (m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \frac{1}{4}\lambda s^2 + 2\lambda\sigma(r - \frac{1}{2})sm)/c$. First, we prove that $\hat{x}_m(s) > x_2$ if and only if $x'_m(s) > x_2$. By Step 3, we have shown that if $\hat{x}_m(s) > x_2$, then $\hat{x}_m(s) = x'_m(s)$. Thus, if $\hat{x}_m(s) > x_2$, then we have $x'_m(s) = \hat{x}_m(s) > x_2$. On the other hand, if $x'_m(s) > x_2$, then we have

$$\begin{aligned}
u_m(x'_m(s)c|x'_m(s), s) =& \mathbb{E}[\max\{(a - bx'_m(s)Y), m\}Y + sI - x'_m(s)c] \\
& - \lambda \mathbb{V}\mathrm{ar}[\max\{(a - bx'_m(s)Y), m\}Y + sI - x'_m(s)c] \\
=& \mathbb{E}[mY + sI - x'_m(s)c] - \lambda \mathbb{V}\mathrm{ar}[mY + sI - x'_m(s)c] \\
=& 0,
\end{aligned}$$

where the second equality holds because, by the definition of x_2, we have $m > a - bx'_m(s)(\mu - \sigma) > a - bx'_m(s)(\mu + \sigma)$ and the third equality has been shown in (c) of Step 3. Thus, $x'_m(s)$ is a solution to the equation $u_m(xc|x,s) = 0$. Moreover, we have shown in Step 1 that $\hat{x}_m(s)$ is the unique positive solution to $u_m(xc|x,s) = 0$. Hence, we have $\hat{x}_m(s) = x'_m(s)$, which implies $x'_m(s) > x_2$. Collectively, we have $x'_m(s) > x_2$ if and only if $\hat{x}_m(s) > x_2$.

Then, in order to prove (2), it is sufficient to prove that if $x'_m(s) > x_2$ for some floor price m, then $x'_m(s) > x_2$ for any higher floor price. That is, it is sufficient to prove that if $x'_m(s) - x_2 > 0$ for some floor price m, then $x'_m(s) - x_2 > 0$ for any higher floor price. Since x_2 is linear in m and $x'_m(s)$ is a quadratic concave function in m, $x'_m(s) - x_2$ must be a quadratic concave function in m. Recall that we consider $m \le a$. Then, it is sufficient to prove that when $m = a$, we have $x'_m(s) - x_2 = x'_m(s) > 0$. Since $s \in [0, 2a\sigma]$ and $x'_m(s)$ is a quadratic concave function of s, it is sufficient to prove when $m = a$, we have $x'_m(0) > 0$ and $x'_m(2a\sigma) > 0$. When $m = a$, we have

$$x'_m(0) = \frac{a\mu - \lambda a^2\sigma^2}{c} > \frac{1}{c}\left(a\mu - \frac{\mu+\sigma}{2a\sigma^2}a^2\sigma^2\right) = \frac{1}{c}\left(a\mu - a\frac{\mu+\sigma}{2}\right) \ge 0,$$

where the first inequality holds because we consider $\lambda < \frac{\mu+\sigma}{2a\sigma^2}$ and the second inequality holds because we consider $\mu \ge \sigma$; we also have

$$\begin{aligned} x'_m(2a\sigma) &= \frac{1}{c}\left(a\mu - \lambda a^2\sigma^2 + a\sigma - \lambda a^2\sigma^2 + 4\lambda\sigma^2\left(r - \frac{1}{2}\right)a^2\right) \\ &\ge \frac{1}{c}\left(a(\mu+\sigma) - 2\lambda a^2\sigma^2\right) \\ &> \frac{1}{c}\left(a(\mu+\sigma) - 2\left(\frac{\mu+\sigma}{2a\sigma^2}\right)a^2\sigma^2\right) \\ &= 0, \end{aligned}$$

where the first inequality holds because $r \ge \frac{1}{2}$. Hence, we have $x'_m(s) > x_2$ when $m = a$. We conclude that if $\hat{x}_m(s) > x_2$ for some floor price m, then $\hat{x}_m(s) > x_2$ for any higher floor price.

Combining all four steps completes the proof when $m \le a$.

We investigate the case where $m \in (a, \bar{m}]$ (this case does not exist if $\bar{m} \le a$). In this case, by the definition of $\bar{s}$, we have $\bar{s} = \max\{a - bx^*_m(s)(\mu+\sigma), m\}(\mu+\sigma) - \max\{a - bx^*_m(s)(\mu-\sigma), m\}(\mu-\sigma) = 2m\sigma$. Then, for any given $s \in [0, 2m\sigma]$ and $m \in (a, \bar{m}]$, we derive the equilibrium planting amount and its closed-form expression in the following two steps:

Step 1: We show that there is a unique positive solution to $u_m(xc|x,s) = 0$. To do so, it is sufficient to prove the following three parts for any given $m \in (a, \frac{\mu}{2\lambda\sigma^2}]$ (we prove this step for any $m \in (a, \frac{\mu}{2\lambda\sigma^2}]$, and at the end of Step 2, we show $\bar{m} \le \frac{\mu}{2\lambda\sigma^2}$): (1) $u_m(xc|x,s)$ strictly decreases in x for $x \ge 0$, (2) $u_m(0|0,s) > 0$ and (3) $u_m(xc|x,s) \le 0$ for some $x > 0$.

First, for part (1), since $m > a$ and a is the maximum possible market-clearing price, we have

$$\begin{aligned} u_m(xc|x,s) &= \mathbb{E}[mY + sI - h] - \lambda \mathbb{V}\text{ar}[mY + sI - h] \\ &= m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \frac{1}{4}\lambda s^2 + 2\lambda\sigma\left(r - \frac{1}{2}\right)sm - xc, \end{aligned}$$

and thus

$$\frac{\partial}{\partial x}u_m(xc|x,s) = -c < 0.$$

So, $u_m(xc|x,s)$ strictly decreases in x.

Second, for part (2), since $u_m(xc|x,s)$ is a quadratic concave function in s, in order to prove $u_m(0|0,s) > 0$ for any $s \in [0, 2m\sigma]$, it is sufficient to prove that $u_m(0|0,0) > 0$ and $u_m(0|0,2m\sigma) > 0$. As we consider $m \in (a, \frac{\mu}{2\lambda\sigma^2}]$, we have

$$u_m(0|0,0) = m\mu - \lambda m^2\sigma^2 \geq m\mu - \frac{m\mu}{2} = \frac{m\mu}{2} > \frac{a\mu}{2} > 0$$

and

$$\begin{aligned} u_m(0|0,2m\sigma) &= m\mu - \lambda m^2\sigma^2 + m\sigma - \lambda m^2\sigma^2 + 4\lambda\sigma^2\left(r - \frac{1}{2}\right)m^2 \\ &\geq m(\mu+\sigma) - 2\lambda m^2\sigma^2 \\ &\geq m(\mu+\sigma) - m\mu \\ &> 0, \end{aligned}$$

where the first inequality holds because $r \geq \frac{1}{2}$.

Finally, for part (3), since we have shown that $u_m(xc|x,s)$ linearly decreases in x, there must exist an $x > 0$ such that $u_m(xc|x,s) \leq 0$.

Collectively, we conclude that there must exist a unique positive solution to $u_m(xc|x,s) = 0$, which we denote as $\hat{x}_m(s)$. By Equation (A.24), we have

$$\hat{x}_m(s) = \frac{m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \frac{1}{4}\lambda s^2 + 2\lambda\sigma(r - \frac{1}{2})sm}{c}. \tag{A.25}$$

Step 2: Similar to the Step 2 when $m \leq a$, we have that a unique equilibrium planting amount $x_m^*(s) = \min\{\hat{x}_m(s), 1\}$.

Finally, we prove that $\bar{m} \leq \frac{\mu}{2\lambda\sigma^2}$. Consider any fixed $m \in (a, \bar{m}]$. Recall that $\bar{m}$ is defined as the smallest m such that either $x_m^*(0) \geq x^{opt}$ or $\frac{\partial}{\partial m}x_m^*(0) \geq 0$. Then, since $x_m^*(s) = \min\{\hat{x}_m(s), 1\}$, we must have $\frac{\partial}{\partial m}\hat{x}_m(0) \geq 0$. By Equation (A.25), we have

$$\frac{\partial}{\partial m}\hat{x}_m(0) = \frac{\partial}{\partial m}(m\mu - \lambda m^2\sigma^2) = \mu - 2\lambda m\sigma^2 \geq 0,$$

which implies that $m \leq \frac{\mu}{2\lambda\sigma^2}$. Since it is true for any $m \in [0, \bar{m}]$, we must have $\bar{m} \leq \frac{\mu}{2\lambda\sigma^2}$.

Combining Step 1 and Step 2 completes the proof when $m \in (a, \bar{m}]$.

Finally, we prove that there exists a threshold c_3 such that if $c \geq c_3$, then the equilibrium planting amount under given floor price $m \in [0, \bar{m}]$ and subsidy amount $s \in [0, \bar{s}]$ is $x_m^*(s) = \hat{x}_m(s) \in (0,1)$. To do so, it is sufficient to prove that there exists a threshold c_3 such that if $c \geq c_3$, then $u_m(c|1,s) < 1$ for any $m \in [0, \bar{m}]$ and $s \in [0, \bar{s}]$. Consider a fixed $m \in [0, \bar{m}]$ and a fixed $s \in [0, \bar{s}]$. We have

$$u_m(c|1,s) = \mathbb{E}[\max\{(a - bY), m\}Y + sI] - c - \lambda \mathbb{V}\text{ar}[\max\{(a - bY), m\}Y + sI].$$

From its expression, we have that $u_m(c|1,s)$ linearly decreases in c and that $u_m(c|1,s)$ is continuous in both m and s. Moreover, when $\bar{m} \leq a$, we have shown that $\bar{s} \leq 2a\sigma$; when $\bar{m} > a$, we have shown that $\bar{m} \leq \frac{\mu}{2\lambda\sigma^2}$ and $\bar{s} = 2m\sigma \leq 2\bar{m}\sigma \leq \frac{\mu}{\lambda\sigma}$. Collectively, it implies that $\bar{m} \leq \max\{a, \frac{\mu}{2\lambda\sigma^2}\}$ and $\bar{s} \leq \max\{2a\sigma, \frac{\mu}{\lambda\sigma}\}$. Therefore, there must exist a threshold c_3 such that if $c \geq c_3$, then $u_m(c|1,s) < 1$ for any $m \in [0, \bar{m}]$ and $s \in [0, \bar{s}]$. □

Before we move on to the proof of Proposition 4, we prove an auxiliary lemma to study the relationship between the equilibrium planting amount $x_m^*(s)$ and the index accuracy r.

Lemma A.6. *If $x_m^*(s) = \hat{x}_m(s)$, where $\hat{x}_m(s)$ is defined in Lemma A.5, then $x_m^*(s)$ increases in the index accuracy r.*

Proof of Lemma A.6: Consider a fixed $m \in [0, \bar{m}]$ and a fixed $s \in [0, \bar{s}]$. Then, given s, according to Lemma A.5, there exists two thresholds m_1 and m_2. Accordingly, we separate the analysis into three cases: In Case 1, $m < m_1$; in Case 2, $m \in [m_1, m_2]$; and in Case 3, $m > m_2$.

Case 1: We study the scenario where $m < m_1$. According to Lemma A.5, in this case, $x_m^*(s) = x^*(s)$. By Corollary 1, we have already proven that $x^*(s)$ increases in r. Therefore, $x_m^*(s)$ increases in r.

Case 2: We study the scenario where $m \in [m_1, m_2]$. Since we consider $x_m^*(s) = \hat{x}_m(s)$, in order to prove $x_m^*(s)$ increases in r, we only need to prove $\hat{x}_m(s)$ increases in r.

Following the same logic as in the proof of Lemma A.1, when $m \in [m_1, m_2]$, $\frac{\partial}{\partial r}\hat{x}_m(s)$ and $\frac{\partial}{\partial r}u_m(xc|x,s)|_{x=x_m^*(s)}$ have the same sign. Then, from Equation (A.21), we have

$$\frac{\partial}{\partial r}u_m(xc|x,s)\Big|_{x=x_m^*(s)} = s[m(\mu + \sigma) - R_L(x_m^*(s))] \geq 0,$$

where the inequality holds because $\bar{s}$ is defined as the lowest subsidy amount that satisfies $R_L(x_m^*(s)) + s \geq R_H(x_m^*(s))$ and thus for any $s \in [0, \bar{s}]$, we have $R_H(x_m^*(s)) - R_L(x_m^*(s)) = m(\mu + \sigma) - R_L(x_m^*(s)) \geq s \geq 0$. Therefore, $\hat{x}_m(s)$ increases in r, which implies $x_m^*(s)$ increases in r.

Case 3: We study the scenario where $m > m_2$. By Lemma A.5, we have

$$\hat{x}_m(s) = \frac{m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \frac{1}{4}\lambda s^2 + 2\lambda\sigma(r - \frac{1}{2})sm}{c},$$

and we have

$$\frac{\partial}{\partial r}\hat{x}_m(s) = \frac{2\lambda\sigma sm}{c} \geq 0.$$

Since we consider $x_m^*(s) = \hat{x}_m(s)$, $x_m^*(s)$ increases in r.

Combining all three cases, we conclude that $x_m^*(s)$ increases in r if $x_m^*(s) = \hat{x}_m(s)$,. □

Proof of Proposition 4: We prove this proposition in the following four steps:

Step 1: We prove that given $m \in [0, \bar{m}]$ and $r \in [\frac{1}{2}, 1]$, there exists a threshold $s_{m,1} > 0$ such that the equilibrium planting amount $x_m^*(s)$ increases in s if and only if $s \leq s_{m,1}$. First, as a building block for further analysis, we prove that $x_m^*(s)$ is continuous in s on $[0, \bar{s}]$. Since we consider $x_m^*(s) = \hat{x}_m(s)$ for any $s \in [0, \bar{s}]$, where $\hat{x}_m(s)$ is the unique positive solution to $u_m(xc|x, s) = 0$, it is sufficient to prove that given x, $u_m(xc|x, s)$ is continuous in s. As defined in the proof of Lemma A.5, we have

$$\begin{aligned} u_m(xc|x,s) =& \mathbb{E}[\max\{(a - bxY), m\}Y + sI - xc] - \lambda \mathbb{V}\text{ar}[\max\{(a - bxY), m\}Y + sI - xc] \\ =& \mathbb{E}[\max\{(a - bxY), m\}Y - xc] + \frac{1}{2}s - \lambda \mathbb{V}\text{ar}[\max\{(a - bxY), m\}Y] - \frac{1}{4}\lambda s^2 \\ & - 2\lambda \text{Cov}(\max\{(a - bxY), m\}Y, I)s, \end{aligned}$$

which is continuous in s. Hence, we conclude that $x_m^*(s)$ is continuous in s on $[0, \bar{s}]$.

Then, given $m \in [0, \bar{m}]$ and $r \in [\frac{1}{2}, 1]$, we define $s_{m,1} := \sup\{s : s = \arg\max_{s \in [0,\bar{s}]} x_m^*(s)\}$. To prove the conclusion in Step 1, it is sufficient to separately prove that $x_m^*(s)$ increases in s for $s \in [0, s_{m,1}]$ and $x_m^*(s)$ decreases in s for $s \in (s_{m,1}, \bar{s}]$ (if $s_{m,1} = \bar{s}$, then $(s_{m,1}, \bar{s}]$ is empty).

We first prove that $x_m^*(s)$ increases in s for $s \in [0, s_{m,1}]$ by contradiction. Suppose there exists a subsidy amount $s' \in [0, s_{m,1}]$ such that $x_m^*(s)$ strictly decreases in s at $s = s'$. Recall that in the proof of Lemma A.5, we define $x_1 = \frac{a-m}{b(\mu+\sigma)}$ (i.e., $a - bx_1(\mu + \sigma) = m$) and $x_2 = \frac{a-m}{b(\mu-\sigma)}$ (i.e., $a - bx_2(\mu - \sigma) = m$). If $x_m^*(s') < x_1$, then, by the definition of $s_{m,1}$ and the continuity of $x_m^*(s)$ in s, there must exist an $s'' \in (s', s_{m,1})$ such that $x_m^*(s'') < x_1$ and $x_m^*(s)$ strictly increases in s at $s = s''$. However, as implied by the discussion in Step 3 of the proof of Lemma A.5, whenever $x_m^*(s) < x_1$, we have $x_m^*(s) = x^*(s)$, which is strictly concave in s as proven in Lemma A.2. Therefore, the existence of such s' and s'' contradicts with the concavity of $x_m^*(s)$.

Further, if $x_m^*(s') \in [x_1, x_2]$ or $x_m^*(s') > x_2$, through the same logic, in order to show that the existence of such s' leads to contradiction, we only need to show that $x_m^*(s)$ is concave whenever $x_m^*(s) \in [x_1, x_2]$ or $x_m^*(s) > x_2$.

As implied by the discussion in Step 3 of the proof of Lemma A.5, whenever $x_m^*(s) \in [x_1, x_2]$, we have $x_m^*(s) = \frac{-\beta_m(s) + \sqrt{\beta_m^2(s) - 4\alpha_m\gamma_m(s)}}{2\alpha_m}$. Moreover, since α_m is independent of s, $\beta_m(s)$ is linear in s and $\beta_m^2(s) - 4\alpha_m\gamma_m(s)$ can be written as a quadratic concave function in s, we have that

$\frac{-\beta_m(s)+\sqrt{\beta_m^2(s)-4\alpha_m\gamma_m(s)}}{2\alpha_m}$ is a strictly concave function in s. Therefore, whenever $x_m^*(s) \in [x_1, x_2]$, $x_m^*(s)$ is concave in s. On the other hand, whenever $x_m^*(s) > x_2$, we have $x_m^*(s) = (m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \frac{1}{4}\lambda s^2 + 2\lambda\sigma(r-\frac{1}{2})sm)/c$, which is a quadratic concave function in s. Therefore, whenever $x_m^*(s) > x_2$, $x_m^*(s)$ is concave in s. Collectively, we conclude that such s' does not exist and thus $x_m^*(s)$ increases in s for $s \in [0, s_{m,1}]$.

Next, we prove that $x_m^*(s)$ decreases in s for $s \in (s_{m,1}, \bar{s}]$ by contradiction. Suppose there exists an $s' \in (s_{m,1}, \bar{s}]$ such that $x_m^*(s)$ strictly increases in s at $s = s'$. If $x_m^*(s') < x_1$, then, by the definition of $s_{m,1}$ and the continuity of $x_m^*(s)$ in s, there must exist an $s'' \in (s_{m,1}, s')$ such that $x_m^*(s'') < x_1$ and $x_m^*(s)$ strictly decreases in s at $s = s''$, which contradicts with the fact that $x_m^*(s) = x^*(s)$ is concave in s. If $x_m^*(s') \in [x_1, x_2]$ or $x_m^*(s') > x_2$, through the same logic, we can show that the existence of such s' leads to the contradiction with the concavity of $\frac{-\beta_m(s)+\sqrt{\beta_m^2(s)-4\alpha_m\gamma_m(s)}}{2\alpha_m}$ and $(m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \frac{1}{4}\lambda s^2 + 2\lambda\sigma(r-\frac{1}{2})sm)/c$ respectively. Collectively, we conclude that $x_m^*(s)$ increases in s if and only if $s \leq s_{m,1}$.

Furthermore, we prove that $s_{m,1} = \sup\{s : s = \arg\max_{s\in[0,\bar{s}]} x_m^*(s)\} = \arg\max_{s\in[0,\bar{s}]} x_m^*(s)$. First, we have proven that $x_m^*(s)$ increases in s if and only if $s \leq s_{m,1}$. Moreover, based on the strict concavity proven earlier, we have $\frac{\partial^2}{\partial s^2} x_m^*(s)|_{s=s_{m,1}} < 0$ if $x_m^*(s_{m,1}) < x_1$, $x_m^*(s_{m,1}) \in [x_1, x_2]$ or $x_m^*(s_{m,1}) > x_2$. Thus, $x_m^*(s)$ has a unique maximizer, which implies that $s_{m,1} = \arg\max_{s\in[0,\bar{s}]} x_m^*(s)$.

Finally, we prove that $s_{m,1} > 0$. By the definition of $s_{m,1}$, it is sufficient to prove that $\frac{\partial}{\partial s} x_m^*(s)|_{s=0} > 0$. If $x_m^*(0) < x_1$, then $x_m^*(0) = x^*(0)$ and we have shown in Lemma A.2 that $\frac{\partial}{\partial s} x^*(s)|_{s=0} > 0$. Thus, $\frac{\partial}{\partial s} x_m^*(s)|_{s=0} > 0$. If $x_m^*(0) \in [x_1, x_2]$, then, following the same logic as in the proof of Lemma A.1, we have

$$\frac{\partial}{\partial s} x_m^*(s)\Big|_{s=0} = \frac{\partial}{\partial s} u_m(xc|x,s)\Big|_{s=0,x=x_m^*(0)} = \frac{1}{2} + \lambda\left(r - \frac{1}{2}\right)(m(\mu+\sigma) - R_L(x_m^*(0))) > 0,$$

where the second equality is based on Equation (A.21) and the inequality holds because we consider $r \in [\frac{1}{2}, 1]$ and we have shown in the proof of Lemma A.6 that $(m(\mu+\sigma) - R_L(x_m^*(0))) \geq s \geq 0$. If $x_m^*(0) > x_2$, then we have, from Equation (A.22),

$$\frac{\partial}{\partial s} x_m^*(s)\Big|_{s=0} = \frac{\frac{1}{2} + 2\lambda\sigma(r-\frac{1}{2})m}{c} > 0.$$

Hence, combining all three cases, we have that $\frac{\partial}{\partial s} x_m^*(s)|_{s=0} > 0$ and thus $s_{m,1} > 0$.

Step 2: We prove that given $m \in [0, \bar{m}]$ and $r \in [\frac{1}{2}, 1]$, if $s_{m,1} < \bar{s}$, then $s_{m,1}$ increases in both r and m. Similar to Step 1, we separate the further proof into three cases: In Case 1, $x_m^*(s_{m,1}) < x_1$; in Case 2, $x_m^*(s_{m,1}) \in [x_1, x_2]$; and in Case 3, $x_m^*(s_{m,1}) > x_2$.

Case 1: We consider the scenario where $x_m^*(s_{m,1}) < x_1$. Then, by the definition of $s_{m,1}$, we have $x_m^*(s) < x_1$ and thus $x_m^*(s) = x^*(s)$ for all $s \in [0, \bar{s}]$. Thus, when $s_{m,1} < \bar{s}$, we have $s_{m,1} = s_1$ defined in Proposition 1. Besides, s_1 increases in r shown in Proposition 1. Hence, $s_{m,1}$ increases in r if $s_{m,1} < \bar{s}$. Further, $s_{m,1}$ is independent of m in this case.

Case 2: We consider the scenario where $x_m^*(s_{m,1}) \in [x_1, x_2]$. According to Lemma A.5, we have $x_m^*(s_{m,1}) = \frac{-\beta_m(s_{m,1}) + \sqrt{\beta_m^2(s_{m,1}) - 4\alpha_m\gamma_m(s_{m,1})}}{2\alpha_m}$. We first prove that if $s_{m,1} < \bar{s}$, then $s_{m,1}$ increases in r. If $s_{m,1} < \bar{s}$, then, since we have shown in Step 1 that $\frac{-\beta_m(s) + \sqrt{\beta_m^2(s) - 4\alpha_m\gamma_m(s)}}{2\alpha_m}$ is strictly concave in s, we have $\frac{\partial}{\partial s} x_m^*(s)|_{s=s_{m,1}} = 0$ and, following the same logic as in the proof of Lemma A.1,

$$\frac{\partial}{\partial s} u_m(xc|x,s)\Big|_{x=x_m^*(s), s=s_{m,1}} = \frac{\partial}{\partial s} x_m^*(s)\Big|_{s=s_{m,1}} = 0.$$

Then, from Equation (A.21), it implies

$$s_{m,1} = \frac{1}{\lambda} + 2(r - \frac{1}{2})[m(\mu + \sigma) - R_L(x_m^*(s_{m,1}))]. \tag{A.26}$$

Taking derivative of $s_{m,1}$ with respect to r, we have

$$\begin{aligned}
\frac{\partial}{\partial r} s_{m,1} &= 2(r - \frac{1}{2}) b(\mu - \sigma)^2 \left(\frac{\partial}{\partial r} x_m^*(s)\Big|_{s=s_{m,1}} + \frac{\partial}{\partial s} x_m^*(s)\Big|_{s=s_{m,1}} \frac{\partial}{\partial r} s_{m,1} \right) + 2[m(\mu + \sigma) - R_L(x_m^*(s_{m,1}))] \\
&= 2(r - \frac{1}{2}) b(\mu - \sigma)^2 \frac{\partial}{\partial r} x_m^*(s)\Big|_{s=s_{m,1}} + 2[m(\mu + \sigma) - R_L(x_m^*(s_{m,1}))],
\end{aligned}$$

where the second equality holds because $\frac{\partial}{\partial s} x_m^*(s)|_{s=s_{m,1}} = 0$. We have shown in Lemma A.6 that $x_m^*(s)$ increases in r. Moreover, we have shown in the proof of Lemma A.6 that $m(\mu + \sigma) - R_L(x_m^*(s)) \geq s \geq 0$ for any $s \in [0, \bar{s}]$. Thus, $\frac{\partial}{\partial r} s_{m,1} \geq 0$. We conclude that $s_{m,1}$ increases in r if $s_{m,1} < \bar{s}$.

Next, we prove that if $s_{m,1} < \bar{s}$, then $s_{m,1}$ increases in m. To do so, as we already know that $x_m^*(s)$ increases in s if and only if $s \leq s_{m,1}$ and that $\frac{\partial}{\partial s} x_m^*(s)|_{s=s_{m,1}} = 0$ if $s_{m,1} < \bar{s}$, it is sufficient to prove that if $s_{m,1} < \bar{s}$, then $\frac{\partial^2}{\partial m \partial s} x_m^*(s)|_{s=s_{m,1}} \geq 0$. Let $\Delta_m(s) = \beta_m^2(s) - 4\alpha_m\gamma_m(s)$. By taking derivative of $x_m^*(s)$ with respect to m and s, we have

$$\begin{aligned}
\frac{\partial^2}{\partial m \partial s} x_m^*(s)\Big|_{s=s_{m,1}} &= \frac{1}{2\alpha_m} \left(\Delta_m^{-\frac{1}{2}}(s_{m,1}) \times \frac{\partial^2}{\partial m \partial s} \Delta_m(s)\Big|_{s=s_{m,1}} + \frac{\partial}{\partial m} \Delta_m(s)\Big|_{s=s_{m,1}} \frac{\partial}{\partial s} \Delta_m^{-\frac{1}{2}}(s)\Big|_{s=s_{m,1}} \right) \\
&= \frac{1}{2\alpha_m} \times \frac{\partial}{\partial m} \Delta_m(s)\Big|_{s=s_{m,1}} \frac{\partial}{\partial s} \Delta_m^{-\frac{1}{2}}(s)\Big|_{s=s_{m,1}},
\end{aligned} \tag{A.27}$$

where the second equality holds because, through algebraic calculation, we have that $\frac{\partial^2}{\partial m \partial s} \Delta_m(s) = 0$ for any s. Moreover, we have $\frac{\partial}{\partial m} \Delta_m(s)|_{s=s_{m,1}} = \lambda b(\mu - \sigma)^2(\mu + \sigma)(c + b(\mu - \sigma)^2) \geq 0$. Therefore, to prove that $\frac{\partial^2}{\partial m \partial s} x_m^*(s)|_{s=s_{m,1}} \geq 0$, it remains to show that $\frac{\partial}{\partial s} \Delta_m^{-\frac{1}{2}}(s)|_{s=s_{m,1}} \geq 0$.

Since $\frac{\partial}{\partial s} x_m^*(s)|_{s=s_{m,1}} = 0$ when $s_{m,1} < \bar{s}$, we have

$$\frac{\partial}{\partial s} x_m^*(s)\Big|_{s=s_{m,1}} = \frac{\partial}{\partial s} \frac{-\beta(s) + \Delta_m^{\frac{1}{2}}(s)}{2\alpha_m}\Big|_{s=s_{m,1}} = 0.$$

Because we know from their expressions in Lemma A.5 that $\frac{\partial}{\partial s}(-\beta(s)) \geq 0$ and α_m is independent of s, we must have $\frac{\partial}{\partial s}\Delta_m^{\frac{1}{2}}(s)|_{s=s_{m,1}} \leq 0$. Therefore, we have $\frac{\partial}{\partial s}\Delta_m^{-\frac{1}{2}}(s)|_{s=s_{m,1}} \geq 0$, which implies that $\frac{\partial^2}{\partial m \partial s}x_m^*(s)|_{s=s_{m,1}} \geq 0$. We conclude that $s_{m,1}$ increases in m if $s_{m,1} < \bar{s}$.

Case 3: We consider the scenario where $x_m^*(s_{m,1}) > x_2$. According to Lemma A.5, we have $x_m^*(s_{m,1}) = (m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s_{m,1} - \frac{1}{4}\lambda s_{m,1}^2 + 2\lambda\sigma(r-\frac{1}{2})s_{m,1}m)/c$. Since $s_{m,1} < \bar{s}$ and $(m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \frac{1}{4}\lambda s^2 + 2\lambda\sigma(r-\frac{1}{2})sm)/c$ is a quadratic concave function in s, we have

$$\frac{\partial}{\partial s}x_m^*(s)\Big|_{s=s_{m,1}} = \frac{\partial}{\partial s}\frac{m\mu - \lambda m^2\sigma^2 + \frac{1}{2}s - \frac{1}{4}\lambda s^2 + 2\lambda\sigma(r-\frac{1}{2})sm}{c}\Big|_{s=s_{m,1}} = 0.$$

Thus, we have

$$s_{m,1} = \frac{1}{\lambda} + 4(r - \frac{1}{2})m\sigma, \tag{A.28}$$

which increases in r and m.

Combining all three cases, we conclude that given $m \in [0, \bar{m}]$ and $r \in [\frac{1}{2}, 1]$, if $s_{m,1} < \bar{s}$, then $s_{m,1}$ increases in both r and m.

Step 3: We prove that there exists a threshold $r_{m,1}$ such that $s_{m,1} < \bar{s}$ if $r < r_{m,1}$. In order to construct the threshold $r_{m,1}$, we let $r'_{m,1} \in [\frac{1}{2}, 1]$ denote the lowest index accuracy such that $s_{m,1} = \bar{s}$, and we next prove the existence of such $r'_{m,1}$. To do so, it is sufficient to prove that when $r = 1$, we have $s_{m,1} = \bar{s}$. Similar to Step 2, we separate the further proof into three cases: In Case 1, $x_m^*(s_{m,1}) < x_1$; in Case 2, $x_m^*(s_{m,1}) \in [x_1, x_2]$; and in Case 3, $x_m^*(s_{m,1}) > x_2$.

Case 1: We consider the scenario where $x_m^*(s_{m,1}) < x_1$. Then, as discussed in Step 2, we have $s_{m,1} = s_1$. In the proof of Proposition 1, we have shown that $s_1 = \bar{s}$ when $r = 1$. Thus, we have $s_{m,1} = \bar{s}$ when $r = 1$.

Case 2: We consider the scenario where $x_m^*(s_{m,1}) \in [x_1, x_2]$. We prove the result by contradiction. Suppose $s_{m,1} < \bar{s}$ when $r = 1$. Then, by the definition of $\bar{s}$, we have

$$\begin{aligned} s_{m,1} &< \max\{a - bx_m^*(s_{m,1})(\mu+\sigma), m\}(\mu+\sigma) - \max\{a - bx_m^*(s_{m,1})(\mu-\sigma), m\}(\mu-\sigma) \\ &= m(\mu+\sigma) - R_L(x_m^*(s_{m,1})). \end{aligned}$$

The equality holds because we have $a - bx_m^*(s_{m,1})(\mu-\sigma) \geq m \geq a - bx_m^*(s_{m,1})(\mu+\sigma)$ by the definition of x_1 and x_2.

On the other hand, when $r = 1$ and $s_{m,1} < \bar{s}$, by Equation (A.26), we have

$$s_{m,1} = \frac{1}{\lambda} + 2(r - \frac{1}{2})[m(\mu+\sigma) - R_L(x_m^*(s_{m,1}))] > m(\mu+\sigma) - R_L(x_m^*(s_{m,1})),$$

which leads to a contradiction. Hence, in this case, $s_{m,1} = \bar{s}$ when $r = 1$.

Case 3: We consider the scenario where $x_m^*(s_{m,1}) > x_2$. We prove the result by contradiction. Suppose $s_{m,1} < \bar{s}$ when $r = 1$. Then, by the definition of $\bar{s}$, we have

$$\begin{aligned} s_{m,1} &< \max\{a - bx_m^*(s_{m,1})(\mu+\sigma), m\}(\mu+\sigma) - \max\{a - bx_m^*(s_{m,1})(\mu-\sigma), m\}(\mu-\sigma) \\ &= m(\mu+\sigma) - m(\mu-\sigma) \\ &= 2m\sigma \end{aligned}$$

The first equality holds because we have $m > a - bx_m^*(s_{m,1})(\mu-\sigma) > a - bx_m^*(s_{m,1})(\mu+\sigma)$ by the definition of x_2.

On the other hand, when $r = 1$ and $s_{m,1} < \bar{s}$, by Equation (A.28), we have

$$s_{m,1} = \frac{1}{\lambda} + 4(r - \frac{1}{2})m\sigma > 2m\sigma,$$

which leads to a contradiction. Hence, combining all three cases, we conclude that $s_{m,1} = \bar{s}$ when $r = 1$. This result implies the existence of $r'_{m,1} \in [\frac{1}{2}, 1]$, defined as the lowest index accuracy such that $s_{m,1} = \bar{s}$. Therefore, letting $r_{m,1} = r'_{m,1}$ is sufficient for the proposition to hold. We next construct an alternative threshold $r_{m,1}$ such that we can use the same threshold $r_{m,1}$ in this proposition and Theorem 2.

Recall that in §5, x^{opt} is the value of $x_m^*(s)$ that maximizes the net benefit $v_m(s)$. Let $r_c \in [\frac{1}{2}, 1]$ denote the lowest index accuracy under which $x_m^*(s_m^*) = x^{opt}$. Such r_c is guaranteed to exist because by following the same arguments as in the proof of Theorem 1, it can be easily shown that $x^*(s_m^*) = x^{opt}$ when $r = 1$. Define $r_{m,1} := \min\{r'_{m,1}, r_c\}$. Then, we have that $s_{m,1} < \bar{s}$ if $r < r_{m,1}$.

Step 4: We prove that there exists a threshold $r_{m,2}$ such that $s_{m,1} = \bar{s}$ if $r \geq r_{m,2}$. First, let $r'_{m,2} \in [\frac{1}{2}, 1]$ denote the lowest index accuracy that satisfies $s_{m,1} = \bar{s}$ for any $r \geq r'_{m,2}$. Such $r'_{m,2}$ is guaranteed to exist because we have shown in Step 3 that $s_{m,1} = \bar{s}$ when $r = 1$. Therefore, letting $r_{m,2} = r'_{m,2}$ is sufficient for the proposition to hold. We next construct an alternative threshold $r_{m,2}$ such that we can use the same threshold $r_{m,2}$ in this proposition and Theorem 2.

Next, we prove that there exists a threshold $r''_{m,2} < 1$ such that for any $r \geq r''_{m,2}$, we have $\frac{\partial}{\partial m} x_m^*(s) \geq 0$ for all $s \in [0, \bar{s}]$. Recall that we consider m such that $\frac{\partial}{\partial m} x_m^*(0) \geq 0$. Thus, it is sufficient to prove that there exists a threshold $r''_{m,2} < 1$ such that for any $r \geq r''_{m,2}$, we have $\frac{\partial^2}{\partial s \partial m} x_m^*(s) \geq 0$ for any $m \in [0, \bar{m}]$ and $s \in [0, \bar{s}]$. Consider a fixed $m \in [0, \bar{m}]$ and a fixed $s \in [0, \bar{s}]$. By Lemma A.5, given s, there exist two thresholds m_1 and m_2. Accordingly, we prove $\frac{\partial^2}{\partial s \partial m} x_m^*(s) \geq 0$ in the following three cases:

Case 1: If $m < m_1$, then, by Lemma A.5, we have $x_m^*(s) = x^*(s)$ and thus $\frac{\partial^2}{\partial s \partial m} x_m^*(s) = 0$.

Case 2: If $m \in [m_1, m_2]$, then, by Equation (A.27), we have

$$\frac{\partial^2}{\partial s \partial m} x_m^*(s) = \frac{1}{2\alpha_m} \times \frac{\partial}{\partial m} \Delta_m(s) \frac{\partial}{\partial s} \Delta_m^{-\frac{1}{2}}(s),$$

where $\Delta_m(s) = \beta_m^2(s) - 4\alpha_m\gamma_m(s)$ and we have shown in Step 2 that $\frac{\partial}{\partial m}\Delta_m(s) \geq 0$. Therefore, to prove that $\frac{\partial^2}{\partial s \partial m} x_m^*(s) \geq 0$, it remains to show $\frac{\partial}{\partial s}\Delta_m^{-\frac{1}{2}}(s) \geq 0$. By rearranging terms of $\Delta_m(s)$, we have that $\Delta_m(s)$ is a concave function in s. Moreover, we have $\frac{\partial}{\partial s}\Delta_m(s)|_{s=0} = -\lambda b(\mu-\sigma)^2[c(2r-1) + b(r-1)(\mu-\sigma)^2]$. From this expression, it is clear that we have $\frac{\partial}{\partial s}\Delta_m(s)|_{s=0} < 0$ when $r = 1$. Therefore, there must exist an $r''_{m,2} < 1$ (independent of s and m) such that for any $r \geq r''_{m,2}$, we have $\frac{\partial}{\partial s}\Delta_m(s)|_{s=0} \leq 0$ and thus $\frac{\partial}{\partial s}\Delta_m^{-\frac{1}{2}}(s) \geq 0$ for any s, which implies that $\frac{\partial^2}{\partial s \partial m} x_m^*(s) \geq 0$.

Case 3: If $m > m_2$, then, by Lemma A.5, we have

$$\frac{\partial^2}{\partial s \partial m} x_m^*(s) = \frac{2\lambda\sigma(r - \frac{1}{2})}{c} \geq 0.$$

Combining all three cases, we conclude that if $r \geq r''_{m,2}$, then we have $x_m^*(s)$ increases in m for any $s \in [0, \bar{s}]$.

Finally, recall that r_c is a threshold defined in Step 3. Define $r_{m,2} := \max\{r'_{m,2}, r''_{m,2}, r_c\}$. Then, we have that $s_{m,1} = \bar{s}$ if $r \geq r_{m,2}$. Moreover, by the definition of $r'_{m,1}$ and $r'_{m,2}$, we have that $r'_{m,1} \leq r'_{m,2}$, which implies $r_{m,1} \leq r_{m,2}$.

Combining all four steps, we conclude that if $r < r_{m,1}$, then $x_m^*(s)$ increases in s if and only if $s \leq s_{m,1}$, where $s_{m,1} \in (0, \bar{s})$ increases in both r and m; if $r \geq r_{m,2}$, then $x^*(s)$ increases in s for all $s \in [0, \bar{s}]$. □

Proof of Theorem 2: From Equation (6), given $m \in [0, \bar{m}]$ and $s \in [0, \bar{s}]$, we have

$$\begin{aligned} v_m(s) = &\int_0^{x_m^*(s)} \mathbb{E}\left[\max\{(a - bx_m^*(s)Y), m\} \times Y + sI - xc\right] \mathrm{d}x - x_m^*(s)\mathbb{E}[sI] \\ &- x_m^*(s)\mathbb{E}[\max\{m - (a - bx_m^*(s)Y), 0\} \times Y] \\ =& v_1 \times x_m^*(s) - v_2 \times (x_m^*(s))^2. \end{aligned}$$

where $v_1 = a\mu$ and $v_2 = b(\mu^2 + \sigma^2) + \frac{1}{2}c$. We have $x^{opt} = \min\{\frac{v_1}{2v_2}, 1\}$ as the value of $x_m^*(s)$ that maximizes $v_m(s)$. Thus, the optimal s_m^* must make $x_m^*(s)$ as close to x^{opt} as possible. In addition, since we consider m such that $x_m^*(0) \leq x^{opt}$, s_m^* must satisfy one of the following two conditions: (1) $x_m^*(s_m^*) < x^{opt}$ and $s_m^* = \inf\{\arg\max_{s\in[0,\bar{s}]} x_m^*(s)\}$; (2) $s_m^* = \inf\{s : x_m^*(s) = x^{opt}\}$.

We first consider $r < r_{m,1}$. By the definition of $r_{m,1}$ in the proof of Proposition 4, we have that $x_m^*(s) < x^{opt}$ for any $s \in [0, \bar{s}]$. In this case, as discussed earlier, s_m^* must satisfy $x_m^*(s_m^*) < x^{opt}$ and $s_m^* = \inf\{\arg\max_{s\in[0,\bar{s}]} x_m^*(s)\}$. Moreover, we have $x_m^*(s) < x^{opt} \leq 1$ for all $s \in [0, \bar{s}]$ and thus the maximizer of $x_m^*(s)$ must be unique as implied by Step 1 of the proof of Proposition 4. Hence, we have $s_m^* = \arg\max_{s\in[0,\bar{s}]} x_m^*(s) = s_{m,1}$ defined in Proposition 4. Further, since we have shown in Proposition 4 that $s_{m,1}$ increases in m, we have s_m^* increases in m.

We then consider $r \geq r_{m,2}$. Recall that x^{opt} is independent of r and we have shown in Lemma A.6 that $x_m^*(s)$ increases in r. Therefore, if for some index accuracy r, there exists an $s \in [0, \bar{s}]$ such that $x_m^*(s) = x^{opt}$, then, for any higher index accuracy, there must also exist an $s \in [0, \bar{s}]$ such that $x_m^*(s) = x^{opt}$. In this case, by the definition of $r_{m,2}$ in the proof of Proposition 4, s_m^* must satisfy $x_m^*(s_m^*) = x^{opt}$. Recall that x^{opt} is independent of m and we have shown in Proposition 4 that for any $r \geq r_{m,2}$, $x_m^*(s)$ increases in m. Moreover, we have $\frac{\partial}{\partial s} x_m^*(s)|_{s=s_m^*} \geq 0$ by the definition of s_m^*. Hence, s_m^* must decrease in m to maintain $x_m^*(s_m^*) = x^{opt}$. $\square$

Appendix A.3 Proofs of Analytical Results in §6

Proof of Proposition 5: From Equation (7) in §6, we have

$$\begin{aligned} v(s,r) &= \int_0^{x^*(s,r)} \mathbb{E}[(a - bx^*(s,r)Y)Y + sI - xc]\,\mathrm{d}x - \psi(s,r) - \phi(r) \\ &= \int_0^{x^*(s,r)} \mathbb{E}[(a - bx^*(s,r)Y)Y + sI - xc]\,\mathrm{d}x - x^*(s,r)\mathbb{E}[sI] - \kappa(r - r_0)^2 \\ &= v_1 \times x^*(s,r) - v_2 \times (x^*(s,r))^2 - \kappa(r - r_0)^2, \end{aligned} \tag{A.29}$$

where $v_1 = a\mu$ and $v_2 = b(\mu^2 + \sigma^2) + \frac{1}{2}c$. Then, we prove this proposition in the following three steps:

Step 1: We prove that for any given budget level $B > 0$, there exists a threshold σ_1 such that if $\sigma \leq \sigma_1$, then $r^* = r_0$ and s^* increases in σ. Recall that we consider scenarios where $x_0^* \leq x^{opt}$ throughout the paper. Since $\frac{\partial}{\partial \sigma} x_0^* < \frac{\partial}{\partial \sigma} x^{opt} < 0$ (as implied by Lemma A.4), the condition $x_0^* \leq x^{opt}$ is equivalent to that σ is higher than a certain threshold, which we denote as $\underline{\sigma}$. Therefore, to prove our conclusion under the condition $x_0^* \leq x^{opt}$, it is sufficient to consider $\sigma \geq \underline{\sigma}$.

Consider a fixed $B > 0$. For any given index accuracy r and yield variability σ, we say that the planting amount x^{opt} is achievable if there exists an $s \in [0, \bar{s}]$ that satisfies $x^*(s,r) = x^{opt}$ and $\psi(s,r) + \phi(r) \leq B$. Let σ_1 be the largest yield variability such that the planting amount x^{opt} is achievable when the index accuracy is $r = r_0$. Such σ_1 must exist because it is straightforward to check that if $\sigma = \underline{\sigma}$, then we have $x^*(0, r_0) = x_0^* = x^{opt}$ and $\psi(0, r_0) + \phi(r_0) = 0 < B$. It can be seen from Equation (A.29) that if there exists an s that satisfies $x^*(s, r_0) = x^{opt}$ and $\psi(s, r_0) + \phi(r_0) \leq B$, then this s must be the optimal subsidy s^* and we also have $r^* = r_0$. Therefore, if $\sigma \leq \sigma_1$, then we have $r^* = r_0$. Moreover, we have shown in Proposition 3 that for any given r, the optimal subsidy amount increases in σ. Hence, when $\sigma \leq \sigma_1$, as we already showed that $r^* = r_0$, we have s^* increases in σ.

Step 2: We prove that there exists a constant $\Delta_\sigma > 0$ such that r^* is continuous in σ if $\sigma < \sigma_1 + \Delta_\sigma$. Since we have shown in Step 1 that $r^* = r_0$ when $\sigma \leq \sigma_1$, we have that r^* is continuous in σ when $\sigma < \sigma_1$. Then, to prove the existence of such Δ_σ, by the definition of continuity, it is sufficient to prove that r^* is continuous in σ at $\sigma = \sigma_1$.

For ease of exposition, under the yield variability σ, let $s^*(\sigma)$ and $r^*(\sigma)$ be s^* and r^*; let $v(s,r|\sigma)$ be $v(s,r)$; and let $v_2(\sigma)$ and $x^{opt}(\sigma)$ be v_2 and x^{opt}. Then, to prove the continuity of r^* at $\sigma=\sigma_1$, since we already have $r^*=r_0$ when $\sigma\leq\sigma_1$, it is sufficient to prove that for any $\epsilon>0$, there exists a constant $\delta>0$ such that for any $\sigma\in[\sigma_1,\sigma_1+\delta]$, we have $r^*(\sigma)-r^*(\sigma_1)=r^*(\sigma)-r_0\leq\epsilon$.

Consider a fixed $\epsilon>0$. Since $x^*(s,r)$ is continuous in σ, by Equation (A.29), $v(s,r|\sigma)$ is also continuous in σ. Thus, there must exist a $\delta>0$ such that for any $\sigma\in[\sigma_1,\sigma_1+\delta]$, we have

$$v(s^*(\sigma_1),r^*(\sigma_1)|\sigma_1)-v(s^*(\sigma_1),r^*(\sigma_1)|\sigma)=v(s^*(\sigma_1),r_0|\sigma_1)-v(s^*(\sigma_1),r_0|\sigma)\leq\kappa\epsilon^2. \tag{A.30}$$

Then, for any $\sigma\in[\sigma_1,\sigma_1+\delta]$, we claim that $r^*(\sigma)-r^*(\sigma_1)=r^*(\sigma)-r_0\leq\epsilon$. We prove this claim by contradiction. Suppose there exists a $\sigma'\in[\sigma_1,\sigma_1+\delta]$ such that $r^*(\sigma')>r_0+\epsilon$. Then, by Equation (A.29), we have

$$\begin{aligned} v(s^*(\sigma_1),r^*(\sigma_1)|\sigma_1)-v(s^*(\sigma'),r^*(\sigma')|\sigma') &= \left[v_1x^{opt}(\sigma_1)-v_2(\sigma_1)\big(x^{opt}(\sigma_1)\big)^2\right]-v(s^*(\sigma'),r^*(\sigma')|\sigma') \\ &> \left[v_1x^{opt}(\sigma_1)-v_2(\sigma_1)\big(x^{opt}(\sigma_1)\big)^2\right] \\ &\quad -\left[v_1x^{opt}(\sigma')-v_2(\sigma')\big(x^{opt}(\sigma')\big)^2\right]+\kappa\epsilon^2 \\ &\geq \kappa\epsilon^2, \end{aligned} \tag{A.31}$$

where the first equality holds because at $\sigma=\sigma_1$, we have $x^*(s^*(\sigma_1),r^*(\sigma_1))=x^{opt}(\sigma_1)$ and $r^*(\sigma_1)=r_0$; the first inequality holds because $x^{opt}(\sigma')$ is defined as the planting amount that maximizes $v_1x^*(s,r)-v_2(\sigma')(x^*(s,r))^2$ and we assumed $r^*(\sigma')>r_0+\epsilon$; and the last inequality holds because $x^{opt}(\sigma)=\min\{\frac{v_1}{2v_2(\sigma)},1\}$ and $\frac{\partial}{\partial\sigma}[v_1\frac{v_1}{2v_2(\sigma)}-v_2(\sigma)(\frac{v_1}{2v_2(\sigma)})^2]=\frac{\partial}{\partial\sigma}\frac{v_1^2}{4v_2(\sigma)}=\frac{-v_1^2b\sigma}{2v_2^2(\sigma)}<0$, which implies $v_1x^{opt}(\sigma)-v_2(\sigma)\big(x^{opt}(\sigma)\big)^2$ decreases in σ.

However, by Equation (A.30), we have that $v(s^*(\sigma_1),r^*(\sigma_1)|\sigma_1)-v(s^*(\sigma_1),r^*(\sigma_1)|\sigma')\leq\kappa\epsilon^2$. Comparing with Equation (A.31), it implies that

$$v(s^*(\sigma_1),r^*(\sigma_1)|\sigma')>v(s^*(\sigma'),r^*(\sigma')|\sigma'),$$

which leads to a contradiction with the optimality of $s^*(\sigma')$ and $r^*(\sigma')$. Hence, we conclude that r^* is continuous at $\sigma=\sigma_1$ and thus there must exist a $\Delta_\sigma>0$ such that r^* is continuous in σ if $\sigma<\sigma_1+\Delta_\sigma$.

Step 3: We prove that for any given budget level $B>0$, there exists a threshold σ_2 such that if $\sigma_1\leq\sigma<\sigma_2$, r^* increases in σ. Recall that r_0 is the lowest possible index accuracy, and we have shown that $r^*=r_0$ when $\sigma=\sigma_1$. Thus, as we have proven in Step 2 that, for $\sigma<\sigma_1+\Delta_\sigma$, r^* is continuous in σ, to reach the conclusion of this step, we only need to prove that there is no $\delta\in(0,\Delta_\sigma)$ such that $r^*=r_0$ for any $\sigma\in(\sigma_1,\sigma_1+\delta)$. Further, it is equivalent to prove that for any $\delta\in(0,\Delta_\sigma)$, there exists a $\sigma\in(\sigma_1,\sigma_1+\delta)$ under which $r^*>r_0$.

Consider a fixed $\delta \in (0, \Delta_\sigma)$. We would like to prove that there exists a $\sigma \in (\sigma_1, \sigma_1 + \delta)$ such that $r^* > r_0$. Note that by the definition of σ_1, we have that for any $\delta' > 0$, there must exist a $\sigma \in (\sigma_1, \sigma_1 + \delta')$ under which x^{opt} is not achievable when $r = r_0$. That is, when $r = r_0$, for any $s \in [0, \bar{s}]$ that satisfies $\psi(s, r_0) + \phi(r_0) \leq B$, we have $x^*(s, r_0) < x^{opt}$. Thus, as we consider a fixed δ, there must exist a $\sigma' \in (\sigma_1, \sigma_1 + \delta)$ such that x^{opt} is not achievable when $\sigma = \sigma'$.

We next prove that when $\sigma = \sigma'$, we have $r^* > r_0$. To do so, it is sufficient to prove that when $\sigma = \sigma'$ and $r = r_0$, for any $s \in (0, \bar{s}]$ that satisfies $\psi(s, r_0) + \phi(r_0) \leq B$,[10] we have (1) $\frac{\partial}{\partial r} v(s, r)|_{r=r_0} > 0$ and (2) $\frac{\partial}{\partial \sigma}[\psi(s, r_0) + \phi(r_0)] < 0$.

Consider $\sigma = \sigma'$ and a fixed $s \in (0, \bar{s}]$ such that $\psi(s, r_0) + \phi(r_0) \leq B$. For (1), from Equation (A.29), we have

$$\frac{\partial}{\partial r} v(s, r)\Big|_{r=r_0} = (v_1 - 2v_2 x^*(s, r_0)) \frac{\partial}{\partial r} x^*(s, r)\Big|_{r=r_0} - \frac{\partial}{\partial r}\phi(r)\Big|_{r=r_0} = (v_1 - 2v_2 x^*(s, r_0)) \frac{\partial}{\partial r} x^*(s, r)\Big|_{r=r_0},$$

where the last equality holds because $\frac{\partial}{\partial r}\phi(r)|_{r=r_0} = \frac{\partial}{\partial r}\kappa(r - r_0)^2|_{r=r_0} = 0$. As x^{opt} is not achievable when $\sigma = \sigma'$, we have that $x^*(s, r_0) < x^{opt} \leq \frac{v_1}{2v_2}$, which implies $v_1 - 2v_2 x^*(s, r_0) > 0$. Moreover, since $s > 0$, by Equation (A.13), we have $\frac{\partial}{\partial r} x^*(s, r)|_{r=r_0} > 0$. Hence, we have $\frac{\partial}{\partial r} v(s, r)|_{r=r_0} > 0$. For (2), we have

$$\frac{\partial}{\partial \sigma}[\psi(s, r_0) + \phi(r_0)] = \frac{s}{2} \frac{\partial}{\partial \sigma} x^*(s, r_0) < 0,$$

where the last inequality holds because we have $\frac{\partial}{\partial \sigma} x^*(s, r) < 0$ by Lemma A.4. Hence, when $\sigma = \sigma'$, we must have $r^* > r_0$.

Finally, as the previous result holds for any $\delta \in (0, \Delta_\sigma)$, we can conclude that there is no $\delta \in (0, \Delta_\sigma)$ such that $r^* = r_0$ for any $\sigma \in (\sigma_1, \sigma_1 + \delta)$. Therefore, there must exist a $\sigma_2 \in (\sigma_1, \sigma_1 + \Delta_\sigma)$ such that r^* increases in σ when $\sigma \in [\sigma_1, \sigma_2)$. $\square$

Proof of Theorem 3: According to Proposition 5, under a given budget level B, there exist two thresholds σ_1 and σ_2 such that if $\sigma \leq \sigma_1$, then $r^*(B) = r_0$; if $\sigma \in [\sigma_1, \sigma_2]$, r^* increases in σ. For ease of exposition, let $\sigma_1(B)$ and $\sigma_2(B)$ denote σ_1 and σ_2 under a budget B respectively.

Consider any two budget levels B_1 and B_2 such that $0 < B_1 < B_2$. By the definition of $\sigma_1(B)$ shown in the proof of Proposition 5, $\sigma_1(B)$ increases in B. Therefore, we have $\sigma_1(B_1) \leq \sigma_1(B_2)$. Let $I_\sigma = [\sigma_1(B_1), \min\{\sigma_1(B_2), \sigma_2(B_1)\}]$. We next prove that if $\sigma \in I_\sigma$, we have $r^*(B_1) \geq r^*(B_2)$ and $s^*(B_1) \leq s^*(B_2)$.

First, if $\sigma \in I_\sigma$, then by definition of I_σ, we have $\sigma \leq \sigma_1(B_2)$, which implies that $r^*(B_2) = r_0$. Moreover, we have $\sigma \in [\sigma_1(B_1), \sigma_2(B_1)]$. Since we already know that $r^*(B_1)$ increases in σ if $\sigma \in [\sigma_1(B_1), \sigma_2(B_1)]$, we have $r^*(B_1) \geq r_0$. Hence, we must have $r^*(B_1) \geq r^*(B_2)$ if $\sigma \in I_\sigma$.

[10] We consider $s > 0$ because it is straightforward that $s = 0$ cannot be jointly optimal.

Second, since we have shown that $r^*(B_1) \geq r^*(B_2)$, we must have $s^*(B_1) \leq s^*(B_2)$. Otherwise, if $s^*(B_1) > s^*(B_2)$, consider the following two cases: Suppose $x^*\big(s^*(B_1), r^*(B_1)\big) > x^*\big(s^*(B_2), r^*(B_2)\big)$. By the definition of $\sigma_1(B_2)$, we have $x^*\big(s^*(B_2), r^*(B_2)\big) = x^{opt}$ if $\sigma \in I_\sigma$. Then, we have $x^*\big(s^*(B_1), r^*(B_1)\big) > x^{opt}$. By Equation (A.29), it implies that $\psi\big(s^*(B_2), r^*(B_2)\big) + \phi\big(r^*(B_2)\big) < B_1$ and $v\big(s^*(B_2), r^*(B_2)\big) > v\big(s^*(B_1), r^*(B_1)\big)$, which leads to a contradiction with the joint optimality of $s^*(B_1)$ and $r^*(B_1)$. Suppose $x^*\big(s^*(B_1), r^*(B_1)\big) \leq x^*\big(s^*(B_2), r^*(B_2)\big)$. Then, we can also conclude $(s^*(B_1), r^*(B_1))$ must not be the optimal solution, because $(s^*(B_1), r^*(B_1))$ must be dominated by $(s^*(B_2), r^*(B_2))$, where the latter leads to a planting amount of x^{opt} and results in a lower expenditure. □

Appendix B Numerical Calibration

In this section, we describe in detail the steps taken to estimate some model parameters used in numerical experiments either from the real-world data or from reasonable assumptions.

First, we present how we generate the values the parameters of farmer's yield, μ and σ. In §7.1, we have each individual farmer's yields as y_i, where $y_i \sim \mathcal{N}(\mu, \sigma^2)$ for $i = 1, ..., n$. Since the variance of individual farmer's yield σ^2 is hard to evaluate from data. We separate y_i into two parts: Let $y_i = y_a + \epsilon_i$, where $y_a \sim \mathcal{N}(\mu, \sigma_a^2)$ and $\epsilon_1, ..., \epsilon_n \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$. We also assume ϵ_i are independent of y_a for all i, which implies $\sigma^2 = \sigma_a^2 + \sigma_\epsilon^2$. Therefore, to find the distribution of individual farmer's yield, we need to evaluate the values of μ, σ_a and σ_ϵ.

For μ and σ_a, we get their values from USDA's Grain and Feed Annual Reports[11]. For each year from 2011 to 2021, there is a report containing the average corn yield in Indonesia (e.g., USDA (2021) is the report of 2021). From the mean and variance of those average yield data, we can estimate that $\mu = 3.029$ tons per hectare and $\sigma_a = 0.324$ tons per hectare.

For σ_ϵ, we can get its value from σ_a and the correlation ρ. In §7.1, we have $\rho = \text{corr}(y_i, y_j)$ for $i \neq j$, where $\rho \in [0, 1]$ is a constant. Given ρ and σ_a, we have, for any $i \neq j$,

$$\text{corr}(y_i, y_j) = \frac{\text{cov}(y_i, y_j)}{\sigma^2} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_\epsilon^2} = \rho$$

Thus, we get $\sigma_\epsilon = \sigma_a\sqrt{(1-\rho)/\rho}$ and $\sigma = \sqrt{\sigma_\epsilon^2 + \sigma_a^2} = \sigma_a \frac{1}{\sqrt{\rho}}$. We then obtain the value of farmers' yield correlation ρ from the result of an empirical study. According to Wang et al. (1998), the correlation between the corn yield of an individual farmer and the average corn yield in a county

[11] Go to website https://gain.fas.usda.gov/#/search, then click "Advanced Search". In the search page, under "Countries", select "Indonesia"; under "Categories", check "Grain and Feed"; and under "Report Name" at "Additional Key Word Filter", type in "Grain and Feed Annual". Finally, click "Search". In the search results, for each year, there is a link to the corresponding report.

in Iowa is approximately 0.8. Define $\bar{y} := \frac{1}{n}\sum_{i=1}^{n} y_i$. By incorporating this empirical result in our framework, we have

$$\mathrm{corr}(y_i, \bar{y}) = \frac{\mathrm{cov}(y_i, \bar{y})}{\sigma^2\sqrt{\rho + \frac{1-\rho}{n}}} = \frac{\frac{1}{n}\sigma^2 + \frac{n-1}{n}\rho\sigma^2}{\sigma^2\sqrt{\rho + \frac{1-\rho}{n}}} \approx \frac{\rho}{\sqrt{\rho}} = 0.8$$

The last approximation holds because the value of n is very large compared to ρ and $1-\rho$. The solution to the last equation above is 0.64 and thus we set $\rho = 0.6$ for the numerical study, as we mentioned in §7.1. Collectively, $\sigma = \sigma_a \frac{1}{\sqrt{\rho}} = 0.324 \times \frac{1}{\sqrt{0.6}} \approx 0.418$ tons per hectare.

Next, we present how we estimate the market-clearing price parameters a and b, and the production cost parameter c. In the same USDA's Grain and Feed Annual reports mentioned above, they also include the market price and the total supply of corn in Indonesia. We adjust the price with the inflation data available at the World Bank[12]. Then, by fitting a linear model to the adjusted price and the total supply of corn, we get $a = \$467.32$ per ton and b =\$0.00189 per ton. For c, we obtain its value from USDA's Economic Research Service (USDA 2023)[13]. Though this data is generated from the surveys in the U.S, we select costs of the farming activities that are common in both developed and developing countries to approximate the costs in our context.

Finally, we discuss the selection of the value of risk aversion coefficient λ. By the same logic behind the assumption made in §3, we need $\lambda < \frac{\mu+\sigma}{2a\sigma^2} = 0.02$. Moreover, λ cannot be so small that all the farmers choose to plant even without any subsidy. Considering these constraints, we let $\lambda = 0.01$ in the numerical experiments such that when the subsidy amount $s = 0$, only a fraction of n farmers will choose to plant.

Appendix C Extensions

C.1 Conditional Value at Risk (CVaR) for Farmer Risk Aversion

In §8.1, we use an alternative risk measure, the conditional value at risk (CVaR), to model farmers' risk aversion. In this section, we show that our results are robust under this framework.

In the CVaR model, due to a discrete distribution for crop yield, the optimal subsidy amount can be discontinuous in index accuracy. This complicates the analysis, and thus reproducing all the analytical results derived from our base model is challenging. However, we are able to prove that the optimal subsidy amount remains non-monotonic in index accuracy, as shown in the following proposition.

[12] Go to https://data.worldbank.org/country/indonesia, then go to the tab "Inflation, consumer prices (annual %)"

[13] Go to the URL in the citation. Click "Corn" dataset under "Recent Cost and Returns" and download the excel. In the excel, we can find the operating cost for 2021 is \$354.14 per acre. Changing the unit from acre to hectare and multiplying it by 2, we have $c = \$1750.2$.

Proposition C.1. *When farmers make planting decisions based on their conditional value at risk (CVaR), the optimal subsidy amount first increases and then decreases in the index accuracy r.*

Proposition C.1 states that our non-monotonicity result presented in Theorem 1 continues to hold under the CVaR model, which demonstrates that our insights are not driven by the particular risk-averse model used in the main text (i.e., the mean-variance utility model). Moreover, we further numerically verify that our other main insights, including how the index accuracy affects the interaction between price and yield protection and how the budget level affects the jointly optimal subsidy amount and index accuracy, also remain intact under the CVaR model.

C.2 Consumer Surplus

In §8.2, we investigate an alternative government objective that considers consumer surplus in addition to farmer surplus when designing an index-based yield protection policy. In this section, we show that our results are robust under this framework. The following proposition presents the robustness of our key results in §4-6 as well as an additional insight.

Proposition C.2. *When the government includes consumer surplus in its objective function, all propositions and theorems presented in §4-6 continue to hold. Moreover, the optimal subsidy amount is higher compared to that in the base model without considering consumer surplus.*

As shown in Proposition C.2, all of our key insights in §4-6 remain intact when the government also considers consumer surplus when designing the index-based yield protection policy. Intuitively speaking, though the inclusion of the consumer surplus changes the government's objective, from the farmers' perspective, it does not change how a specific policy affects the farmers' utility and thus the farmers' planting decision. Therefore, the main drivers behind our key results on the optimal subsidy amount, such as high income variance when the index accuracy is low and oversupply issue when the index accuracy is high, remain effective with the inclusion of consumer surplus. The detailed proof is included in §C.5.

We also make an additional discovery that the government should offer greater subsidies to farmers under this framework compared to the base model. The logic behind this result is as follows. As discussed in §4, an increase in the subsidy amount can incentivize more farmers to plant, which results in a higher supply and a lower market-clearing price. In particular, when the index accuracy is high, a higher subsidy can more effectively motivating planting and boosting supply, which can largely benefit the consumers. Therefore, the government, aiming to optimize the sum of farmer and consumer surplus, would have incentives to entice more farmers to plant by further increasing the subsidy amount.

C.3 Different Distributions for Yield and Index

In §8.3, we consider different distributions for Y and I by introducing a new parameter, payment probability $p \in (0,1]$. The new joint probability of Y and I is presented in Table 2. In this section, we formally discuss the robustness of our key results under this new framework, and provide an additional insight.

First, given the joint probabilities in Table 2, we study the range of the index accuracy r. Consider the conditional probability $\mathbb{P}(I=1|Y=\mu-\sigma)=2rp$. By the same logic as in §3 that the index should be better than a random guess, we have

$$2rp = \mathbb{P}(I=1|Y=\mu-\sigma) \geq \frac{\mathbb{P}(I=1)\mathbb{P}(Y=\mu-\sigma)}{\mathbb{P}(Y=\mu-\sigma)} = p.$$

Besides, we have $\mathbb{P}(I=1|Y=\mu-\sigma)=2rp \leq 1$, which implies $r \leq \min\{1, 1/2p\}$. Collectively, we have $r \in [1/2, \min\{1, 1/2p\}]$, where $p \in (0,1]$.

Then, with the range of the index accuracy, we can show that all of our key insights generated from the base model continue to hold with the following proposition.

Proposition C.3. *Given a payment probability $p \in (0,1]$ and the joint probabilities in Table 2, all propositions and theorems presented in §4-6 continue to hold.*

Proposition C.3 states that all of our key results in §4-6 are robust in this extension. The general intuition behind this proposition is as follows. The conditional probability $\mathbb{P}(I=1|Y=\mu-\sigma)=2rp$, which, when p is given, is linear in the index accuracy r. Moreover, recall that in our base model, $p=\frac{1}{2}$ and $\mathbb{P}(I=1|Y=\mu-\sigma)=r$. Therefore, such "linear transformation" preserves the dynamics between the model parameters and the decisions of farmers and the government in the base case, and all of our analysis and results under our base model structurally hold. The detailed proof is included in Appendix C.5.

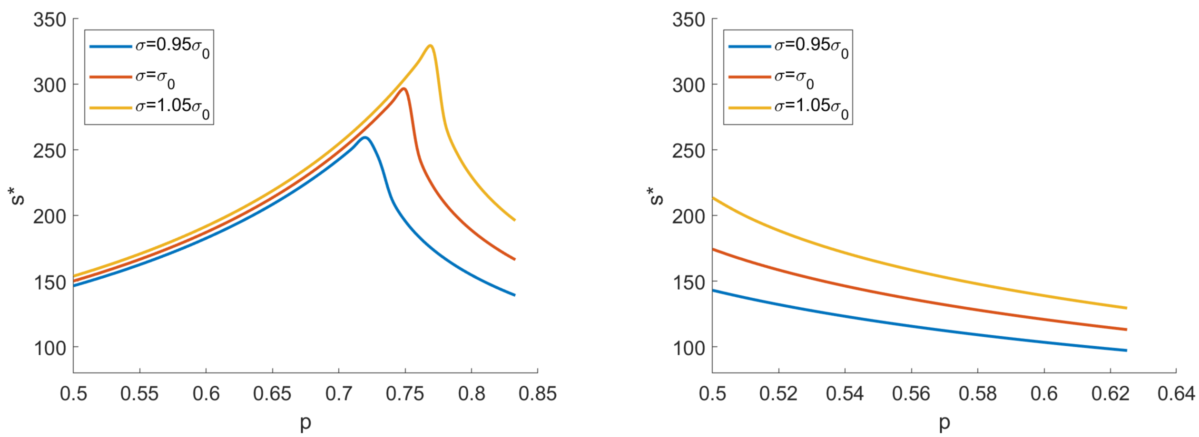
As discussed in §8.3, in practice, the government can control the value of payment probability p by adjusting the index design. Therefore, we are interested in exploring how the payment probability affects the optimal subsidy amount given the index accuracy. As solving for how s^* changes in p is analytically intractable, we resort to extensive numerical experiments to find their relationship. We find that, when the index accuracy is low, a higher payment probability p leads to a higher optimal subsidy payment; when the index accuracy is high, the optimal subsidy amount decreases in p. Figure C.1 presents a representative set of results.[14]

The reasoning behind this result is as follows. When the index accuracy is low, a higher payment probability, as indicated by the conditional probability $\mathbb{P}(I=1|Y=\mu-\sigma)$, leads to a higher chance

[14] All the relevant parameters take the same values as in §7, as we did in §C.2.

for the farmers to receive the subsidy payment when the yield is low. It makes the policy more effective in improving farmer surplus and motivating planting, as we discussed in §4. Therefore, when the index accuracy is low, a higher payment probability allows the government to offer a higher subsidy amount. On the other hand, when the index accuracy is high, the subsidy amount needs to be low to avoid the oversupply issue. With a higher payment probability, the equilibrium planting amount increases even faster in the subsidy amount, and so does the negative effect of the oversupply issue. Hence, when the index accuracy is high, the subsidy amount should be lower with a higher p.

Figure C.1 Optimal Subsidy Amount under Different Payment Probability



(a) $r = 0.6$ (b) $r = 0.8$

C.4 Effect of Premium

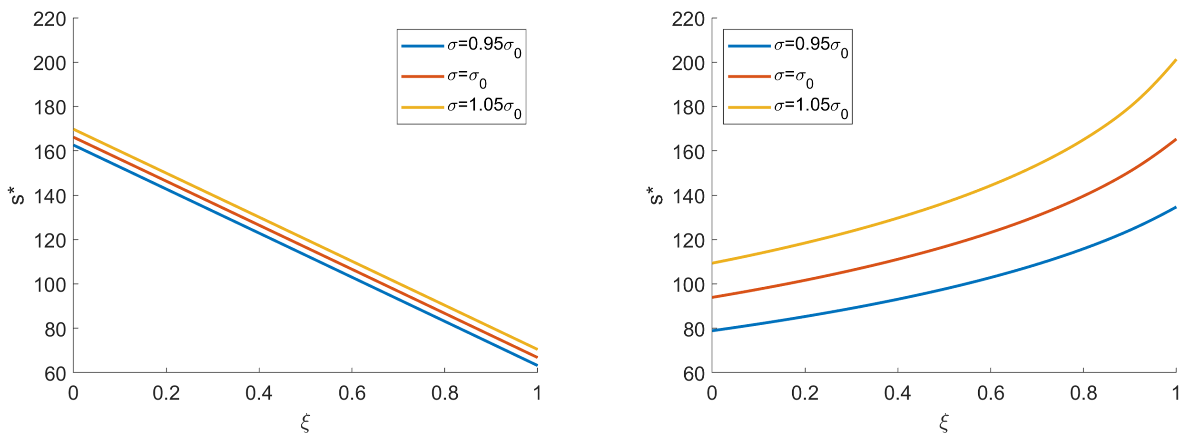
In this section, we focus on a scenario described in §8.4 where the farmers must pay a premium to enroll in the index-based yield protection program. We can analytically show that all of our key insights generated from the base model carry on in this new framework. The following proposition demonstrates such robustness.

Proposition C.4. *When the farmers need to pay a fraction $\xi \in [0,1]$ of the expected subsidy amount received to enroll in an index-based yield protection program, all propositions and theorems presented in §4-6 continue to hold.*

Proposition C.4 states that all of our key results in §4-6 continue to hold in this alternative model setup. In the detailed proof in Appendix C.5, we show that while the presence of premium reduces the farmers' utility and thus affects their planting decisions, in Appendix C.5, the characteristics of the equilibrium planting amount, such as the ones described in Proposition 1 and Proposition 4, are preserved in this new framework. In addition, the government's objective function, the net benefit, remains unchanged for any value of ξ. Therefore, our critical results regarding to the optimal policy design continue to hold structurally.

Moreover, an additional insight can be generated by comparing the optimal subsidy amount under different ξ. Through extensive numerical analysis, we find that when the index accuracy is low, a higher value of ξ leads to a lower optimal payment s; however, when the index accuracy is high, a higher value of ξ leads to a higher optimal payment. Figure C.2 illustrates a representative group of results.[15]

Figure C.2 Optimal Subsidy Amount under Different Premium Fraction



(a) $r = 0.6$ (b) $r = 0.9$

C.5 Proofs of Analytical Results in Appendix C

Proof of Proposition C.1: As discussed in §8.1, if $r \le 1 - 2\alpha$ (which is equivalently to $\frac{1-r}{2} \ge \alpha$), we have $CVaR_\alpha(h|x,s) = (a - bx(\mu - \sigma)))(\mu - \sigma) - h$. If $r > 1 - 2\alpha$ (which is equivalently to $\frac{1-r}{2} < \alpha$), we have $CVaR_\alpha(h|x,s) = (a - bx(\mu - \sigma)))(\mu - \sigma) - h + \frac{(\alpha - \frac{1-r}{2})s}{\alpha}$. We next characterize the optimal subsidy amount by considering $r \le 1 - 2\alpha$ and $r > 1 - 2\alpha$, respectively.

First, consider $r \le 1 - 2\alpha$. In this case, because there is a significant probability that farmers suffer from low crop yield without receiving subsidy, increasing the subsidy amount s cannot increase the lower tail conditional expected payoff. Therefore, in this range of r, we have $s^* = 0$.

Second, consider $r > 1 - 2\alpha$. In this case, following a similar analysis as for the base model, it can be shown that the equilibrium planting amount $x^*(s)$ must satisfy either $x^*(s) = 1$ or

$$(a - bx^*(s)(\mu - \sigma)))(\mu - \sigma) - cx^*(s) + \frac{\left(\alpha - \frac{1-r}{2}\right)s}{\alpha} = 0.$$

[15] All the parameters take the same values as in §7, where $\mu = 3.029$, $\sigma_0 = 0.418$, $a = 467.32$, $b = 18.9$, $c = 1750.2$, $\lambda = 0.01$.

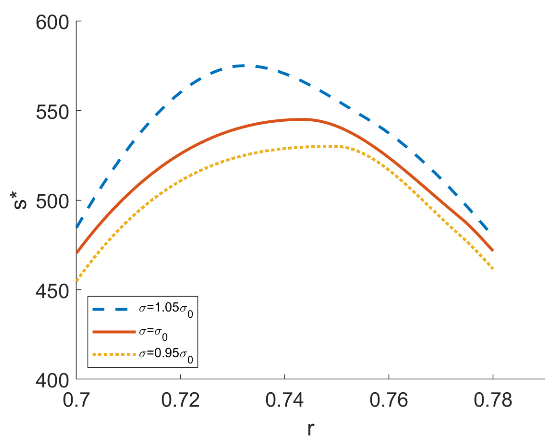
Then, the equilibrium planting amount is given by

$$x^*(s) = \min\left\{ \frac{a(\mu-\sigma) + \frac{\left(\alpha - \frac{1-r}{2}\right)s}{\alpha}}{b(\mu-\sigma)^2 + c}, 1 \right\}.$$

From this expression, it is clear that $x^*(s)$ increases in s, and it increases at a faster rate when r is higher. Further, since $x^*(s)$ strictly increases in s until it hits one, we must have $x^*(s) \geq x^{opt}$ when s is sufficiently high. Recall that the optimal subsidy amount s^* must satisfy that the corresponding equilibrium planting amount $x^*(s^*)$ is equal to or as close to x^{opt} as possible. Therefore, for $r > 1 - 2\alpha$, the optimal subsidy amount s^* must decrease in r to ensure that $x^*(s^*) = x^{opt}$.

To summarize, when r increases from a small number, the optimal subsidy amount s^* initially remains zero, then jumps to a positive number, and then decreases in r, which demonstrates that the non-monotonic effect of index accuracy on the optimal subsidy amount continues to hold under the CVaR model. Further, we remark that the main reason why there is a jump in the optimal subsidy amount as a function of r is that the crop yield follows a two-point distribution. We numerically verify that, when the crop yield follows a continuous distribution as in §7, the optimal subsidy also becomes continuous in r, and it first increases and then decreases in r. Figure C.3 illustrates a representative group of results. $\square$

Figure C.3 Optimal Subsidy Amount under Different Index Accuracies with 5% CVaR



Proof of Proposition C.2: Since including the consumer surplus in the net benefit does not affect the farmers' utility, the equilibrium planting amount and all its characteristics remain unchanged. Therefore, the results before Theorem 1 are the same.

For Theorem 1, the government has a new objective function including the consumer surplus. Now, the net benefit is, as shown in §8.2,

$$
\begin{aligned}
v_c(s) = &\int_0^{x^*(s)} \mathbb{E}[(a - bx^*(s)Y)Y + sI - xc] + \mathbb{E}[(a - bxY)Y - (a - bx^*(s)Y)Y]\,\mathrm{d}x \\
&- x^*(s)\mathbb{E}[sI] \\
=& a\mu \times x^*(s) - \frac{1}{2}c \times (x^*(s))^2
\end{aligned}
$$

Let $x_c^{opt} = \min\{\frac{a\mu}{c}, 1\}$ be the planting amount that maximizes $v_c(s)$. By replacing $v(s)$ with $v_c(s)$ and x^{opt} with x_c^{opt} in the proof of Theorem 1, as the results in Proposition 1 are preserved, we can still show that there must exist a threshold r_2 such that s^* increases in r if $r < r_2$ and decreases if $r \geq r_2$. Note that in this extension, if $x^*(s) < x^{opt}$ for any $r \in [\frac{1}{2}, 1]$ and $s \in [0, \bar{s}]$, then let $r_2 = r_1$; if $x^*(s) = x^{opt}$ for some $r \in [\frac{1}{2}, 1]$ and $s \in [0, \bar{s}]$, we still have $r_2 = \min\{r_c, r_1\}$, where r_c is defined as the lowest index accuracy under which $x^*(s) = x^{opt}$ for some $s \in [0, \bar{s}]$.

Moreover, recall that, as discussed in §3, the optimal s^* must make $x^*(s)$ as close to x^{opt} (x_c^{opt} in this extension) as possible. Since we have

$$
x_c^{opt} = \min\left\{\frac{a\mu}{c}, 1\right\} \geq \min\left\{\frac{a\mu}{2b(\mu^2 + \sigma^2) + c}\right\} = x^{opt},
$$

the optimal subsidy amount increases as we consider the consumer surplus.

Lemma A.4 still holds with x_c^{opt} because we have $\frac{\partial}{\partial \sigma} x_c^{opt} = 0$. In addition, all the second derivatives of $x^*(s)$ remain the same. Thus, through the same analysis as in Appendix A, Proposition 3 still stands.

Similarly, both Lemma A.5 and Proposition 4 remain unchanged. For Theorem 2, similar to the earlier discussion on Theorem 1, the insight still stands by replacing $v(s)$ with $v_c(s)$.

Finally, for Proposition 5 and Theorem 3, since $x^*(s)$ remains the same in this extension, the results continue to hold by going through the same analysis as in the proofs of Proposition 5 and Theorem 3 in Appendix A and replacing x^{opt} with x_c^{opt}. □

Proof of Proposition C.3: Consider a fixed $p \in [0, 1]$ and a fixed $s \in [0, 2a\sigma]$. With the joint probabilities in Table 2, we have

$$
\begin{aligned}
u(xc|x, s) &= \mathbb{E}[(a - bxY)Y + sI] - xc - \lambda\mathbb{V}\mathrm{ar}[(a - bxY)Y + sI] \\
&= -\alpha x^2 - \beta_d(s)x - \gamma_d(s),
\end{aligned}
$$

where $\alpha = 4\lambda b^2\sigma^2\mu^2$, $\beta_d(s) = b(\mu^2 + \sigma^2) + c - 4\lambda ab\sigma^2\mu + 4\lambda b(2rp - p)\sigma\mu s$ and $\gamma_d(s) = \lambda a^2\sigma^2 - a\mu - 2\lambda a(2rp - p)\sigma s - ps + \lambda p(1 - p)s^2$. Then, according to the logic introduced in the proof of Lemma 2, we generate a unique equilibrium planting amount $x_d^*(s)$ based on the solution to $u(xc|x, s) = 0$, which is

$$
\hat{x}_d(s) = \frac{-\beta_d + \sqrt{\beta_d^2 - 4\alpha\gamma_d}}{2\alpha}.
$$

Further, we have $\bar{s} \leq 2a\sigma$. Thus, given any $s \in [0, \bar{s}]$, the unique equilibrium planting amount is $x_d^*(s)$.

Since Equation (A.6) still holds for $u(xc|x,s)$ and $x_d^*(s)$, Lemma A.1 remains true. Through the same analysis in the proof of Lemma A.2, we can show $\beta_d^2(s) - 4\alpha\gamma_d(s)$ is a quadratic concave function in s and thus $x_d^*(s)$ is concave in s. Then, through the same analysis in the proof in Appendix A, the results of Proposition 1 structurally hold for $x_d^*(s)$. In addition, the government's objective is still to optimize $v(s)$ defined in Equation (4). Therefore, Theorem 1 can be proven by the same analysis as in Appendix A.

For Proposition 3, since the inclusion of p does not structurally affect how σ and λ impact the equilibrium planting amount respectively, all the relevant first and second order derivatives of $x_d^*(s)$ keep the same sign. Hence, Proposition 3 continues to hold.

For Lemma A.5, the new equilibrium planting amount in (i), (ii) and (iii) in Appendix A can be achieved by replacing $\frac{1}{2}s$ with ps, $\frac{1}{4}\lambda s^2$ with $\lambda p(1-p)s^2$, and $r - \frac{1}{2}$ with $2rp - p$. Then, as we can prove that how m affects the equilibrium planting amount structurally remains the same as in the base model, through the same analysis in its proof in Appendix A, the results of Proposition 4 are preserved. Furthermore, as the government's objective function is still $v(s)$, Theorem 2 continues to hold.

For Proposition 5 and Theorem 3, since all the relevant characteristics of $x_d^*(s)$ remain the same as $x^*(s)$ and the objective function of the government remains the same as in §6, the statements in both propositions can be verified through the same analysis as in Appendix A with $x_d^*(s)$ instead of $x^*(s)$. □

Proof of Proposition C.4: Consider a fixed $\xi \in [0,1]$ and a $s \in [0, 2a\sigma]$. Let $u(h|x,s)$ to denote the farmer's utility introduced in §8.4 and we have

$$\begin{aligned} u(xc|x,s) &= \mathbb{E}[(a - bxY)Y + sI - h - \xi\mathbb{E}[sI]] - \lambda\mathbb{V}\text{ar}[(a - bxY)Y + sI - h - \xi\mathbb{E}[sI]] \\ &= \mathbb{E}[(a - bxY)Y - xc + sI] - \xi\mathbb{E}[sI] - \lambda\mathbb{V}\text{ar}[(a - bxY)Y + sI] \\ &= -\alpha x^2 - \beta(s)x - \gamma_p(s), \end{aligned}$$

where $\alpha = 4\lambda b^2\sigma^2\mu^2$, $\beta(s) = b(\mu^2 + \sigma^2) + c - 4\lambda ab\sigma^2\mu + 4\lambda b(r - \frac{1}{2})\sigma\mu s$ and $\gamma_p(s) = \lambda a^2\sigma^2 - a\mu - 2\lambda a(r - \frac{1}{2})\sigma s + \frac{1}{4}\lambda s^2 - \frac{1}{2}(1-\xi)s$. Then, according to the logic introduced in the proof of Lemma 2, we generate a unique equilibrium planting amount $x_p^*(s)$ based on the solution to $u(xc|x,s) = 0$, which is

$$\hat{x}_p(s) = \frac{-\beta(s) + \sqrt{\beta(s)^2 - 4\alpha\gamma_p(s)}}{2\alpha},$$

Further, we have $\bar{s} \leq 2a\sigma$. Thus, given any $s \in [0, \bar{s}]$, the unique equilibrium planting amount is $x_p^*(s)$.

Since Equation (A.6) still holds for $u(xc|x,s)$ and $x_p^*(s)$, Lemma A.1 remains true. Through the same analysis in the proof of Lemma A.2, we can show $\beta(s)^2 - 4\alpha\gamma_p(s)$ is a quadratic concave function in s and thus $x_p^*(s)$ is concave in s. Then, through the same analysis in its proof in Appendix A, the results of Proposition 1 structurally hold for $x_p^*(s)$. In addition, the government's objective is still to optimize $v(s)$ defined in Equation (4). Therefore, Theorem 1 can be proven by the same analysis as in Appendix A.

For Proposition 3, since the inclusion of ξ does not structurally affect how σ and λ impact the equilibrium planting amount respectively, all the relevant first and second order derivatives of $x_p^*(s)$ keep the same sign. Hence, Proposition 3 continues to hold.

With a price protection, the utility of the farmer with production cost h is

$$u_m(h|x,s) = \mathbb{E}[\max\{(a-bxY),m\}Y + sI - \xi\mathbb{E}[sI] - h] - \lambda\mathbb{V}\text{ar}[\max\{(a-bxY),m\}Y + sI - \xi\mathbb{E}[sI] - h].$$

Following the same analysis in the proof of Lemma A.5, we can get the equilibrium planting amount $x_{mp}^*(s)$: There exist two thresholds m_1 and m_2 such that

(i): If $m \in [0, m_1)$, then we have $x_{mp}^*(s) = x_p^*(s)$

(ii): If $m \in [m_1, m_2]$, then we have

$$x_{mp}^*(s) = \min\left\{\frac{-\beta_m(s) + \sqrt{\beta_m^2(s) - 4\alpha_m\gamma_{mp}}}{2\alpha_m}, 1\right\},$$

where $\alpha_m = \frac{1}{4}\lambda b^2(\mu-\sigma)^4$, $\beta_m(s) = \frac{1}{2}b(\mu-\sigma)^2 + c + \frac{1}{2}\lambda b(\mu-\sigma)^2(m(\mu+\sigma) - a(\mu-\sigma)) - \lambda b(r - \frac{1}{2})s(\mu-\sigma)^2$ and $\gamma_{mp}(s) = \frac{1}{4}\lambda(m(\mu+\sigma) - a(\mu-\sigma))^2 - \frac{1}{2}m(\mu+\sigma) - \frac{1}{2}a(\mu-\sigma) - \lambda(r-\frac{1}{2})s(m(\mu+\sigma) - a(\mu-\sigma)) - \frac{1}{2}(1-\xi)s + \lambda\frac{1}{4}s^2$.

(iii): If $m \in (m_2, \bar{m}]$, then we have

$$x_{mp}^*(s) = \min\left\{\frac{m\mu - \lambda m^2\sigma^2 - \frac{1}{4}\lambda s^2 + 2\lambda\sigma(r-\frac{1}{2})sm}{c}, 1\right\}.$$

For Proposition 4, as we can prove that how m affects the equilibrium planting amount structurally remains the same as in the base model, through the same analysis in its proof in Appendix A, the results of Proposition 4 are preserved. Furthermore, as the government's objective function is still $v(s)$, Theorem 2 continues to hold.

For Proposition 5 and Theorem 3, since all the relevant characteristics of $x_p^*(s)$ remain the same as $x^*(s)$ and the objective function of the government remains the same as in §6, the statements in both propositions can be verified through the same analysis as in Appendix A with $x_p^*(s)$ instead of $x^*(s)$. □