ONLINE APPENDIX TO "CROP PLANNING IN SUSTAINABLE AGRICULTURE: DYNAMIC FARMLAND ALLOCATION IN THE PRESENCE OF CROP ROTATION BENEFITS"

Onur Boyabatlı

Lee Kong Chian School of Business, Singapore Management University

oboyabatli@smu.edu.sg

Javad Nasiry

School of Business and Management, Hong Kong University of Science and Technology

nasiry@ust.hk

Yangfang (Helen) Zhou

Lee Kong Chian School of Business, Singapore Management University

helenzhou@smu.edu.sg

A Proofs for Technical Statements

We use the following notation and results throughout this section. Let $\phi(.)$ and $\Phi(.)$ denote the probability density function and the cumulative distribution function of the standard normal random variable, respectively. Let $\ln x$ denotes the natural logarithm of x. We use the following identities based on the standard normal distribution $\phi^{'}(z) = -z\phi(z)$ and $\phi(z) = \phi(-z)$. We also use the following result from Cain (1994):

Lemma 1 *Let $(\tilde{X}_1, \tilde{X}_2)$ follow a bivariate normal distribution with mean vector $\boldsymbol{\mu} = (\mu_1, \mu_2)$, and covariance matrix $\mathbf{\Sigma}$ where $\mathbf{\Sigma}_{jj} = \sigma_j^2$ for $j = 1, 2$ and $\mathbf{\Sigma}_{12} = \rho\sigma_1\sigma_2$ and ρ denotes the correlation coefficient.*

$$\mathbb{E}[\max(\tilde{X}_1, \tilde{X}_2)] = \mu_1\Phi\left(\frac{\mu_1 - \mu_2}{\theta}\right) + \mu_2\Phi\left(\frac{\mu_2 - \mu_1}{\theta}\right) + \theta\phi\left(\frac{\mu_2 - \mu_1}{\theta}\right),$$

where $\theta \doteq \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$.

We also use the following result, the proof of which is omitted for brevity.

Lemma 2 *For $-1 < y \leq 0$, $\frac{y}{2}\left(\frac{2+y}{1+y}\right) \leq \ln(1+y)$.*

Proof of Proposition 1: See the proof of the T-period problem in Proposition 8. ■

Proof of Proposition 2: See the proof of the T-period problem in Proposition 8. ■

Proof of Proposition 3: The expressions for $K_2^j(\boldsymbol{r}_1)$ are obtained from (4) by using $\mathbb{E}_2[\tilde{r}_2^j] = e^{-\kappa^j} r_1^j + \left(1 - e^{-\kappa^j}\right) \xi^j$ for $j \in \{c, s\}$. Given $\boldsymbol{r}_0 = (r_0^c, r_0^s)$, $(\tilde{r}_1^c, \tilde{r}_1^s)$ follow a bivariate normal distribution with mean vector (μ_1^c, μ_1^s), and covariance matrix $\boldsymbol{\Sigma}$ with $\boldsymbol{\Sigma}_{11} = \text{VAR}_1^c$, $\boldsymbol{\Sigma}_{22} = \text{VAR}_1^s$, and $\boldsymbol{\Sigma}_{12} = \text{COV}_1$ where $\mu_1^j \doteq \mathbb{E}[\tilde{r}_1^j \mid \boldsymbol{r}_0]$, $\text{VAR}_1^j \doteq \text{VAR}[\tilde{r}_1^j \mid \boldsymbol{r}_0]$, and $\text{COV}_1 \doteq \text{COV}[\tilde{r}_1^c, \tilde{r}_1^s \mid \boldsymbol{r}_0]$ are given by (8) with $t = 1$ and $\hat{t} = 0$. Therefore, $(\underline{u}\tilde{r}_1^j + \underline{v}, \overline{u}\tilde{r}_1^{(-j)} + \overline{v})$ follow a bivariate normal distribution with mean vector $(\underline{u}\mu_1^j + \underline{v}, \overline{\mu}_1^{(-j)} + \overline{v})$ and covariance matrix $\hat{\boldsymbol{\Sigma}}$ with $\hat{\boldsymbol{\Sigma}}_{11} = \underline{u}^2\text{VAR}_1^j$, $\hat{\boldsymbol{\Sigma}}_{22} = \overline{u}^2\text{VAR}_1^{(-j)}$, and $\hat{\boldsymbol{\Sigma}}_{12} = \underline{u}\overline{u}\text{COV}_1$. The identity in (9) follows from Lemma 1. ■

Proof of Proposition 4: From Proposition 2, $V_1(\alpha_0, \boldsymbol{r}_0) = \alpha_0 K_1^c(\boldsymbol{r}_0) + (1 - \alpha_0) K_1^s(\boldsymbol{r}_0)$ where $K_1^j(\boldsymbol{r}_0)$ for $j \in \{c, s\}$ is as given in (6). Under the bivariate normal distribution specified in (8), $\mathbb{E}_1[\tilde{r}_1^j] = e^{-\kappa^j} r_0^j + \left(1 - e^{-\kappa^j}\right) \xi^j$ and, thus, $K_1^j(\boldsymbol{r}_0)$ is impacted by ρ through its effect on $\mathbb{E}_1[K_2^c(\tilde{\boldsymbol{r}}_1)]$ and $\mathbb{E}_1[K_2^s(\tilde{\boldsymbol{r}}_1)]$. From Proposition 3, $\mathbb{E}_1[K_2^j(\tilde{\boldsymbol{r}}_1)] = \mathbb{E}_1[\max\{\underline{u}\tilde{r}_1^j + \underline{v}, \overline{u}\tilde{r}_1^{(-j)} + \overline{v}\}]$ for $j \in \{c, s\}$ where $\underline{u} = e^{-\kappa^j}$, $\underline{v} = \left(1 - e^{-\kappa^j}\right) \xi^j - \omega^j$, $\overline{u} = (1 + b^{(-j)}) e^{-\kappa^{(-j)}}$, and $\overline{v} = (1 + b^{(-j)}) \left(1 - e^{-\kappa^{(-j)}}\right) \xi^{(-j)} - (1 - \gamma^{(-j)})\omega^{(-j)}$. Let $\lambda = \sqrt{\underline{u}^2\text{VAR}_1^j + \overline{u}^2\text{VAR}_1^{(-j)} - 2\underline{u}\overline{u}\text{COV}_1}$ where VAR_1^j, $\text{VAR}_1^{(-j)}$, and COV_1 are given by (8) with $t = 1$ and $\hat{t} = 0$. Using the identity in (9), we obtain

$$\frac{\partial \mathbb{E}_1[K_2^j(\tilde{\boldsymbol{r}}_1)]}{\partial \rho} = \frac{\partial \mathbb{E}_1[\max\{\underline{u}\tilde{r}_1^j + \underline{v}, \overline{u}\tilde{r}_1^{(-j)} + \overline{v}\}]}{\partial \rho} = \phi\left(\frac{\overline{u}\mu_1^{(-j)} + \overline{v} - \underline{u}\mu_1^j - \underline{v}}{\lambda}\right)\frac{\partial \lambda}{\partial \rho},$$

where $\frac{\partial \lambda}{\partial \rho} = -\frac{\underline{u}\overline{u}}{\lambda}\frac{\partial \text{COV}_1}{\partial \rho} < 0$ because $\lambda > 0$ and $\frac{\partial \text{COV}_1}{\partial \rho} = \left(\frac{1 - e^{-(\kappa^c + \kappa^s)}}{\kappa^c + \kappa^s}\right)\sigma^c\sigma^s > 0$. Therefore, $\frac{\partial \mathbb{E}_1[K_2^j(\tilde{\boldsymbol{r}}_1)]}{\partial \rho} < 0$, and thus, $\frac{\partial K_1^j(\tilde{\boldsymbol{r}}_1)}{\partial \rho} < 0$. It follows that $\frac{\partial V_1(\alpha_0, \boldsymbol{r}_0)}{\partial \rho} < 0$. ■

Proof of Proposition 5: Paralleling the proof of Proposition 4, it is easy to establish that

$$\begin{aligned}
K_1^c(\boldsymbol{r}_0) &\doteq \max\left\{-\omega^c + \mathbb{E}_1[\tilde{r}_1^c] + \mathbb{E}_1[K_2^c(\tilde{\boldsymbol{r}}_1)], -(1 - \gamma^s)\omega^s + \mathbb{E}_1[(1 + b^s)\tilde{r}_1^s] + \mathbb{E}_1[K_2^s(\tilde{\boldsymbol{r}}_1)]\right\}, \\
K_1^s(\boldsymbol{r}_0) &\doteq \max\left\{-(1 - \gamma^c)\omega^c + \mathbb{E}_1[(1 + b^c)\tilde{r}_1^c] + \mathbb{E}_1[K_2^c(\tilde{\boldsymbol{r}}_1)], -\omega^s + \mathbb{E}_1[\tilde{r}_1^s] + \mathbb{E}_1[K_2^s(\tilde{\boldsymbol{r}}_1)]\right\}
\end{aligned}$$

are impacted by σ^j for $j \in \{c, s\}$ through its effect on $\mathbb{E}_1[K_2^c(\tilde{\boldsymbol{r}}_1)]$ and $\mathbb{E}_1[K_2^s(\tilde{\boldsymbol{r}}_1)]$. To prove the impact of σ^j on $V_1(\alpha_0, \boldsymbol{r}_0)$ we will establish the following statements:

(i) $\mathbb{E}_1[K_2^c(\tilde{\boldsymbol{r}}_1)]$ and $\mathbb{E}_1[K_2^s(\tilde{\boldsymbol{r}}_1)]$ first decrease then increase in σ^j,

(ii) $\mathbb{E}_1[K_2^c(\tilde{\boldsymbol{r}}_1)]$ and $\mathbb{E}_1[K_2^s(\tilde{\boldsymbol{r}}_1)]$ are convex in σ^j.

Because convexity is preserved under maximization, (ii) implies that $K_1^c(\boldsymbol{r}_0)$ and $K_1^s(\boldsymbol{r}_0)$ are also convex in σ^j, and thus, $V_1(\alpha_0, \boldsymbol{r}_0) = \alpha_0 K_1^c(\boldsymbol{r}_0) + (1-\alpha_0)K_1^s(\boldsymbol{r}_0)$ is also convex in σ^j. Moreover, it follows from (i) that $\mathbb{E}_1[K_2^c(\tilde{\boldsymbol{r}}_1)]$ and $\mathbb{E}_1[K_2^s(\tilde{\boldsymbol{r}}_1)]$, and thus, $V_1(\alpha_0, \boldsymbol{r}_0)$ decrease in σ^j for sufficiently small σ^j and increase in σ^j for sufficiently large σ^j. Because $V_1(\alpha_0, \boldsymbol{r}_0)$ is convex in σ^j, there exists a unique $\widehat{\sigma}^j$ such that $\frac{\partial V_1(\alpha_0, \boldsymbol{r}_0)}{\partial \sigma^j} \leq 0$ if $\sigma^j \leq \widehat{\sigma}^j$, and $\frac{\partial V_1(\alpha_0, \boldsymbol{r}_0)}{\partial \sigma^j} \geq 0$ if $\sigma^j \geq \widehat{\sigma}^j$.

We now provide the proof for (i). Following similar steps with Proposition 4, we obtain

$$\frac{\partial \mathbb{E}_1[K_2^j(\tilde{\boldsymbol{r}}_1)]}{\partial \sigma^j} = \frac{\partial \mathbb{E}_1[\max\{\underline{u}\tilde{r}_1^j + \underline{v}, \overline{u}\tilde{r}_1^{(-j)} + \overline{v}\}]}{\partial \sigma^j} = \phi\left(\frac{\overline{u}\mu_1^{(-j)} + \overline{v} - \underline{u}\mu_1^j - \underline{v}}{\lambda}\right)\frac{\partial \lambda}{\partial \sigma^j}, \quad (15)$$

where

$$\frac{\partial \lambda}{\partial \sigma^j} = \frac{1}{\lambda}\left[\underline{u}^2\left(\frac{1-e^{-2\kappa^j}}{2\kappa^j}\right)\sigma^j - \underline{u}\overline{u}\left(\frac{1-e^{-(\kappa^c+\kappa^s)}}{\kappa^c+\kappa^s}\right)\rho\sigma^{(-j)}\right]. \quad (16)$$

The term inside the bracket in (16) can be written as $A\sigma^j - B$, where $A \doteq \underline{u}^2\left(\frac{1-e^{-2\kappa^j}}{2\kappa^j}\right) > 0$ and $B \doteq \underline{u}\overline{u}\left(\frac{1-e^{-(\kappa^c+\kappa^s)}}{\kappa^c+\kappa^s}\right)\rho\sigma^{(-j)} > 0$. Therefore, there exists a unique $\sigma_0^j \doteq \frac{B}{A}$ threshold such that $\frac{\partial \mathbb{E}_1[K_2^j(\tilde{\boldsymbol{r}}_1)]}{\partial \sigma^j} < 0$ for $\sigma^j < \sigma_0^j$ and $\frac{\partial \mathbb{E}_1[K_2^j(\tilde{\boldsymbol{r}}_1)]}{\partial \sigma^j} > 0$ for $\sigma^j > \sigma_0^j$. Following similar steps, it can be established that there exists a unique σ_1^j threshold such that $\frac{\partial \mathbb{E}_1[K_2^{(-j)}(\tilde{\boldsymbol{r}}_1)]}{\partial \sigma^j} < 0$ for $\sigma^j < \sigma_1^j$ and $\frac{\partial \mathbb{E}_1[K_2^{(-j)}(\tilde{\boldsymbol{r}}_1)]}{\partial \sigma^j} > 0$ for $\sigma^j > \sigma_1^j$.

To conclude, we now provide the proof for (ii). We will only show the convexity of $\mathbb{E}_1[K_2^j(\tilde{\boldsymbol{r}}_1)]$ in σ^j, that is, $\frac{\partial^2 \mathbb{E}_1[K_2^j(\tilde{\boldsymbol{r}}_1)]}{\partial (\sigma^j)^2} \geq 0$; the convexity of $\mathbb{E}_1[K_2^{(-j)}(\tilde{\boldsymbol{r}}_1)]$ in σ^j can be established in a similar fashion. For expositional purposes, besides A and B defined above, we also define $a \doteq \overline{u}\mu_1^{(-j)} + \overline{v} - \underline{u}\mu_1^j - \underline{v}$ and $C \doteq \overline{u}^2\left(\frac{1-e^{-2\kappa^{(-j)}}}{2\kappa^{(-j)}}\right)(\sigma^{(-j)})^2$. Using this notation, it follows from (15) and (16) that

$$\frac{\partial \mathbb{E}_1[K_2^j(\tilde{\boldsymbol{r}}_1)]}{\partial \sigma^j} = \frac{\partial \mathbb{E}_1[\max\{\underline{u}\tilde{r}_1^j + \underline{v}, \overline{u}\tilde{r}_1^{(-j)} + \overline{v}\}]}{\partial \sigma^j} = \phi\left(\frac{a}{\lambda}\right)\left(\frac{A\sigma^j - B}{\lambda}\right),$$

where $\lambda = \sqrt{A(\sigma^j)^2 + C - 2B\sigma^{(-j)}}$. We obtain

$$\frac{\partial^2 \mathbb{E}_1[K_2^j(\tilde{\boldsymbol{r}}_1)]}{\partial (\sigma^j)^2} = \phi\left(\frac{a}{\lambda}\right)\frac{1}{\lambda}\left[\frac{A\sigma^j - B}{\lambda}\right]^2\left[\frac{a}{\lambda}\right]^2 + \phi\left(\frac{a}{\lambda}\right)\frac{1}{\lambda}\left[A - \left(\frac{A\sigma^j - B}{\lambda}\right)^2\right].$$

To prove $\frac{\partial^2 \mathbb{E}_1[K_2^j(\tilde{\boldsymbol{r}}_1)]}{\partial (\sigma^j)^2} \geq 0$, it is sufficient to show that $A - \left(\frac{A\sigma^j - B}{\lambda}\right)^2 \geq 0$. Using the definitions of A, B, and λ, it is easy to establish that this condition is equivalent to

$$\left(\frac{1-e^{-2\kappa^c}}{2\kappa^c}\right)\left(\frac{1-e^{-2\kappa^s}}{2\kappa^s}\right) \geq \left(\frac{1-e^{-(\kappa^c+\kappa^s)}}{\kappa^c+\kappa^s}\right)^2 \rho^2.$$

Because $\rho \leq 1$, it is sufficient to show

$$\left(\frac{1-e^{-2\kappa^c}}{2\kappa^c}\right)\left(\frac{1-e^{-2\kappa^s}}{2\kappa^s}\right) \geq \left(\frac{1-e^{-(\kappa^c+\kappa^s)}}{\kappa^c+\kappa^s}\right)^2. \tag{17}$$

The condition in (17) is satisfied with equality for $\kappa^c = \kappa^s$; therefore, we focus on the case $\kappa^c \neq \kappa^s$ hereafter. We define $u \doteq e^{-\kappa^c}$ and $v \doteq e^{-\kappa^s}$ where $0 < u, v < 1$, $\kappa^c = -\ln u$, and $\kappa^s = -\ln v$. After some algebra, (17) can be equivalently written as

$$\left(\frac{1-uv}{\ln uv}\right)^2 \geq \left(\frac{u-v}{\ln u - \ln v}\right)^2. \tag{18}$$

In (18), the right-hand side $l \doteq \frac{u-v}{\ln u - \ln v}$ is the logarithmic mean of u and v. It follows from the ordering among the geometric, logarithmic, and arithmetic means (see, e.g., Burk 1987) that $\sqrt{uv} \leq l \leq \frac{u+v}{2}$. Using $\sqrt{uv} \leq l$, to prove (18), it suffices to show

$$\left(\frac{1-l^2}{\ln l^2}\right)^2 \geq l^2. \tag{19}$$

Because $u, v < 1$, we have $\frac{u+v}{2} < 1$ and so $l \leq \frac{u+v}{2} < 1$ from the ordering of logarithmic and arithmetic means. Therefore, $1 - l^2 > 0$ and $\ln l < 0$. It follows that the condition in (19) is equivalent to $\ln l \geq \frac{l^2-1}{2l}$, which is satisfied because as follows from Lemma 2, we have

$$\ln l = \ln\left(1 + (l-1)\right) \geq \frac{l-1}{2}\left(\frac{2+(l-1)}{1+(l-1)}\right) = \frac{l^2-1}{2l},$$

where $-1 < l - 1 < 0$ as $0 < l < 1$. This concludes the proof. ■

Proof of Proposition 6: When $\kappa^c = \kappa^s = \kappa$, $\xi^c = \xi^s = \xi$, $\sigma^c = \sigma^s = \sigma$, and $r_0^c = r_0^s = r_0$, we obtain from (8) with $t = 1$ and $\hat{t} = 0$ that $\mu_1^c = \mu_1^s = \mu_1 = e^{-\kappa} r_0 + (1 - e^{-\kappa})\,\xi$, $\mathrm{VAR}_1^c = \mathrm{VAR}_1^s = \mathrm{VAR}_1 = \left(\frac{1-e^{-2\kappa}}{2\kappa}\right)\sigma^2$, and $\mathrm{COV}_1 = \left(\frac{1-e^{-2\kappa}}{2\kappa}\right)\rho\sigma^2$. Let $\lambda \doteq \sqrt{(1+(1+b)^2)e^{-2\kappa}\mathrm{VAR}_1 - 2(1+b)e^{-2\kappa}\mathrm{COV}_1} > 0$ where $b^c = b^s = b$. Recall that $\Gamma = \mathbb{E}_1\left[K_2^c(\tilde{\boldsymbol{r}}_1)\right] - \mathbb{E}_1\left[K_2^s(\tilde{\boldsymbol{r}}_1)\right]$. Paralleling the proof of Proposition 4, we obtain

$$\frac{\partial \Gamma}{\partial \rho} = \left(\frac{1+b}{\lambda}\right) e^{-2\kappa}\left(\frac{1-e^{-2\kappa}}{2\kappa}\right)\sigma^2\left[\phi\left(\frac{t^s}{\lambda}\right) - \phi\left(\frac{t^c}{\lambda}\right)\right], \tag{20}$$

where $t^j = X + \gamma^{(-j)}\omega$ for $j \in \{c, s\}$ with $\omega^c = \omega^s = \omega$ and $X \doteq b\left(e^{-\kappa}\mu_1 + (1-e^{-\kappa})\xi\right) > 0$. It follows from (20) that the sign of $\frac{\partial \Gamma}{\partial \rho}$ is determined by the sign of $\left[\phi\left(\frac{t^s}{\lambda}\right) - \phi\left(\frac{t^c}{\lambda}\right)\right]$. Because $X > 0$ (and $\gamma^{(-j)} > 0$ by definition), we have $t^j > 0$ for $j \in \{c, s\}$. Because the standard normal density function $\phi(z)$ decreases in z for $z > 0$, when $\gamma^c > \gamma^s$, we have $t^s > t^c$, and thus, $\left[\phi\left(\frac{t^s}{\lambda}\right) - \phi\left(\frac{t^c}{\lambda}\right)\right] < 0$, and $\frac{\partial \Gamma}{\partial \rho} < 0$. When $\gamma^c < \gamma^s$, we have $t^s < t^c$, and thus, $\left[\phi\left(\frac{t^s}{\lambda}\right) - \phi\left(\frac{t^c}{\lambda}\right)\right] > 0$, and $\frac{\partial \Gamma}{\partial \rho} > 0$. ■

Proof of Proposition 7: We only provide the proof for the result related to σ^c. The result related to σ^s can be established in a similar fashion. Paralleling the proof of Proposition 5, it can be established that

$$\frac{\partial \Gamma}{\partial \sigma^c} = \phi\left(\frac{t^c}{\lambda^c}\right)\left[\frac{A\sigma^c - B}{\lambda^c}\right] - \phi\left(\frac{t^s}{\lambda^s}\right)\left[\frac{C\sigma^c - D}{\lambda^s}\right], \tag{21}$$

for some scalars $\lambda^j > 0$ and $t^j > 0$ (whose definitions are omitted for brevity) for $j \in \{c, s\}$, where

$$\begin{aligned} A &\doteq e^{-2\kappa^c}\left(\frac{1-e^{-2\kappa^c}}{2\kappa^c}\right) > 0, \\ B &\doteq (1+b^s)e^{-(\kappa^c+\kappa^s)}\left(\frac{1-e^{-(\kappa^c+\kappa^s)}}{\kappa^c+\kappa^s}\right)\rho\sigma^s > 0, \\ C &\doteq (1+b^c)^2 e^{-2\kappa^c}\left(\frac{1-e^{-2\kappa^c}}{2\kappa^c}\right) > 0, \\ D &\doteq (1+b^c)e^{-(\kappa^c+\kappa^s)}\left(\frac{1-e^{-(\kappa^c+\kappa^s)}}{\kappa^c+\kappa^s}\right)\rho\sigma^s > 0. \end{aligned}$$

Let $U^c \doteq \frac{e^{-\kappa^s}}{e^{-\kappa^c}}\left(\frac{1-e^{-(\kappa^c+\kappa^s)}}{1-e^{-2\kappa^c}}\right)\left(\frac{2\kappa^c}{\kappa^c+\kappa^s}\right)\rho\sigma^s$. It follows that $\frac{B}{A} = (1+b^s)U^c$ and $\frac{D}{C} = \frac{1}{(1+b^c)}U^c$. When $\sigma^c \in \left[\frac{1}{(1+b^c)}U^c, (1+b^s)U^c\right]$, in (21), $\left[\frac{A\sigma^c - B}{\lambda^c}\right]$ is negative whereas $\left[\frac{C\sigma^c - D}{\lambda^s}\right]$ is positive, and thus, $\frac{\partial \Gamma}{\partial \sigma^c} < 0$. ■

Proof of Proposition 8: Let $\pi_t(\alpha_t) \doteq L(\alpha_t \mid \alpha_{t-1}, \boldsymbol{r}_{t-1}) + \mathbb{E}_t[V_{t+1}(\alpha_t, \tilde{\boldsymbol{r}}_t)]$ denote the objective function of the optimization problem in period $t \in [1, T]$ where the immediate payoff $L(\alpha_t \mid \alpha_{t-1}, \boldsymbol{r}_{t-1})$ is as given in (1) and the optimal value function from period $t+1$ onward $V_{t+1}(\alpha_t, \tilde{\boldsymbol{r}}_t)$ is as given in (2).

We first prove that the characterization of $V_t(\alpha_{t-1}, \boldsymbol{r}_{t-1})$ given in (13) holds for period $1, \cdots, T+1$ by induction. In period $T+1$, $V_{T+1}(\alpha_T, \boldsymbol{r}_T) = \alpha_T K^c_{T+1}(\boldsymbol{r}_T) + (1-\alpha_t)K^s_{T+1}(\boldsymbol{r}_T)$ holds trivially because $V_{T+1}(\cdot) = 0$, $K^c_{T+1}(\cdot) = 0$, and $K^s_{T+1}(\cdot) = 0$ hold by definition. Let us assume that the claimed property holds for period $t+1, \cdots, T$. We next show this property also holds for period t. We have $\pi_t(\alpha_t) \doteq L(\alpha_t \mid \alpha_{t-1}, \boldsymbol{r}_{t-1}) + \mathbb{E}_t[V_{t+1}(\alpha_t, \tilde{\boldsymbol{r}}_t)]$ where $\mathbb{E}_t[V_{t+1}(\alpha_t, \tilde{\boldsymbol{r}}_t)] = \alpha_t \mathbb{E}_t[K^c_{t+1}(\boldsymbol{r}_t)] + (1-\alpha_t)\mathbb{E}_t[K^s_{t+1}(\boldsymbol{r}_t)]$ due to the induction assumption. We obtain

$$\frac{\partial \pi_t(\alpha_t)}{\partial \alpha_t} = \begin{cases} h^1_t \doteq \mathbb{E}_t[(1+b^c)\tilde{r}^c_t + K^c_{t+1}(\boldsymbol{r}_t)] - (1-\gamma^c)\omega^c - \mathbb{E}_t[\tilde{r}^s_t + K^s_{t+1}(\boldsymbol{r}_t)] + \omega^s & \text{if } 0 \le \alpha_t \le 1 - \alpha_{t-1}, \\ h^2_t \doteq \mathbb{E}_t[\tilde{r}^c_t + K^c_{t+1}(\boldsymbol{r}_t)] - \omega^c - \mathbb{E}_t[(1+b^s)\tilde{r}^s_t + K^s_{t+1}(\boldsymbol{r}_t)] + (1-\gamma^s)\omega^s & \text{if } 1 - \alpha_{t-1} < \alpha_t \le 1, \end{cases}$$

where $h^1_t - h^2_t = \Sigma_j\left(b^j \mathbb{E}_T[\tilde{r}^j_t] + \gamma^j \omega^j\right) > 0$. Therefore, $\pi_t(\alpha_t)$ is piecewise linear and concave in α_t with a kink at $\alpha_t = 1 - \alpha_{t-1}$. We have three cases to consider:

(i) $\frac{\partial \pi_t}{\partial \alpha_t}\Big|_{0^+} \leq 0$, that is, $h_t^1 \leq 0$. In this case, $\alpha_t^* = 0$ and $\pi_t(0) = -\omega^s + \mathbb{E}_t[\tilde{r}_t^s + K_{t+1}^s(\boldsymbol{r}_t)] + \alpha_{t-1}(b^s \mathbb{E}_t[\tilde{r}_t^s] + \gamma^s \omega^s)$. Because $h_t^1 \leq 0$, $K_t^j(\boldsymbol{r}_{t-1})$ for $j \in \{c, s\}$ as defined in (12) is such that $K_t^c(\boldsymbol{r}_{t-1}) = \mathbb{E}_t[(1+b^s)\tilde{r}_t^s + K_{t+1}^s(\boldsymbol{r}_t)] - (1-\gamma^s)\omega^s$ and $K_t^s(\boldsymbol{r}_{t-1}) = \mathbb{E}_t[\tilde{r}_t^s + K_{t+1}^s(\boldsymbol{r}_t)] - \omega^s$. It is easy to establish that $\alpha_{t-1}K_t^c(\boldsymbol{r}_{t-1}) + (1-\alpha_{t-1})K_t^s(\boldsymbol{r}_{t-1})$ equals $\pi_t(0)$ and, thus, the characterization of $V_t(\alpha_{t-1}, \boldsymbol{r}_{t-1})$ as given in (13) holds.

(ii) $\frac{\partial \pi_t}{\partial \alpha_t}\Big|_{1^-} \geq 0$, that is, $h_t^2 \geq 0$. In this case, $\alpha_t^* = 1$ and $\pi_t(1) = -\omega^c + \mathbb{E}_t[\tilde{r}_t^c + K_{t+1}^c(\boldsymbol{r}_t)] + (1-\alpha_{t-1})(b^c \mathbb{E}_t[\tilde{r}_t^c] + \gamma^c \omega^c)$. Because $h_t^2 \geq 0$, $K_t^j(\boldsymbol{r}_{t-1})$ for $j \in \{c, s\}$ as defined in (12) is such that $K_t^c(\boldsymbol{r}_{t-1}) = \mathbb{E}_t[\tilde{r}_t^c + K_{t+1}^c(\boldsymbol{r}_t)] - \omega^c$ and $K_t^s(\boldsymbol{r}_{t-1}) = \mathbb{E}_t[(1+b^c)\tilde{r}_t^c + K_{t+1}^c(\boldsymbol{r}_t)] - (1-\gamma^c)\omega^c$. It is easy to establish that $\alpha_{t-1}K_t^c(\boldsymbol{r}_{t-1}) + (1-\alpha_{t-1})K_t^s(\boldsymbol{r}_{t-1})$ equals $\pi_t(1)$ and, thus, the characterization of $V_t(\alpha_{t-1}, \boldsymbol{r}_{t-1})$ as given in (13) holds.

(iii) $\frac{\partial \pi_t}{\partial \alpha_t}\Big|_{0^+} > 0$ and $\frac{\partial \pi_t}{\partial \alpha_t}\Big|_{1^-} < 0$, that is, $h_t^1 > 0$ and $h_t^2 < 0$. In this case, $\alpha_t^* = 1 - \alpha_{t-1}$ and

$$\pi_t(1-\alpha_{t-1}) = \alpha_{t-1}\left(-(1-\gamma^s)\omega^s + \mathbb{E}_t[(1+b^s)\tilde{r}_t^s + K_{t+1}^s(\boldsymbol{r}_t)]\right) + (1-\alpha_{t-1})\left(-(1-\gamma^c)\omega^c + \mathbb{E}_t[(1+b^c)\tilde{r}_t^c + K_{t+1}^c(\boldsymbol{r}_t)]\right).$$

Because $h_t^1 > 0$ and $h_t^2 < 0$, $K_t^j(\boldsymbol{r}_{t-1})$ for $j \in \{c, s\}$ as defined in (12) is such that $K_t^c(\boldsymbol{r}_{t-1}) = -(1-\gamma^s)\omega^s + \mathbb{E}_t[(1+b^s)\tilde{r}_t^s + K_{t+1}^s(\boldsymbol{r}_t)]$ and $K_t^s(\boldsymbol{r}_{t-1}) = -(1-\gamma^c)\omega^c + \mathbb{E}_t[(1+b^c)\tilde{r}_t^c + K_{t+1}^c(\boldsymbol{r}_t)]$. It follows that $\alpha_{t-1}K_t^c(\boldsymbol{r}_{t-1}) + (1-\alpha_{t-1})K_t^s(\boldsymbol{r}_{t-1})$ equals $\pi_t(1-\alpha_{t-1})$ and, thus, the characterization of $V_t(\alpha_{t-1}, \boldsymbol{r}_{t-1})$ as given in (13) holds.

Based on the principle of mathematical induction, (13) holds for all periods $1, \cdots, T+1$.

For each period $t = 1, \cdots, T$, given that (13) holds, (11) and (12) follow as seen from the steps above. ■

B Detailed Analysis of Section 6: Extensions

B.1 Calculating the Proportion of Farmland Allocated to Rotated Crops

In this section, we describe in detail how we numerically compute the (expected) proportion of farmland allocated to rotated crops per growing season under the one-period lookeahed policy and the optimal policy. To calculate this metric, we first define $R_t^H(\alpha_{t-1}, \boldsymbol{r}_{t-1})$ for $t \in [1, T]$ as the total expected proportion of farmland allocated to rotated crops over the planning horizon under the allocation policy H from period t onward given α_{t-1} and $\boldsymbol{r}_{t-1}$, which satisfies

$$R_t^H(\alpha_{t-1}, \boldsymbol{r}_{t-1}) = \min\{\alpha_t^H, 1 - \alpha_{t-1}\} + \min\{1 - \alpha_t^H, \alpha_{t-1}\} + \mathbb{E}_t[R_{t+1}^H(\alpha_t^H, \tilde{\boldsymbol{r}}_t)], \tag{22}$$

with a boundary condition $R_{T+1}^H(\cdot) = 0$, where α_t^H denotes the corn allocation in period t under the policy H. Here, H refers to the one-period lookahead heuristic policy or the optimal policy. For the one-period lookahead policy, α_t^H can be obtained from (5) of Proposition 2 by substituting $t = 1$ ($t-1 = 0$, $t+1 = 2$) with an arbitrary t ($t-1$, $t+1$), and using the identities given in Proposition 3. For the optimal policy, α_t^H follows from Proposition 8.

In (22), $\min\{\alpha_t^H, 1 - \alpha_{t-1}\}$ and $\min\{1 - \alpha_t^H, \alpha_{t-1}\}$ refer to the proportion of corn and soybeans grown on rotated farmland under the policy H in period t, respectively. The expected proportion of farmland allocated to rotated crops per growing season under the policy H is given by $\frac{R_1^H(\alpha_0, \boldsymbol{r}_0)}{T} * 100$, where α_0 and $\boldsymbol{r}_0$ denote the observed corn allocation and crop revenues at the beginning of the planning horizon, respectively. We calculate this metric using the same 312,500 numerical instances as in Section 5.2.3 and using the numerical computation procedure as described in Section 5.2.1.

B.2 Introducing Fallow Farmland

Throughout the paper, we assume that farmland is always fully allocated to the two crops in each growing season. In practice, farmers may also let the farmland lay fallow—that is, deliberately not use the farmland to grow any crop—to rejuvenate the soil and increase the revenue for the crop grown on this farmland in the subsequent seasons. In this section, we generalize our model to incorporate the farmer's decision to let some portion of the farmland lay fallow in each growing season. For ease of notation, we capture this decision through what we call a "fallow crop." In particular, the farmland that is laid fallow (where no crop is grown) in a growing season in practice corresponds to the farmland allocated to the fallow crop in that growing season in our model.

The remainder of this section is organized as follows. We describe the model in Section B.2.1. We characterize the optimal allocation policy in Section B.2.2 and demonstrate that although the optimal policy is more complex (due to a larger set of decision variables) it follows the same structure as the optimal allocation policy presented in our paper. Section B.2.3 numerically examines the value of crop planning based on principles of sustainable agriculture and compares the optimal allocation policy's performance with that of heuristic allocation policies. We demonstrate that our key insights from the numerical analysis in our paper continue to hold in the presence of a fallow crop.

B.2.1 Model Description

We consider a farmer who allocates a single acre of farmland among two cash crops (corn and soybeans) and a fallow crop in each growing season to maximize the expected total profit over a finite number of growing seasons. We use superscript c, s, and f to denote the corn, soybeans, and fallow crop related parameters, respectively.

Decision variables. Let $\alpha_t^j \in [0,1]$ for $j \in \{c,s,f\}$ denote the proportion of farmland allocated to corn in time period (growing season) t on which crop j was grown in period $t-1$. For notational convenience, we denote the total proportion of farmland allocated to corn in period t as $\alpha_t \doteq \sum_{j\in\{c,s,f\}} \alpha_t^j$, and corn allocations in period t as $\boldsymbol{\alpha}_t \doteq \left(\alpha_t^c, \alpha_t^s, \alpha_t^f\right)$. Let $\beta_t^j \in [0,1]$ for $j \in \{c,s,f\}$ denote the proportion of farmland allocated to a fallow crop in period t on which crop j was grown in period $t-1$. We denote the total proportion of farmland allocated to a fallow crop in period t as $\beta_t \doteq \sum_{j\in\{c,s,f\}} \beta_t^j$, and fallow crop allocations in period t as $\boldsymbol{\beta}_t \doteq \left(\beta_t^c, \beta_t^s, \beta_t^f\right)$. The remaining proportion of farmland, $1-\alpha_t-\beta_t$, is allocated to soybeans in period t. Within this soybean allocation, it is important to determine the proportion of farmland on which crop $j \in \{c,s,f\}$ was grown in period $t-1$. As we will discuss shortly, this can be determined without defining another three decision variables (which denote the proportion of farmland allocated to soybeans in period t on which crop $j \in \{c,s,f\}$ was grown in period $t-1$).

Revenue uncertainty. Let $\tilde{r}_t^c$ and $\tilde{r}_t^s$ denote the uncertain corn and soybean revenue per acre in period t, respectively. Because the farmland allocated to a fallow crop in each growing season in our model corresponds to the farmland that is laid fallow (where no crop is grown) in that growing season in practice, the fallow crop revenue per acre is assumed to be zero in each period. As in the paper, we assume that $\tilde{\boldsymbol{r}}_t = (\tilde{r}_t^c, \tilde{r}_t^s)$ follow correlated stochastic processes with Markovian property.

Crop rotation benefits. To capture the revenue-enhancing crop rotation benefits, we assume that the uncertain revenue per acre of (cash) crop $j \in \{c, s\}$ in period t is $\tilde{r}_t^j$ if it is grown on nonrotated farmland, $(1+b_1^j)\tilde{r}_t^j$ if it is grown on rotated farmland where the other cash crop was grown in the previous period, and $(1+b_2^j)\tilde{r}_t^j$ if it is grown on rotated farmland where a fallow crop was grown in the previous period. We assume $b_2^j \geq b_1^j \geq 0$; that is, revenue-enhancing crop rotation benefits for cash crop $j \in \{c, s\}$ are (stochastically) larger on the rotated farmland where a fallow crop was grown in the previous season than on the rotated farmland where the other cash crop was grown in the previous season. To capture the cost-reducing crop rotation benefits, we assume that the unit farming cost of cash crop $j \in \{c, s\}$ is ω^j if it is grown on nonrotated farmland, $(1 - \gamma_1^j)\omega^j$ if it is grown on rotated farmland where the other cash crop was grown in the previous period, and $(1-\gamma_2^j)\omega^j$ if it is grown on rotated farmland where a fallow crop was grown in the previous period. Similarly, we assume $\gamma_2^j \geq \gamma_1^j \geq 0$. Because the farmland allocated to the fallow crop in each growing season in our model corresponds to the farmland that is laid fallow (where no crop is grown) in that growing season in practice, we assume that there is no farming cost associated with the fallow crop in each period.

Formulation. We formulate the farmer's problem as a finite horizon stochastic dynamic program. In each period $t \in [1, T]$, the sequence of events is as follows:

(i) At the beginning of period t, the farmer observes the total corn allocation α_{t-1}, the total fallow crop allocation β_{t-1}, and corn and soybean revenues $\boldsymbol{r}_{t-1} = (r_{t-1}^c, r_{t-1}^s)$ from period $t-1$. The farmer then chooses the corn allocations $\boldsymbol{\alpha}_t = (\alpha_t^c, \alpha_t^s, \alpha_t^f)$ and the fallow crop allocations $\boldsymbol{\beta}_t = (\beta_t^c, \beta_t^s, \beta_t^f)$ constrained by the available farmland where crop $j \in \{c, s, f\}$ was grown in period $t-1$, that is, $\alpha_t^c + \beta_t^c \leq \alpha_{t-1}, \alpha_t^f + \beta_t^f \leq \beta_{t-1}, \alpha_t^s + \beta_t^s \leq 1 - \alpha_{t-1} - \beta_{t-1}$. For example, the first constraint ensures that the sum of the proportion of farmland allocated to corn and a fallow crop in this period where corn was grown in the previous period cannot be larger than the proportion of farmland where corn was grown in the previous period.

(ii) At the end of period t, the corn and soybean revenues $\tilde{\boldsymbol{r}}_t = (\tilde{r}_t^c, \tilde{r}_t^s)$ are realized and the farmer collects the revenues from the crop sales.

The farmer's immediate payoff in period $t \in [1, T]$ is given by

$$\begin{aligned} L(\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t \mid \alpha_{t-1}, \beta_{t-1}, \boldsymbol{r}_{t-1}) \doteq\ & \alpha_t^f \mathbb{E}_t\left[(1+b_2^c)\tilde{r}_t^c - (1-\gamma_2^c)\omega^c\right] + \alpha_t^s \mathbb{E}_t\left[(1+b_1^c)\tilde{r}_t^c - (1-\gamma_1^c)\omega^c\right] + \alpha_t^c \mathbb{E}_t\left[\tilde{r}_t^c - \omega^c\right] \\ &+ \min\left(1-\alpha_t-\beta_t, \beta_{t-1}-\alpha_t^f-\beta_t^f\right)\mathbb{E}_t\left[(1+b_2^s)\tilde{r}_t^s - (1-\gamma_2^s)\omega^s\right] \\ &+ \min\left((1-\alpha_t-\beta_t-(\beta_{t-1}-\alpha_t^f-\beta_t^f))^+, \alpha_{t-1}-\alpha_t^c-\beta_t^c\right)\mathbb{E}_t\left[(1+b_1^s)\tilde{r}_t^s - (1-\gamma_1^s)\omega^s\right] \\ &+ \left((1-\alpha_t-\beta_t-(\beta_{t-1}-\alpha_t^f-\beta_t^f))^+ - (\alpha_{t-1}-\alpha_t^c-\beta_t^c)\right)^+ \mathbb{E}_t\left[\tilde{r}_t^s - \omega^s\right], \end{aligned} \tag{23}$$

where $\mathbb{E}_t[.]$ denotes the expectation operator conditional on the available information at time t, that is, $\mathbb{E}_t[.] = \mathbb{E}[.|\boldsymbol{r}_{t-1}]$. In (23), the first line corresponds to the total expected profit from growing corn in period t. It is the sum of expected profit from growing corn on three different farmlands: rotated farmland where a fallow crop was grown in the previous period, rotated farmland where soybeans were grown in the previous period, and nonrotated farmland. The remaining three lines in (23) denote the total expected profit from growing soybeans in period t. For $1-\alpha_t-\beta_t$ proportion of the farmland that is allocated to soybeans, to leverage crop rotation benefits, the farmer starts planting soybeans from the rotated farmland where the fallow crop was grown in the previous period which remains (if any) from the corn and the fallow crop allocation in this period (that is given by $\beta_{t-1}-\alpha_t^f-\beta_t^f$). Therefore, rotation benefits for soybean plantation on fallow farmland in the previous period are only relevant for $\min\left(1-\alpha_t-\beta_t, \beta_{t-1}-\alpha_t^f-\beta_t^f\right)$ proportion of the farmland. For the remaining $(1-\alpha_t-\beta_t-(\beta_{t-1}-\alpha_t^f-\beta_t^f))^+$ proportion of the farmland allocated to soybeans, again to leverage crop rotation benefits, the farmer starts planting soybeans from the rotated farmland where corn was grown in the previous period which remains (if any) from the corn and the fallow crop allocation in this period (that is given by $\alpha_{t-1}-\alpha_t^c-\beta_t^c$). Therefore, rotation benefits for soybean plantation on corn farmland in the previous period are only relevant for $\min\left((1-\alpha_t-\beta_t-(\beta_{t-1}-\alpha_t^f-\beta_t^f))^+, \alpha_{t-1}-\alpha_t^c-\beta_t^c\right)$ proportion of the farmland. The remaining proportion of the farmland allocated to soybeans is from the nonrotated farmland (where soybeans were grown in the previous period), which has no rotation benefit.

Let $V_t(\alpha_{t-1}, \beta_{t-1}, \boldsymbol{r}_{t-1})$ for $t \in [1, T]$ denote the optimal value function from period t

onward given α_{t-1}, β_{t-1}, and $\boldsymbol{r}_{t-1}$, which satisfies

$$\begin{aligned} V_t(\alpha_{t-1}, \beta_{t-1}, \boldsymbol{r}_{t-1}) = \max_{\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t} \quad & \left\{ L(\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t \mid \alpha_{t-1}, \beta_{t-1}, \boldsymbol{r}_{t-1}) + \mathbb{E}_t \left[V_{t+1}(\alpha_t = \sum_{j \in \{c,s,f\}} \alpha_t^j, \beta_t = \sum_{j \in \{c,s,f\}} \beta_t^j, \tilde{\boldsymbol{r}}_t) \right] \right\} \\ \text{s.t.} \quad & \alpha_t^c + \beta_t^c \leq \alpha_{t-1}, \alpha_t^f + \beta_t^f \leq \beta_{t-1}, \alpha_t^s + \beta_t^s \leq 1 - \alpha_{t-1} - \beta_{t-1}, \\ & 0 \leq \alpha_t^j \leq 1, 0 \leq \beta_t^j \leq 1 \text{ for } j \in \{c, s, f\}, \end{aligned} \tag{24}$$

with a boundary condition $V_{T+1}(\cdot) = 0$. The farmer's optimal total expected profit over the entire planning horizon is given by $V_1(\alpha_0, \beta_0, \boldsymbol{r}_0)$, where α_0, β_0, and $\boldsymbol{r}_0$ denote the observed corn allocation, fallow crop allocation, and crop revenues at the beginning of the planning horizon, respectively. It is easy to establish that in the absence of a fallow crop, the optimal value function above is identical to the optimal value function presented in (2) of our paper. In this case, $\beta_{t-1} = 0$, $\beta_t^j = 0$, and $\alpha_t^f = 0$, and the corn allocation decision in period t can be captured through α_t using $\alpha_t^s = \min(\alpha_t, 1 - \alpha_{t-1})$ and $\alpha_t^c = (\alpha_t - (1 - \alpha_{t-1}))^+$.

B.2.2 Optimal Allocation Policy

We now solve for the farmer's optimization problem stated in (24) and characterize the optimal allocation decision and the optimal value function in period $t \in [1, T]$. For this purpose, similar to the analysis in our paper, we first define the following recursive operators:

$$\begin{aligned} K_t^c(\boldsymbol{r}_{t-1}) &= \max\left\{C_t^{(0)}, S_t^{(1)}, F_t\right\}, \\ K_t^s(\boldsymbol{r}_{t-1}) &= \max\left\{C_t^{(1)}, S_t^{(0)}, F_t\right\}, \\ K_t^f(\boldsymbol{r}_{t-1}) &= \max\left\{C_t^{(2)}, S_t^{(2)}, F_t\right\}, \end{aligned} \tag{25}$$

where

$$\begin{aligned} C_t^{(0)} &\doteq -\omega^c + \mathbb{E}_t\left[\tilde{r}_t^c + K_{t+1}^c(\tilde{\boldsymbol{r}}_t)\right], \\ C_t^{(1)} &\doteq -(1-\gamma_1^c)\omega^c + \mathbb{E}_t\left[(1+b_1^c)\tilde{r}_t^c + K_{t+1}^c(\tilde{\boldsymbol{r}}_t)\right], \\ C_t^{(2)} &\doteq -(1-\gamma_2^c)\omega^c + \mathbb{E}_t\left[(1+b_2^c)\tilde{r}_t^c + K_{t+1}^c(\tilde{\boldsymbol{r}}_t)\right], \\ S_t^{(0)} &\doteq -\omega^s + \mathbb{E}_t\left[\tilde{r}_t^s + K_{t+1}^s(\tilde{\boldsymbol{r}}_t)\right], \\ S_t^{(1)} &\doteq -(1-\gamma_1^s)\omega^s + \mathbb{E}_t\left[(1+b_1^s)\tilde{r}_t^s + K_{t+1}^s(\tilde{\boldsymbol{r}}_t)\right], \\ S_t^{(2)} &\doteq -(1-\gamma_2^s)\omega^s + \mathbb{E}_t\left[(1+b_2^s)\tilde{r}_t^s + K_{t+1}^s(\tilde{\boldsymbol{r}}_t)\right], \\ F_t &\doteq \mathbb{E}_t\left[K_{t+1}^f(\tilde{\boldsymbol{r}}_t)\right], \end{aligned}$$

with $K_{T+1}^j(\boldsymbol{r}_T) = 0$ for $j \in \{c, s, f\}$. It is easy to establish that $C_t^{(0)} \leq C_t^{(1)} \leq C_t^{(2)}$ and $S_t^{(0)} \leq S_t^{(1)} \leq S_t^{(2)}$ hold by definition.

In (25), $K_t^j(\boldsymbol{r}_{t-1})$ denotes the expected marginal profit of farmland in the remaining planning horizon (from period t onward) where crop j was grown in period $t-1$. Consider, for example, $K_t^c(\boldsymbol{r}_{t-1})$. It is given by the maximum profit from three options available to the farmer: (i) growing corn in period t and optimally using the farmland in the remaining periods (which yields the expected marginal profit $\mathbb{E}_t\left[K_{t+1}^c(\tilde{\boldsymbol{r}}_t)\right]$), (ii) growing soybeans in period t and optimally using the farmland in the remaining periods (which yields the expected marginal profit $\mathbb{E}_t\left[K_{t+1}^s(\tilde{\boldsymbol{r}}_t)\right]$), and (iii) growing a fallow crop in period t (which brings zero expected profit in this period) and optimally using the farmland in the remaining periods (which yields the expected marginal profit $\mathbb{E}_t\left[K_{t+1}^f(\tilde{\boldsymbol{r}}_t)\right]$).

The farmer's optimization problem stated in (24) has a piecewise linear structure—that is, the objective function is piecewise linear in the decision variables α_t^j and β_t^j for $j \in \{c, s, f\}$. Therefore, the optimal allocation decisions can be presented in terms of two other variables: the total proportion of farmland allocated to corn in period t, that is, $\alpha_t = \sum_{j\in\{c,s,f\}} \alpha_t^j$, and the total proportion of farmland allocated to a fallow crop in period t, that is, $\beta_t = \sum_{j\in\{c,s,f\}} \beta_t^j$. As we will discuss shortly, there is a one-to-one correspondence between these two variables and the decision variables α_t^j and β_t^j for $j \in \{c, s, f\}$. Focusing on α_t and β_t to characterize the optimal allocation decisions also enables us to make direct comparisons with the optimal allocation decision presented in Proposition 8 of our paper.

Proposition 9 *In period $t \in [1, T]$, the optimal corn allocation α_t^* and the optimal fallow* *crop allocation* β_t^* *are given by*

$$(\alpha_t^*,\beta_t^*) = \begin{cases} (0,1) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = F_t, \\ (1,0) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = C_t^{(2)} \text{ and } K_t^c(\boldsymbol{r}_{t-1}) = C_t^{(0)} \\ (1-\beta_{t-1},0) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = S_t^{(2)} \text{ and } K_t^c(\boldsymbol{r}_{t-1}) = C_t^{(0)} \\ (1-\alpha_{t-1},0) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = C_t^{(2)} \text{ and } K_t^s(\boldsymbol{r}_{t-1}) = C_t^{(1)} \text{ and } K_t^c(\boldsymbol{r}_{t-1}) = S_t^{(1)} \\ (1-\alpha_{t-1}-\beta_{t-1},0) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = S_t^{(2)} \text{ and } K_t^s(\boldsymbol{r}_{t-1}) = C_t^{(1)} \text{ and } K_t^c(\boldsymbol{r}_{t-1}) = S_t^{(1)} \\ (\beta_{t-1},1-\alpha_{t-1}-\beta_{t-1}) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = C_t^{(2)} \text{ and } K_t^s(\boldsymbol{r}_{t-1}) = F_t \text{ and } K_t^c(\boldsymbol{r}_{t-1}) = S_t^{(1)} \\ (0,1-\alpha_{t-1}-\beta_{t-1}) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = S_t^{(2)} \text{ and } K_t^s(\boldsymbol{r}_{t-1}) = F_t \text{ and } K_t^c(\boldsymbol{r}_{t-1}) = S_t^{(1)} \\ (\beta_{t-1},0) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = C_t^{(2)} \text{ and } K_t^s(\boldsymbol{r}_{t-1}) = S_t^{(0)} \\ (0,0) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = S_t^{(2)} \text{ and } K_t^s(\boldsymbol{r}_{t-1}) = S_t^{(0)} \\ (1-\alpha_{t-1},\alpha_{t-1}) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = C_t^{(2)} \text{ and } K_t^s(\boldsymbol{r}_{t-1}) = C_t^{(1)} \text{ and } K_t^c(\boldsymbol{r}_{t-1}) = F_t \\ (1-\alpha_{t-1}-\beta_{t-1},\alpha_{t-1}) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = S_t^{(2)} \text{ and } K_t^s(\boldsymbol{r}_{t-1}) = C_t^{(1)} \text{ and } K_t^c(\boldsymbol{r}_{t-1}) = F_t \\ (\beta_{t-1},1-\beta_{t-1}) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = C_t^{(2)} \text{ and } K_t^s(\boldsymbol{r}_{t-1}) = F_t \text{ and } K_t^c(\boldsymbol{r}_{t-1}) = F_t \\ (0,1-\beta_{t-1}) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = S_t^{(2)} \text{ and } K_t^s(\boldsymbol{r}_{t-1}) = F_t \text{ and } K_t^c(\boldsymbol{r}_{t-1}) = F_t. \end{cases}$$

The intuition behind Proposition 9 is similar to the intuition behind Proposition 8 of our paper. In particular, the optimal allocation decisions are characterized based on which of the three options, growing (i) corn, (ii) soybeans, or (iii) a fallow crop in period t (and optimally using the farmland in the remaining periods), is the most profitable on the farmland where crop $j \in \{c,s,f\}$ was grown in the previous period, as captured by the recursive operators $K_t^j(\boldsymbol{r}_{t-1})$ given in (25). For example, consider the first case presented in Proposition 9. When $K_t^f(\boldsymbol{r}_{t-1}) = F_t$, because $C_t^{(0)} \le C_t^{(1)} \le C_t^{(2)}$ and $S_t^{(0)} \le S_t^{(1)} \le S_t^{(2)}$ by definition, we have $K_t^c(\boldsymbol{r}_{t-1}) = F_t$ and $K_t^s(\boldsymbol{r}_{t-1}) = F_t$. In other words, growing a fallow crop is the most profitable option regardless of which crop was grown in the previous period. Therefore, the whole farmland is optimally allocated to a fallow crop; that is, $\beta_t^* = 1$ (and, thus, $\alpha_t^* = 0$). Consider another example (seventh case presented in Proposition 9): what is the optimal allocation when $K_t^s(\boldsymbol{r}_{t-1}) = F_t$, $K_t^f(\boldsymbol{r}_{t-1}) = S_t^{(2)}$, and $K_t^c(\boldsymbol{r}_{t-1}) = S_t^{(1)}$? In this case, growing a fallow crop is the most profitable option on the farmland where soybeans were grown in the previous period while growing soybeans is the most profitable option on the farmland where a fallow crop or corn was grown in the previous period. Therefore, $\beta_t^* = 1-\alpha_{t-1}-\beta_{t-1}$ and $1-\alpha_t^*-\beta_t^* = \beta_{t-1}+\alpha_{t-1}$; in other words, no farmland is allocated to corn, that is, $\alpha_t^* = 0$. The other cases are characterized in a similar fashion.

The optimal levels for the original decision variables $\alpha_t^{j^*}$ and $\beta_t^{j^*}$ for $j \in \{c,s,f\}$ in the farmer's optimization problem stated in (24) can be obtained from α_t^* and β_t^* characterizations, respectively, using the previous period's allocation for each crop $(\alpha_{t-1}, \beta_{t-1}, 1-\alpha_{t-1}-\beta_{t-1})$. For example, when $\alpha_t^* = 1 - \alpha_{t-1}$, because $1 - \alpha_{t-1} = (\beta_{t-1}) + (1 - \alpha_{t-1} - \beta_{t-1})$, it follows that $\alpha_t^{f^*} = \beta_{t-1}$, $\alpha_t^{s*} = 1 - \alpha_{t-1} - \beta_{t-1}$, and $\alpha_t^{c*} = 0$. Consider another example, $\beta_t^* = 1 - \beta_{t-1}$. In this case, because $1 - \beta_{t-1} = (\alpha_{t-1}) + (1 - \alpha_{t-1} - \beta_{t-1})$, we obtain $\beta_t^{c*} = \alpha_{t-1}$, $\beta_t^{s*} = 1 - \alpha_{t-1} - \beta_{t-1}$, and $\beta_t^{f^*} = 0$.

Note that Proposition 8 in our paper is a special case of Proposition 9 where there is no fallow crop and the whole farmland is allocated to corn and soybeans. This special case can be obtained from Proposition 9 by setting $\beta_{t-1} = 0$, $K_t^f(\mathbf{r}_{t-1}) = \max\left\{C_t^{(2)}, S_t^{(2)}\right\}$ for $t \in [1, T]$, $C_t^{(2)} = C_t^{(1)}$, and $S_t^{(2)} = S_t^{(1)}$ (because $b_2^j = b_1^j$ and $\gamma_2^j = \gamma_1^j$ for $j \in \{c, s\}$). It follows that cases 2 and 3 in Proposition 9 correspond to the case in Proposition 6 where the whole farmland is allocated to corn; cases 4 and 5 correspond to the one where the farmer rotates in period t; and cases 8 and 9 correspond to the case in Proposition 8 where the whole farmland is allocated to soybeans. The remaining cases (1, 6, 7, 10, 11, 12, and 13) in Proposition 9 become infeasible with the parameter restrictions of this special case and, thus, do not appear in Proposition 8 of our paper.

We note that Proposition 9 identifies three strategies that emerge as a part of the optimal allocation policy. In particular, there is *monoculture* strategy where only one of the crops is grown on the entire farmland—this strategy corresponds to cases *i*, *ii*, and *ix* in Proposition 9. There is *rotate* strategy where each crop is only grown on rotated farmland—this corresponds to cases *iv*, *v*, *vi*, *vii*, *x*, *xi*, *xii*, and *xiii*. Finally, there is *mixed* strategy (the remaining cases in Proposition 9) where one of the cash crops is grown on rotated farmland where a fallow crop was grown in the previous period, and the other cash crop is grown both on rotated farmland where the other cash crop was grown in the previous period and on nonrotated farmland. In the absence of a fallow crop, mixed strategy is not possible and rotate strategy is represented by only a single case as there are only two crops, consistent with our optimal policy characterization in Proposition 8.

The characterization of the optimal total expected profit from period t onward follows the same structure as the characterization in Proposition 8 of our paper:

Proposition 10 *The optimal value function from period t onward is given by*

$$V_t(\alpha_{t-1}, \beta_{t-1}, \boldsymbol{r}_{t-1}) = \alpha_{t-1} K_t^c(\boldsymbol{r}_{t-1}) + \beta_{t-1} K_t^f(\boldsymbol{r}_{t-1}) + (1 - \alpha_{t-1} - \beta_{t-1}) K_t^s(\boldsymbol{r}_{t-1}),$$

where $K_t^j(\boldsymbol{r}_{t-1})$ for $j \in \{c, s, f\}$ is given by (25).

Similar to our paper, the optimal total expected profit from period t onward is given by the product of the proportion of farmland allocated to crop $j \in \{c, s, f\}$ in the previous period (α_{t-1} for corn, β_{t-1} for a fallow crop, and $1 - \alpha_{t-1} - \beta_{t-1}$ for soybeans) and its corresponding expected marginal profit $K_t^j(\boldsymbol{r}_{t-1})$ in (25).

We close this section with an important observation. Similar to our paper, once the farmer optimally follows a monoculture allocation policy in period t—that is, the whole farmland is only allocated to a single crop—the farmer also optimally follows a monoculture policy in the subsequent periods. The following corollary formalizes this observation.

Corollary 1 *i) When the whole farmland is allocated to corn in period $t-1 \in [0, T-1]$, that is, $\alpha_{t-1} = 1$ and $\beta_{t-1} = 0$, the optimal corn and fallow crop allocation (α_t^*, β_t^*) in period t are given by*

$$(\alpha_t^*, \beta_t^*) = \begin{cases} (1,0) & \text{if } K_t^c(\boldsymbol{r}_{t-1}) = C_t^{(0)}, \\ (0,0) & \text{if } K_t^c(\boldsymbol{r}_{t-1}) = S_t^{(1)}, \\ (0,1) & \text{if } K_t^c(\boldsymbol{r}_{t-1}) = F_t. \end{cases}$$

ii) When the whole farmland is allocated to a fallow crop in period $t-1$, that is, $\alpha_{t-1} = 0$ and $\beta_{t-1} = 1$, the optimal corn and fallow crop allocation (α_t^*, β_t^*) in period t are given by

$$(\alpha_t^*, \beta_t^*) = \begin{cases} (1,0) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = C_t^{(2)}, \\ (0,0) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = S_t^{(2)}, \\ (0,1) & \text{if } K_t^f(\boldsymbol{r}_{t-1}) = F_t. \end{cases}$$

iii) When the whole farmland is allocated to soybeans in period $t-1$, that is, $\alpha_{t-1} = 0$ and $\beta_{t-1} = 0$, the optimal corn and fallow crop allocation (α_t^*, β_t^*) in period t are given by

$$(\alpha_t^*, \beta_t^*) = \begin{cases} (1,0) & \text{if } K_t^s(\boldsymbol{r}_{t-1}) = C_t^{(1)}, \\ (0,0) & \text{if } K_t^s(\boldsymbol{r}_{t-1}) = S_t^{(0)}, \\ (0,1) & \text{if } K_t^s(\boldsymbol{r}_{t-1}) = F_t. \end{cases}$$

B.2.3 Numerical Analysis

In this section, we replicate our computational experiments in Sections 5.2.3 and 5.2.4 of our paper in the presence of a fallow crop. In particular, we numerically examine the value of crop planning based on multiple crops with rotation benefits in comparison with continuously growing only one of the crops, and compare the optimal allocation policy's performance with that of heuristic allocation policies.

To form our numerical instances, we consider a smaller subset of parameter values that are used in the paper[12] but at the same time introduce new parameters to capture the impact of introducing a fallow crop. In particular, we consider revenue correlation $\rho = \{0.63, 0.73, 0.83\}$, evenly-spaced around the baseline value 0.73; we consider corn (soybean) volatility $\sigma^c \in \{81.17, 108.22, 135.28\}$ ($\sigma^s \in \{59.77, 79.69, 99.61\}$); these parameter levels correspond to $\{-25\%, 0\%, 25\%\}$ of their baseline values. To capture the crop rotation benefits on the farmland where the other cash crop was grown in the previous season, we use the parameters given in the paper. In particular, we consider revenue-enhancing rotation benefit $b_1^c \in \{0.06, 0.08, 0.10\}$ for corn and $b_1^s \in \{0.13, 0.17, 0.21\}$ for soybeans, and cost-reducing rotation benefit $\gamma_1^c \in \{0.08, 0.10, 0.13\}$ for corn—these parameter values correspond to $\{-25\%, 0\%, 25\%\}$ of their baseline values. For crop rotation benefits on the farmland where a fallow crop was grown in the previous season, we set $b_2^c \in \{0.10, 0.12, 0.14\}$, $b_2^s \in \{0.21, 0.26, 0.30\}$, and $\gamma_2^c \in \{0.13, 0.15, 0.18\}$; these parameter values are $\{25\%, 50\%, 75\%\}$ higher than the baseline value for b_1^c, b_1^s, and γ_1^c, respectively. We continue to assume $\gamma_1^s = 0$, and we also set $\gamma_2^s = 0$. We set initial corn allocation $\alpha_0 \in \{0.48, 0.58, 0.68\}$, evenly-spaced around the baseline value 0.58, and initial fallow crop allocation $\beta_0 \in \{0, 0.1, 0.2\}$. We also use different planning horizons, specifically $T \in \{5, 10, 15\}$. In summary, we consider 531,441 numerical instances.

To study the value of crop planning based on multiple crops with rotation benefits in comparison with continuously growing only one of the crops, paralleling Section 5.2.3 of our paper, we numerically compute the percentage profit loss $\Delta * 100$ due to continuously growing the same crop. Recall that Δ is defined as

$$\Delta = \left[\frac{V_1(\alpha_0, \beta_0 \boldsymbol{r}_0) - V_1^B(\alpha_0, \beta_0, \boldsymbol{r}_0)}{V_1(\alpha_0, \beta_0, \boldsymbol{r}_0)} \right],$$

where $V_1(\alpha_0, \beta_0, \boldsymbol{r}_0)$ is the optimal total expected profit over the planning horizon and $V_1^B(\alpha_0, \beta_0, \boldsymbol{r}_0)$ denotes the expected profit under the benchmark case. In the benchmark case the farmer has two options, growing corn or growing soybeans over the entire planning horizon. Here, $V_1^B(\alpha_0, \beta_0, \boldsymbol{r}_0)$ denotes the maximum expected profit from these two options.

[12]Precisely speaking, instead of 5 values considered for ρ, σ^c, σ^s, b_1^c, b_1^s, γ_1^c, and α_0, and 4 values considered for T in the paper, we consider 3 values for these parameters in this section.

Table 2: Performance of heuristic allocation policies in the model with fallow crops

	Always rotate	Always rotate (monoculture)	Myopic	One-period lookahead
Average	0.99%	1.69%	0.65%	0.03%
Min	0.38%	0.65%	0.26%	0.01%
Max	1.89%	3.15%	1.11%	0.10%

Notes. For each heuristic (H), "Average" denotes the average percentage profit loss ($\Delta^H * 100$), whereas "Min" and "Max" denote the minimum and the maximum percentage profit loss observed in all numerical instances, respectively.

We find that the average profit loss in the numerical instances considered is 18.16% with a minimum and a maximum loss of 12.94% and 23.03%, respectively. This result indicates the same conclusion as in our paper: *crop planning based on multiple crops with rotation benefits has a substantial economic value.*

To study the performance of a variety of heuristic allocation policies in comparison with the optimal policy, paralleling Section 5.2.4 of our paper, we numerically compute percentage profit loss $\Delta^H * 100$ due to employing heuristic policy (H). Recall that Δ^H is defined as

$$\Delta^H = \left[\frac{V_1(\alpha_0, \beta_0, \boldsymbol{r}_0) - V_1^H(\alpha_0, \beta_0, \boldsymbol{r}_0)}{V_1(\alpha_0, \beta_0, \boldsymbol{r}_0)}\right],$$

where $V_1^H(\alpha_0, \beta_0, \boldsymbol{r}_0)$ denotes the expected profit over the planning horizon under the heuristic allocation policy. We consider the same set of heuristic allocation policies as in our paper. For the always rotate and the always rotate (monoculture) heuristic policies, we continue to assume that the farmland is allocated to only corn and soybeans—that is, there is no allocation to a fallow crop under these policies. For the always rotate heuristic, we consider two heuristics based on whether the farmer grows corn or soybeans on the previously fallow farmland (β_0) in the first period. In each numerical instance, we only report the better performing heuristic—that is, $V_1^H(\alpha_0, \beta_0, \boldsymbol{r}_0)$ denotes the higher expected profit of the two heuristics.

Table 2 summarizes the average, minimum, and maximum percentage profit losses $\Delta^H *$ 100 observed under each heuristic policy in all numerical instances. We make the same two observations as in our paper:

1) In all numerical instances the profit loss is the smallest with the one-period lookahead policy and the maximum percentage loss with this policy is only 0.10%, as reported in Table 2. In other words, *the one-period lookahead policy not only outperforms the commonly suggested heuristic policies in the literature but also provides a near-optimal performance.*

2) In all numerical instances the profit loss with the always rotate (monoculture) policy is larger than the always rotate policy. Therefore, *when the farmer chooses to follow a rotation-based allocation policy, there is value in considering the possibility of growing more than one crop in the same season.*

C Relaxing the One-season Crop Rotation History Assumption

Throughout the paper, we assume that crop rotation benefits carry through for one growing season and, thus, the farmland allocation decision in each season is affected by only the allocation in the previous season and not the earlier seasons. Although a one-season crop rotation history assumption is reasonable for corn-soybean rotation as considered in our paper, crop rotation benefits may carry through for more than one growing season for some other crops. In this section, we generalize our model to incorporate crop rotation benefits that carry through for two growing seasons. In the presence of a two-season crop history, the farmland allocation decision in each growing season is affected by the allocations in the previous two seasons.

The remainder of this section is organized as follows. We describe the model in Section C.1. We characterize the optimal allocation policy in Section C.2 and demonstrate that although the optimal policy is more complex (due to a larger set of decision variables) than the optimal allocation policy presented in our paper, it follows the same structure.

C.1 Model Description

Paralleling our paper, we consider a farmer who allocates a single acre of farmland among two cash crops (corn and soybeans) in each growing season to maximize the expected total profit over a finite number of growing seasons.

Decision variables. Because we consider a two-season crop history, we define a new set of decision variables to capture this feature. In particular, let $\alpha_t^{i,j} \in [0,1]$ for $i, j \in \{c, s\}$

denote the proportion of farmland allocated to corn in time period (growing season) t on which crop i was grown in period $t-2$ and crop j was grown in period $t-1$. Similarly, let $\beta_t^{i,j} \in [0,1]$ for $i, j \in \{c, s\}$ denote the proportion of farmland allocated to soybeans in period t on which crop i was grown in period $t-2$ and crop j was grown in period $t-1$. By definition, $\sum_i \sum_j (\alpha_t^{i,j} + \beta_t^{i,j}) = 1$. For notational convenience, we define $\boldsymbol{\alpha}_t \doteq (\alpha_t^{c,c}, \alpha_t^{s,c}, \alpha_t^{c,s}, \alpha_t^{s,s})$ and $\boldsymbol{\beta}_t \doteq (\beta_t^{c,c}, \beta_t^{s,c}, \beta_t^{c,s}, \beta_t^{s,s})$ to denote the corn and soybean allocations in period t, respectively.

Crop rotation benefits. Because there is two-season crop history, the rotation benefits are defined based on what was grown on the farmland in the last two periods. To capture the revenue-enhancing crop rotation benefits, we assume that the uncertain revenue per acre of crop $j \in \{c, s\}$ in period t is

i. $\tilde{r}_t^j$ if it is grown on farmland where the same crop j was grown in the previous two periods,

ii. $(1+b_1^j)\tilde{r}_t^j$ if it is grown on farmland where the same crop j was grown in the previous period and the other crop $(-j)$ was grown in period $t-2$, and

iii. $(1+b_2^j)\tilde{r}_t^j$ if it is grown on farmland where the other crop $(-j)$ was grown in the previous period and the same crop j was grown in period $t-2$, and

iv. $(1+b_3^j)\tilde{r}_t^j$ if it is grown on farmland where the other crop $(-j)$ was grown in the previous two periods.

Here, case i represents the nonrotated farmland whereas the other three cases represent the rotated farmland. Consistent with practice, we assume that revenue-enhancing rotation benefit based on the period $t-1$ allocation is larger than the benefit based on the period $t-2$ allocation, that is, $b_2^j > b_1^j > 0$, and the revenue-enhancing crop rotation benefit is the largest on the farmland where the other crop was grown in the previous two seasons, that is, $b_3^j > b_2^j$. To capture the cost-reducing crop rotation benefits, we assume that the unit farming cost of crop j is

i. ω^j if it is grown on farmland where the same crop j was grown in the previous two periods,

ii. $(1-\gamma_1^j)\omega^j$ if it is grown on farmland where the same crop j was grown in the previous period and the other crop $(-j)$ was grown in period $t-2$, and

iii. $(1-\gamma_2^j)\omega^j$ if it is grown on farmland where the other crop $(-j)$ was grown in the previous period and the same crop j was grown in period $t-2$, and

iv. $(1-\gamma_3^j)\omega^j$ if it is grown on rotated farmland where a fallow crop was grown in the previous period.

Similarly, we assume $\gamma_3^j > \gamma_2^j > \gamma_1^j > 0$. The special case where the crop rotation history is only one year, as considered in our paper, can be captured by setting $b_1^j = 0$, $b_2^j = b_3^j = b^j$, $\gamma_1^j = 0$, and $\gamma_2^j = \gamma_3^j = \gamma^j$ for $j \in \{c, s\}$.

Formulation. Paralleling our paper, we formulate the farmer's problem as a finite horizon stochastic dynamic program. In each period $t \in [1, T]$, the sequence of events is as follows:

(i) At the beginning of period t, the farmer observes the corn allocations $\boldsymbol{\alpha}_{t-1} = \left(\alpha_{t-1}^{c,c}, \alpha_{t-1}^{s,c}, \alpha_{t-1}^{c,s}, \alpha_{t-1}^{s,s}\right)$, the soybean allocations $\boldsymbol{\beta}_{t-1} = \left(\beta_{t-1}^{c,c}, \beta_{t-1}^{s,c}, \beta_{t-1}^{c,s}, \beta_{t-1}^{s,s}\right)$, and corn and soybean revenues $\boldsymbol{r}_{t-1} = (r_{t-1}^c, r_{t-1}^s)$ from period $t-1$. The farmer then chooses the corn allocations $\alpha_t^{i,j}$ and the soybean allocations $\beta_t^{i,j}$, where $\sum_i \sum_j (\alpha_t^{i,j} + \beta_t^{i,j}) = 1$, constrained by the available farmland where crop i was grown in period $t-2$ and crop j was grown in period $t-1$, that is, (i) $\beta_t^{s,s} + \alpha_t^{s,s} = \beta_{t-1}^{c,s} + \beta_{t-1}^{s,s}$, (ii) $\beta_t^{s,c} + \alpha_t^{s,c} = \alpha_{t-1}^{s,s} + \alpha_{t-1}^{c,s}$, (iii) $\beta_t^{c,s} + \alpha_t^{c,s} = \beta_{t-1}^{s,c} + \beta_{t-1}^{c,c}$, and (iv) $\beta_t^{c,c} + \alpha_t^{c,c} = \alpha_{t-1}^{s,c} + \alpha_{t-1}^{c,c}$. For example, the first constraint ensures that the sum of the proportion of farmland allocated to corn and soybeans in this period where soybeans were grown in the previous two periods, that is, $\alpha_t^{s,s} + \beta_t^{s,s}$, equals the actual proportion of farmland where soybeans were grown in the previous two periods, which is given by $\beta_{t-1}^{c,s} + \beta_{t-1}^{s,s}$.

(ii) At the end of period t, the corn and soybean revenues $\tilde{\boldsymbol{r}}_t = (\tilde{r}_t^c, \tilde{r}_t^s)$ are realized and the farmer collects the revenues from the crop sales.

The farmer's immediate payoff in period $t \in [1, T]$ is given by

$$\begin{aligned} L(\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t \mid \boldsymbol{\alpha}_{t-1}, \boldsymbol{\beta}_{t-1}, \mathbf{r}_{t-1}) &= (\alpha_t^{c,c} + (1+b_1^c)\alpha_t^{s,c} + (1+b_2^c)\alpha_t^{c,s} + (1+b_3^c)\alpha_t^{s,s})\,\mathbb{E}_t[\tilde{r}_t^c] \quad (26) \\ &+ (\beta_t^{s,s} + (1+b_1^s)\beta_t^{c,s} + (1+b_2^s)\beta_t^{s,c} + (1+b_3^s)\beta_t^{c,c})\,\mathbb{E}_t[\tilde{r}_t^s] \\ &- (\alpha_t^{c,c} + (1-\gamma_1^c)\alpha_t^{s,c} + (1-\gamma_2^c)\alpha_t^{c,s} + (1-\gamma_3^c)\alpha_t^{s,s})\,\omega^c \\ &- (\beta_t^{s,s} + (1-\gamma_1^s)\beta_t^{c,s} + (1-\gamma_2^s)\beta_t^{s,c} + (1-\gamma_3^s)\beta_t^{c,c})\,\omega^s, \end{aligned}$$

where $\mathbb{E}_t[.]$ denotes the expectation operator conditional on the available information at time t, that is, $\mathbb{E}_t[.] = \mathbb{E}[.|\boldsymbol{r}_{t-1}]$. In (26), the first line corresponds to the total expected revenue from growing corn in period t. It is the sum of expected revenue from growing corn on four different farmlands: farmland where corn was grown in the last two periods (which does not experience any revenue-enhancing rotation benefit), farmland where soybeans were grown in period $t-2$ and corn was grown in the previous period, farmland where corn was grown in period $t-2$ and soybeans were grown in the previous period, and farmland where soybeans were grown in the last two periods (which experiences the largest revenue-enhancing rotation benefit). The second line in (26) denotes the total expected revenue from growing soybeans in period t in a similar fashion. The last two lines in (26) denote the total farming cost incurred from growing corn and soybeans in period t, respectively. These expressions capture the cost-reducing rotation benefit experienced for each crop based on the four different farmlands.

Let $V_t(\boldsymbol{\alpha}_{t-1}, \boldsymbol{\beta}_{t-1}, \boldsymbol{r}_{t-1})$ for $t \in [1, T]$ denote the optimal value function from period t onward given $\boldsymbol{\alpha}_{t-1}$, $\boldsymbol{\beta}_{t-1}$, and $\boldsymbol{r}_{t-1}$, which satisfies

$$
\begin{aligned}
V_t(\boldsymbol{\alpha}_{t-1}, \boldsymbol{\beta}_{t-1}, \boldsymbol{r}_{t-1}) = \max_{\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t} \quad & \left\{ L(\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t \mid \boldsymbol{\alpha}_{t-1}, \boldsymbol{\beta}_{t-1}, \boldsymbol{r}_{t-1}) + \mathbb{E}_t \left[V_{t+1}(\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, \tilde{\boldsymbol{r}}_t) \right] \right\} \\
\text{s.t.} \quad & \beta_t^{s,s} + \alpha_t^{s,s} = \beta_{t-1}^{c,s} + \beta_{t-1}^{s,s}, \; \beta_t^{s,c} + \alpha_t^{s,c} = \alpha_{t-1}^{s,s} + \alpha_{t-1}^{c,s}, \\
& \beta_t^{c,s} + \alpha_t^{c,s} = \beta_{t-1}^{s,c} + \beta_{t-1}^{c,c}, \; \beta_t^{c,c} + \alpha_t^{c,c} = \alpha_{t-1}^{s,c} + \alpha_{t-1}^{c,c}, \\
& 0 \le \alpha_t^{i,j} \le 1, \; 0 \le \beta_t^{i,j} \le 1, \; \sum_{i \in \{c,s\}} \sum_{j \in \{c,s\}} (\alpha_t^{i,j} + \beta_t^{i,j}) = 1,
\end{aligned}
\tag{27}
$$

with a boundary condition $V_{T+1}(\cdot) = 0$. The farmer's optimal total expected profit over the entire planning horizon is given by $V_1(\boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, \boldsymbol{r}_0)$, where $\boldsymbol{\alpha}_0$, $\boldsymbol{\beta}_0$, and $\boldsymbol{r}_0$ denote the observed corn allocations, soybean allocations, and crop revenues at the beginning of the planning horizon, respectively.

C.2 Optimal Allocation Policy

We now solve for the farmer's optimization problem stated in (27) and characterize the optimal allocation decisions and the optimal value function in period $t \in [1, T]$. For this purpose, similar to the analysis in our paper, we first define the following recursive operators:

$$\begin{aligned}
K_t^{c,c}(\mathbf{r}_{t-1}) &\doteq \max\left\{-\omega^c + \mathbb{E}_t[\tilde{r}_t^c + K_{t+1}^{c,c}(\tilde{\mathbf{r}}_t)], -(1-\gamma_3^s)\omega^s + \mathbb{E}_t[(1+b_3^s)\tilde{r}_t^s + K_{t+1}^{c,s}(\tilde{\mathbf{r}}_t)]\right\}, \\
K_t^{s,c}(\mathbf{r}_{t-1}) &\doteq \max\left\{-(1-\gamma_1^c)\omega^c + \mathbb{E}_t[(1+b_1^c)\tilde{r}_t^c + K_{t+1}^{c,c}(\tilde{\mathbf{r}}_t)], -(1-\gamma_2^s)\omega^s + \mathbb{E}_t[(1+b_2^s)\tilde{r}_t^s + K_{t+1}^{c,s}(\tilde{\mathbf{r}}_t)]\right\}, \\
K_t^{c,s}(\mathbf{r}_{t-1}) &\doteq \max\left\{-(1-\gamma_2^c)\omega^c + \mathbb{E}_t[(1+b_2^c)\tilde{r}_t^c + K_{t+1}^{s,c}(\tilde{\mathbf{r}}_t)], -(1-\gamma_1^s)\omega^s + \mathbb{E}_t[(1+b_1^s)\tilde{r}_t^s + K_{t+1}^{s,s}(\tilde{\mathbf{r}}_t)]\right\}, \\
K_t^{s,s}(\mathbf{r}_{t-1}) &\doteq \max\left\{-(1-\gamma_3^c)\omega^c + \mathbb{E}_t[(1+b_3^c)\tilde{r}_t^c + K_{t+1}^{s,c}(\tilde{\mathbf{r}}_t)], -\omega^s + \mathbb{E}_t[\tilde{r}_t^s + K_{t+1}^{s,s}(\tilde{\mathbf{r}}_t)]\right\}.
\end{aligned}$$

with $K_{T+1}^{i,j}(\boldsymbol{r}_T) = 0$ for $i, j \in \{c, s\}$. Here, $K_t^{i,j}(\boldsymbol{r}_{t-1})$ denotes the expected marginal profit of farmland in the remaining planning horizon (from period t onward) where crop i was grown in period $t-2$ and crop j was grown in period $t-1$. Consider, for example, $K_t^{c,c}(\boldsymbol{r}_{t-1})$. It is given by the maximum profit from two options: (i) growing corn in period t and optimally using the farmland in the remaining periods (which yields the expected marginal profit $\mathbb{E}_t\left[K_{t+1}^{c,c}(\tilde{\boldsymbol{r}}_t)\right]$) and (ii) growing soybeans in period t and optimally using the farmland in the remaining periods (which yields the expected marginal profit $\mathbb{E}_t\left[K_{t+1}^{c,s}(\tilde{\boldsymbol{r}}_t)\right]$). For the first option, because corn was grown in the previous two periods, growing corn in this period does not accrue any rotation benefits. For the second option, because corn was grown in the previous two periods, growing soybeans in this period accrues the largest rotation benefits, captured by b_3^s and γ_3^s.

For notational convenience, we define

$$\begin{aligned}
C_t^{(0)} &\doteq -\omega^c + \mathbb{E}_t\left[\tilde{r}_t^c + K_{t+1}^{c,c}(\tilde{\boldsymbol{r}}_t)\right], \\
C_t^{(1)} &\doteq -(1-\gamma_1^c)\omega^c + \mathbb{E}_t\left[(1+b_1^c)\tilde{r}_t^c + K_{t+1}^{c,c}(\tilde{\boldsymbol{r}}_t)\right], \\
C_t^{(2)} &\doteq -(1-\gamma_2^c)\omega^c + \mathbb{E}_t\left[(1+b_2^c)\tilde{r}_t^c + K_{t+1}^{s,c}(\tilde{\boldsymbol{r}}_t)\right], \\
C_t^{(3)} &\doteq -(1-\gamma_3^c)\omega^c + \mathbb{E}_t\left[(1+b_3^c)\tilde{r}_t^c + K_{t+1}^{s,c}(\tilde{\boldsymbol{r}}_t)\right], \\
S_t^{(0)} &\doteq -\omega^s + \mathbb{E}_t\left[\tilde{r}_t^s + K_{t+1}^{s,s}(\tilde{\boldsymbol{r}}_t)\right], \\
S_t^{(1)} &\doteq -(1-\gamma_1^s)\omega^s + \mathbb{E}_t\left[(1+b_1^s)\tilde{r}_t^s + K_{t+1}^{s,s}(\tilde{\boldsymbol{r}}_t)\right], \\
S_t^{(2)} &\doteq -(1-\gamma_2^s)\omega^s + \mathbb{E}_t\left[(1+b_2^s)\tilde{r}_t^s + K_{t+1}^{c,s}(\tilde{\boldsymbol{r}}_t)\right], \\
S_t^{(3)} &\doteq -(1-\gamma_3^s)\omega^s + \mathbb{E}_t\left[(1+b_3^s)\tilde{r}_t^s + K_{t+1}^{c,s}(\tilde{\boldsymbol{r}}_t)\right],
\end{aligned}$$

where $C_t^{(0)} < C_t^{(1)}$, $C_t^{(2)} < C_t^{(3)}$, $S_t^{(0)} < S_t^{(1)}$, and $S_t^{(2)} < S_t^{(3)}$ by construction. Using this notation, the recursive operators can be rewritten as:

$$\begin{aligned} K_t^{c,c}(\boldsymbol{r}_{t-1}) &= \max\left\{C_t^{(0)}, S_t^{(3)}\right\}, \\ K_t^{s,c}(\boldsymbol{r}_{t-1}) &= \max\left\{C_t^{(1)}, S_t^{(2)}\right\}, \\ K_t^{c,s}(\boldsymbol{r}_{t-1}) &= \max\left\{C_t^{(2)}, S_t^{(1)}\right\}, \\ K_t^{s,s}(\boldsymbol{r}_{t-1}) &= \max\left\{C_t^{(3)}, S_t^{(0)}\right\}. \end{aligned} \tag{28}$$

Let $\widehat{(.)}$ denote the optimal decision.

Proposition 11 *In period $t \in [1, T]$, the optimal allocation is characterized by the following cases where, in each case, the unlisted decision variables equal 0:*

(i) If $K_t^{c,c}(\mathbf{r}_{t-1}) = C_t^{(0)}$ and $K_t^{c,s}(\mathbf{r}_{t-1}) = C_t^{(2)}$ then $\widehat{\alpha}_t^{c,c} = \alpha_{t-1}^{c,c} + \alpha_{t-1}^{s,c}$, $\widehat{\alpha}_t^{c,s} = \beta_{t-1}^{c,c} + \beta_{t-1}^{s,c}$, $\widehat{\alpha}_t^{s,c} = \alpha_{t-1}^{c,s} + \alpha_{t-1}^{s,s}$, and $\widehat{\alpha}_t^{s,s} = \beta_{t-1}^{s,s} + \beta_{t-1}^{c,s}$.

(ii) If $K_t^{s,c}(\mathbf{r}_{t-1}) = S_t^{(2)}$ and $K_t^{s,s}(\mathbf{r}_{t-1}) = S_t^{(0)}$ then $\widehat{\beta}_t^{s,s} = \beta_{t-1}^{s,s} + \beta_{t-1}^{c,s}$, $\widehat{\beta}_t^{s,c} = \alpha_{t-1}^{c,s} + \alpha_{t-1}^{s,s}$, $\widehat{\beta}_t^{c,s} = \beta_{t-1}^{c,c} + \beta_{t-1}^{s,c}$, and $\widehat{\beta}_t^{c,c} = \alpha_{t-1}^{s,c} + \alpha_{t-1}^{c,c}$.

(iii) If $K_t^{c,s}(\mathbf{r}_{t-1}) = C_t^{(2)}$ and $K_t^{s,c}(\mathbf{r}_{t-1}) = S_t^{(2)}$ then $\widehat{\alpha}_t^{c,s} = \beta_{t-1}^{s,c} + \beta_{t-1}^{c,c}$, $\widehat{\beta}_t^{c,c} = \alpha_{t-1}^{s,c} + \alpha_{t-1}^{c,c}$, $\widehat{\beta}_t^{s,c} = \alpha_{t-1}^{c,s} + \alpha_{t-1}^{s,s}$, and $\widehat{\alpha}_t^{s,s} = \beta_{t-1}^{c,s} + \beta_{t-1}^{s,s}$.

(iv) If $K_t^{c,c}(\mathbf{r}_{t-1}) = C_t^{(0)}$ and $K_t^{s,s}(\mathbf{r}_{t-1}) = S_t^{(0)}$ then $\widehat{\alpha}_t^{c,c} = \alpha_{t-1}^{s,c} + \alpha_{t-1}^{c,c}$, $\widehat{\alpha}_t^{s,c} = \alpha_{t-1}^{c,s} + \alpha_{t-1}^{s,s}$, $\widehat{\beta}_t^{c,s} = \beta_{t-1}^{s,c} + \beta_{t-1}^{c,c}$, and $\widehat{\beta}_t^{s,s} = \beta_{t-1}^{c,s} + \beta_{t-1}^{s,s}$.

(v) If $K_t^{c,c}(\mathbf{r}_{t-1}) = C_t^{(0)}$, $K_t^{c,s}(\mathbf{r}_{t-1}) = S_t^{(1)}$, and $K_t^{s,s}(\mathbf{r}_{t-1}) = C_t^{(3)}$ then $\widehat{\alpha}_t^{c,c} = \alpha_{t-1}^{s,c} + \alpha_{t-1}^{c,c}$, $\widehat{\alpha}_t^{s,c} = \alpha_{t-1}^{c,s} + \alpha_{t-1}^{s,s}$, $\widehat{\beta}_t^{c,s} = \beta_{t-1}^{s,c} + \beta_{t-1}^{c,c}$, and $\widehat{\alpha}_t^{s,s} = \beta_{t-1}^{c,s} + \beta_{t-1}^{s,s}$.

(vi) If $K_t^{c,c}(\mathbf{r}_{t-1}) = S_t^{(3)}$, $K_t^{s,c}(\mathbf{r}_{t-1}) = C_t^{(1)}$, and $K_t^{s,s}(\mathbf{r}_{t-1}) = S_t^{(0)}$ then $\widehat{\beta}_t^{c,c} = \alpha_{t-1}^{s,c} + \alpha_{t-1}^{c,c}$, $\widehat{\beta}_t^{c,s} = \beta_{t-1}^{s,c} + \beta_{t-1}^{c,c}$, $\widehat{\alpha}_t^{s,c} = \alpha_{t-1}^{c,s} + \alpha_{t-1}^{s,s}$, and $\widehat{\beta}_t^{s,s} = \beta_{t-1}^{c,s} + \beta_{t-1}^{s,s}$.

(vii) If $K_t^{c,c}(\mathbf{r}_{t-1}) = S_t^{(3)}$, $K_t^{c,s}(\mathbf{r}_{t-1}) = S_t^{(1)}$, $K_t^{s,c}(\mathbf{r}_{t-1}) = C_t^{(1)}$, and $K_t^{s,s}(\mathbf{r}_{t-1}) = C_t^{(3)}$ then $\widehat{\beta}_t^{c,c} = \alpha_{t-1}^{s,c} + \alpha_{t-1}^{c,c}$, $\widehat{\beta}_t^{c,s} = \beta_{t-1}^{s,c} + \beta_{t-1}^{c,c}$, $\widehat{\alpha}_t^{s,c} = \alpha_{t-1}^{c,s} + \alpha_{t-1}^{s,s}$, and $\widehat{\alpha}_t^{s,s} = \beta_{t-1}^{c,s} + \beta_{t-1}^{s,s}$.

(viii) If $K_t^{s,c}(\mathbf{r}_{t-1}) = S_t^{(2)}$, $K_t^{c,s}(\mathbf{r}_{t-1}) = S_t^{(1)}$, and $K_t^{s,s}(\mathbf{r}_{t-1}) = C_t^{(3)}$ then $\widehat{\beta}_t^{c,c} = \alpha_{t-1}^{s,c} + \alpha_{t-1}^{c,c}$, $\widehat{\beta}_t^{c,s} = \beta_{t-1}^{s,c} + \beta_{t-1}^{c,c}$, $\widehat{\beta}_t^{s,c} = \alpha_{t-1}^{c,s} + \alpha_{t-1}^{s,s}$, and $\widehat{\alpha}_t^{s,s} = \beta_{t-1}^{c,s} + \beta_{t-1}^{s,s}$.

(ix) If $K_t^{c,c}(\mathbf{r}_{t-1}) = S_t^{(3)}$, $K_t^{s,c}(\mathbf{r}_{t-1}) = C_t^{(1)}$, and $K_t^{c,s}(\mathbf{r}_{t-1}) = C_t^{(2)}$ then $\widehat{\beta}_t^{c,c} = \alpha_{t-1}^{s,c} + \alpha_{t-1}^{c,c}$, $\widehat{\alpha}_t^{c,s} = \beta_{t-1}^{s,c} + \beta_{t-1}^{c,c}$, $\widehat{\alpha}_t^{s,c} = \alpha_{t-1}^{c,s} + \alpha_{t-1}^{s,s}$, and $\widehat{\alpha}_t^{s,s} = \beta_{t-1}^{c,s} + \beta_{t-1}^{s,s}$.

The intuition behind Proposition 11 is similar to the intuition behind Proposition 8 of our paper. In particular, the optimal allocation decisions are characterized based on which of the two options, growing corn or soybeans in period t (and optimally using the farmland in the remaining periods), is the most profitable option on the farmland where crop j was grown in the previous period and crop i was grown in period $t-2$, as captured by the recursive operators $K_t^{i,j}(\boldsymbol{r}_{t-1})$ for $i, j \in \{c, s\}$ given in (28). For example, consider the first case presented in Proposition 11. When $K_t^{c,c}(\mathbf{r}_{t-1}) = C_t^{(0)}$ and $K_t^{c,s}(\mathbf{r}_{t-1}) = C_t^{(2)}$, because $C_t^{(0)} < C_t^{(1)}$, $C_t^{(2)} < C_t^{(3)}$, $S_t^{(0)} < S_t^{(1)}$, and $S_t^{(2)} < S_t^{(3)}$ by definition, we have $K_t^{s,c}(\boldsymbol{r}_{t-1}) = C_t^{(1)}$ and $K_t^{s,s}(\boldsymbol{r}_{t-1}) = C_t^{(3)}$. In other words, growing corn is the most profitable option regardless of which crop was grown in the previous two periods. Therefore, the whole farmland is optimally allocated to corn, that is, $\sum_i \sum_j \widehat{\alpha}_t^{i,j} = 1$ where $\widehat{\alpha}_t^{c,c} = \alpha_{t-1}^{c,c} + \alpha_{t-1}^{s,c}$, $\widehat{\alpha}_t^{c,s} = \beta_{t-1}^{c,c} + \beta_{t-1}^{s,c}$, $\widehat{\alpha}_t^{s,c} = \alpha_{t-1}^{c,s} + \alpha_{t-1}^{s,s}$, and $\widehat{\alpha}_t^{s,s} = \beta_{t-1}^{s,s} + \beta_{t-1}^{c,s}$ (and, thus, $\widehat{\beta}_t^{i,j} = 0$ for $i, j \in \{c, s\}$). Consider another example: what is the optimal allocation when $K_t^{c,c}(\mathbf{r}_{t-1}) = S_t^{(3)}$, $K_t^{c,s}(\mathbf{r}_{t-1}) = S_t^{(1)}$, $K_t^{s,c}(\mathbf{r}_{t-1}) = C_t^{(1)}$, and $K_t^{s,s}(\mathbf{r}_{t-1}) = C_t^{(3)}$ (case *vii* in Proposition 11)? In this case, growing corn is the most profitable option on the farmland where soybeans were grown in period $t-2$ regardless of the crop grown in period $t-1$, while growing soybeans is the most profitable option on the farmland where corn was grown in period $t-2$ regardless of the crop grown in period $t-1$. Therefore, the total optimal corn allocation in period t is given by $\sum_i(\alpha_{t-1}^{i,s} + \beta_{t-1}^{i,s})$ whereas the total optimal soybean allocation in period t is given by $\sum_i(\alpha_{t-1}^{i,c} + \beta_{t-1}^{i,c})$. The other cases are characterized in a similar fashion.

The characterization of the optimal total expected profit from period t onward follows the same structure as the characterization in Proposition 8 of our paper:

Proposition 12 *The optimal value function from period t onward is given by*

$$V_t(\boldsymbol{\alpha}_{t-1}, \boldsymbol{\beta}_{t-1}, \boldsymbol{r}_{t-1}) = \sum_{j \in \{c,s\}} \Big((\alpha_{t-1}^{c,j} + \alpha_{t-1}^{s,j}) K_t^{j,c}(\mathbf{r}_{t-1}) \Big) + \sum_{j \in \{c,s\}} \Big((\beta_{t-1}^{c,j} + \beta_{t-1}^{s,j}) K_t^{j,s}(\mathbf{r}_{t-1}) \Big),$$

where $K_t^{i,j}(\boldsymbol{r}_{t-1})$ for $i, j \in \{c, s\}$ is given by (28).

Similar to our paper, the optimal total expected profit from period t onward is given by the product of the proportion of farmland allocated to each crop in the previous period where crop j was grown in period $t-2$ ($\alpha_{t-1}^{c,j} + \alpha_{t-1}^{s,j}$ for corn and $\beta_{t-1}^{c,j} + \beta_{t-1}^{s,j}$ for soybeans) and its corresponding expected marginal profit as given in (28) ($K_t^{j,c}(\boldsymbol{r}_{t-1})$ for corn and $K_t^{j,s}(\boldsymbol{r}_{t-1})$ for soybeans). When the crop history is only one year, that is, $b_1^j = 0$, $b_2^j = b_3^j = b^j$, $\gamma_1^j = 0$, and $\gamma_2^j = \gamma_3^j = \gamma^j$ for $j \in \{c, s\}$, it follows that $K_t^{c,c}(\boldsymbol{r}_{t-1}) = K_t^{s,c}(\boldsymbol{r}_{t-1})$ and $K_t^{c,s}(\boldsymbol{r}_{t-1}) = K_t^{s,s}(\boldsymbol{r}_{t-1})$. In this case, denoting corn allocation in period $t-1$ with $\alpha_{t-1} = \sum_{j\in\{c,s\}} \left((\alpha_{t-1}^{c,j} + \alpha_{t-1}^{s,j})\right)$ and soybean allocation in period $t-1$ with $\beta_{t-1} = \sum_{j\in\{c,s\}} \left((\beta_{t-1}^{c,j} + \beta_{t-1}^{s,j})\right)$, and using $\beta_{t-1} = 1 - \alpha_{t-1}$, it is easy to establish that the optimal value function above is identical to the optimal value function presented in Proposition 8 of our paper.

We close this section with an important remark. In the presence of a one-year crop history, an important property of the optimal allocation policy, as presented in Proposition 8 of our paper, is that once the farmer optimally follows a monoculture allocation policy in period t—that is, the whole farmland is allocated to a single crop—the farmer optimally continues to follow the monoculture policy in the subsequent periods. In the presence of a two-year crop history, a slightly modified version of this property holds for the optimal allocation policy. In particular, once the farmer optimally follows a monoculture allocation policy *in the last two periods*, the farmer optimally continues to follow the monoculture policy in the subsequent periods. The following corollary formalizes this observation. To this end, we first define $\widehat{\alpha}_t \doteq \sum_i \sum_j \widehat{\alpha}_t^{i,j}$ and $\widehat{\beta}_t \doteq \sum_i \sum_j \widehat{\beta}_t^{i,j}$ as the optimal total soybean and corn allocation in period t, respectively.

Corollary 2 *Assume that the whole farmland is allocated to corn in period $t-1 \in [0, T-1]$, that is, $\beta_{t-1}^{i,j} = 0$ for $i, j \in \{c, s\}$.*
i) If the whole farmland is allocated to corn in period $t-2$, that is, $\alpha_{t-1}^{c,s} = \alpha_{t-1}^{s,s} = 0$ and $\alpha_{t-1}^{c,c} + \alpha_{t-1}^{s,c} = 1$, then the optimal total corn and soybean allocation $(\widehat{\alpha}_t, \widehat{\beta}_t)$ in period t are given by

$$(\widehat{\alpha}_t, \widehat{\beta}_t) = \begin{cases} (1,0) & \text{if } K_t^{c,c}(\boldsymbol{r}_{t-1}) = C_t^{(0)}, \\ (0,1) & \text{if } K_t^{c,c}(\boldsymbol{r}_{t-1}) = S_t^{(3)}. \end{cases}$$

ii) If the whole farmland is allocated to soybeans in period $t-2$, that is, $\alpha_{t-1}^{c,c} = \alpha_{t-1}^{s,c} = 0$ and $\alpha_{t-1}^{c,s} + \alpha_{t-1}^{s,s} = 1$, then the optimal total corn and soybean allocation $(\widehat{\alpha}_t, \widehat{\beta}_t)$ in period t are given by

$$(\widehat{\alpha}_t, \widehat{\beta}_t) = \begin{cases} (1,0) & \text{if } K_t^{s,c}(\boldsymbol{r}_{t-1}) = C_t^{(1)}, \\ (0,1) & \text{if } K_t^{s,c}(\boldsymbol{r}_{t-1}) = S_t^{(2)}. \end{cases}$$

Corollary 2 states that if the whole farmland is allocated to corn in perid $t-1$ and the whole farmland is allocated to either corn or soybeans in period $t-2$, the optimal allocation in period t is such that the whole farmland is also allocated to a single crop. A similar result can be obtained for the case where the whole farmland is allocated to soybeans in period $t-1$ and the whole farmland is allocated to either corn or soybeans in period $t-2$.

D Calculating the Expected Number of Periods and the Probability to Achieve Monoculture

As mentioned in Section 5.1, an important characteristic of the optimal allocation policy, as presented in Proposition 8, is that once the farmer optimally follows a monoculture allocation policy in one period, the farmer optimally continues to follow the monoculture policy in the subsequent periods. In this section, we investigate whether monoculture is always achievable in realistic planning horizons. To this end, we compute the expected number of periods for monoculture to set in when following the optimal allocation policy as well as the one-period lookahead heuristic policy. Note that this metric is computed as a conditional expectation: conditional on the event that monoculture occurs, the expected number of periods it takes for each policy to achieve monoculture. To complement this metric, we also compute another metric: the probability of the monoculture occurrence under each policy. Using our model calibration, we compute the values for these two metrics under both policies. We use the same 312,500 numerical instances as in Section 5.2.4.

Computation. To compute these two metrics, it is helpful to view our constructed lattice (the discretized corn and soybean revenue process) equivalently as follows: the discretized corn and soybean revenue evolution over time can follow a finite number of paths, where each path specifies the values for corn and soybean revenues in each period $t \in \{1, T\}$ (throughout the planning horizon) and the probability that the corn and soybean revenue process can follow this path. Specifically, a path o on the lattice is defined as $\{(r_1^{c,o}, r_1^{s,o}), (r_2^{c,o}, r_2^{s,o}), \cdots, (r_T^{c,o}, r_T^{s,o})\}$ and has probability p_o to occur. The probability of monoculture occurrence is given as follows:

$$\sum_{o=1}^{N^d} p_o \cdot I(\text{Monoculture occurs on path } o), \tag{29}$$

where N^d is the number of discretized corn and soybean revenue evolution paths, and $I(\cdot)$ is the indicator function that equals one if the condition inside the parentheses holds and equals zero otherwise.

The expected number of periods to achieve monoculture conditional on its occurrence is given as:

$$\frac{\sum_{o=1}^{N^d} p_o \cdot (\text{the period that monoculture first occurs on path } o \text{ if monoculture occurs on path } o)}{\sum_{o=1}^{N^d} p_o \cdot I(\text{Monoculture occurs on path } o)}.$$

Results. Table 3 summarizes the average, minimum, and maximum values for the probability of monoculture occurrence observed under the optimal policy and the one-period lookahead policy on all numerical instances given different values for the planning horizon, that is, $T \in \{5, 10, 15, 20\}$ years. For each value of the planning horizon, Table 4 summarizes the average, minimum, and maximum values for the expected number of periods to achieve monoculture conditional on its occurrence observed under each policy on all numerical instances.

We find that using the optimal policy does not always end up with monoculture: when $T = 5$ years, monoculture only occurs with a probability of 55.27% on average across all numerical instances. Even though this fraction increases as T increases, it increases only to 81.92% even when $T = 20$ years. Conditional on the event that monoculture does occur, the expected number of periods for monoculture to set in increases from 3.87 years on average when $T = 5$ to 8.35 years on average when $T = 20$; again these expected number of periods only reflect the scenarios where monoculture does occur, that is, they do not reflect the scenarios where monoculture does not occur. We observe similar values for these two metrics under the one-period lookahead policy.

E Replicating The Computational Study Using Simulation

In this section, we describe how we replicate our numerical experiments in Section 5.2 using a simulation approach and how we verify that our main insights do not change. For brevity, we only provide the details for the analysis in Section 5.2.4 where we study the performance of a variety of heuristic policies in comparison with the optimal policy.

Simulation Setup. In the simulation study, to compute the relevant performance measures we first generate 10,000 sample paths of corn and soybean revenue pairs using the continuous revenue process as described in (7) in the paper. On each sample path we map the realized corn and soybean revenues to the nearest node on the lattice (which is described in detail in Section 5.2.1 in the paper), and use the parameters from that node for computation. Using

Table 3: The probability of monoculture occurrence under the optimal policy and the one-period lookahead policy

	$T = 5$ (years)		$T = 10$ (years)	
	Optimal	One-period lookahead	Optimal	One-period lookahead
Average	55.27%	53.51%	70.67%	66.87%
Min	0.02%	0.02%	0.03%	0.03%
Max	99.12%	99.12%	100%	100%
	$T = 15$ (years)		$T = 20$ (years)	
	Optimal	One-period lookahead	Optimal	One-period lookahead
Average	77.78%	73.37%	81.92%	77.34%
Min	0.03%	0.03%	0.03%	0.03%
Max	100%	100%	100%	100%

Notes. For each policy, "Average", "Min", and "Max" denote the average, minimum, and maximum values of this metric observed in all numerical instances, respectively.

Table 4: The expected number of periods to reach monoculture (conditional on its occurrence) under the optimal policy and the one-period lookahead policy

	$T = 5$ (years)		$T = 10$ (years)	
	Optimal	One-period lookahead	Optimal policy	One-period lookahead
Average	3.87	4.04	5.82	6.3
Min	2.4	2.4	2.44	2.44
Max	5	5	9.96	10

	$T = 15$ (years)		$T = 20$ (years)	
	Optimal	One-period lookahead	Optimal policy	One-period lookahead
Average	7.23	7.97	8.35	9.33
Min	2.44	2.44	2.44	2.44
Max	14.95	15	19.93	20

Notes. For each policy, "Average", "Min", and "Max" denote the average, minimum, and maximum values of this metric observed in all numerical instances, respectively.

this approach, we compute the relevant performance measures along each sample path and average across all sample paths.

In Section 5.2.4, we make comparisons between the expected profits under any two policies (heuristic or optimal policy). In making this comparison, as standard in the literature, we use the same 10,000 sample paths for computing the expected profit under each policy.[13] After taking the difference between the profits under these two policies on each sample path, we then use a one-sample two-tailed Student t-test with the Null Hypothesis that the difference between the expected profits under these two policies is zero. We note here that the simulation analysis is not carried out for all the 312,500 instances considered in Section 5.2.4, as the computation time in evaluating the expected profits under all heuristic policies and the optimal policy using simulation becomes prohibitively long. Therefore, paralleling the analysis in Devalkar et al. (2011), the simulation analysis is carried out around the baseline scenario (as described in Section 5.2.1 in the paper) while considering different planning horizons. Besides $\hat{T} = 10$ years (as in the baseline scenario) and the additional values $T \in \{5, 15, 20\}$ considered in the paper, we also consider longer planning horizons $T \in \{30, 40, 50, 100\}$.[14] In summary, the simulation analysis is carried out in our baseline scenario by considering eight different planning horizons.

Results. One of our main results in Section 5.2.4 is that the always rotate policy performs better than the always rotate (monoculture) policy. Is the difference between the profits under the two heuristics statistically significant? Table 5 summarizes the mean and the standard deviation of the expected profit under each heuristic obtained from 10,000 sample paths as well as the corresponding p-value from the Student t-test for different planning horizons. We observe that, for each planning horizon, the mean expected profit is higher under the always rotate policy and the profit difference is statistically significant at a significance level of 1% (we reject the Null Hypothesis with a significance level of 1%—that is, a confidence level of 99%—and accept the alternative hypothesis that the difference between the expected profits under these two policies is not zero).

Our second main result in Section 5.2.4 is that the one-period lookahead policy outper-

[13]Using the same sample paths is a variance reduction technique called Common Random Numbers, and it is standard in simulation in comparing the performance of two systems.

[14]Longer planning horizons are considered based on a suggestion by one of the reviewers.

Table 5: Always Rotate Policy Versus Always Rotate (Monoculture) Policy

Horizon	Always rotate policy		Always rotate (monoculture) policy		p-value
T (in years)	expected profits	standard deviation	expected profits	standard deviation	
5	1276.8	400.27	1259.8	394.26	$< 0.01^{(a)}$
10	2530.9	720.57	2507.1	724.36	$< 0.01^{(a)}$
15	3821.1	974.42	3803.6	973.17	$< 0.01^{(a)}$
20	5090.9	1167.2	5067	1170.6	$< 0.01^{(a)}$
30	7647.9	1484.8	7626.7	1488.9	$< 0.01^{(a)}$
40	10209	1743.3	10187	1749.7	$< 0.01^{(a)}$
50	12750	1967.2	12727	1973.3	$< 0.01^{(a)}$
100	25473	2854.6	25447	2861.1	$< 0.01^{(a)}$

Notes. (a) Statistically significant at a significance level of 1%

forms all the heuristic policies considered. Is the difference between the profit under the one-period lookahead policy and the profit under each of the other heuristic policies statistically significant? To answer this question, we conduct a similar comparative analysis between the one-period lookahead policy and the other heuristic policies. Table 6 summarizes the comparison with the myopic policy, Table 7 summarizes the comparison with the always rotate policy, and Table 8 summarizes the comparison with the always rotate (monoculture) policy. We observe that for all the three policies considered and for each planning horizon the mean expected profit is higher under the one-period lookahead policy, and the profit difference is statistically significant at a significance level of 1%

Our final main result in Section 5.2.4 is that the one-period lookahead policy provides a near-optimal performance. Is the difference between the profit under the one-period lookahead policy and the profit under the optimal policy statistically significant? To answer this question, we conduct a similar comparative analysis between the one-period lookahead policy and the optimal policy. Table 9 summarizes this comparison. As expected, for each planning horizon the profit is always higher under the optimal policy, but the profit difference is *not* statistically significant at a significance level of 1% for the planning horizons $T \in \{5, 10, 30, 40, 50\}$.

Conclusion. Based on the simulation analysis, we verify that our main insights in Section 5.2.4 do not change. In particular, (i) it is statistically significant that the always rotate

Table 6: One-period Lookahead Policy Versus Myopic Policy

Horizon	One-period lookahead policy		Myopic policy		p-value
T (in years)	expected profits	standard deviation	expected profits	standard deviation	
5	1287.4	394.19	1282.1	392.12	$< 0.01^{(a)}$
10	2555.2	704.73	2538.4	698.62	$< 0.01^{(a)}$
15	3855.5	954.16	3829.7	944.59	$< 0.01^{(a)}$
20	5139.1	1138.5	5102.5	1127	$< 0.01^{(a)}$
30	7720.2	1451.5	7662	1436.5	$< 0.01^{(a)}$
40	10306	1706.3	10229	1686.9	$< 0.01^{(a)}$
50	12868	1924.2	12770	1898.6	$< 0.01^{(a)}$
100	25707	2789.7	25506	2753.7	$< 0.01^{(a)}$

Notes. (a) Statistically significant at a significance level of 1%

Table 7: One-period Lookahead Policy Versus Always Rotate Policy

Horizon	One-period lookahead policy		Always rotate policy		p-value
T (in years)	expected profits	standard deviation	expected profits	standard deviation	
5	1287.4	394.19	1276.8	400.27	$< 0.01^{(a)}$
10	2555.2	704.73	2530.9	720.57	$< 0.01^{(a)}$
15	3855.5	954.16	3821.1	974.42	$< 0.01^{(a)}$
20	5139.1	1138.5	5090.9	1167.2	$< 0.01^{(a)}$
30	7720.2	1451.5	7647.9	1484.8	$< 0.01^{(a)}$
40	10306	1706.3	10209	1743.3	$< 0.01^{(a)}$
50	12868	1924.2	12750	1967.2	$< 0.01^{(a)}$
100	25707	2789.7	25473	2854.6	$< 0.01^{(a)}$

Notes. (a) Statistically significant at a significance level of 1%

Table 8: One-period Lookahead Policy Versus Always Rotate (Monoculture) Policy

Horizon	One-period lookahead policy		Always rotate (monoculture) policy		p-value
T (in years)	expected profits	standard deviation	expected profits	standard deviation	
5	1287.4	394.19	1259.8	394.26	$< 0.01^{(a)}$
10	2555.2	704.73	2507.1	724.36	$< 0.01^{(a)}$
15	3855.5	954.16	3803.6	973.17	$< 0.01^{(a)}$
20	5139.1	1138.5	5067	1170.6	$< 0.01^{(a)}$
30	7720.2	1451.5	7626.7	1488.9	$< 0.01^{(a)}$
40	10306	1706.3	10187	1749.7	$< 0.01^{(a)}$
50	12868	1924.2	12727	1973.3	$< 0.01^{(a)}$
100	25707	2789.7	25447	2861.1	$< 0.01^{(a)}$

Notes. (a) Statistically significant at a significance level of 1%

Table 9: Optimal policy Versus One-period Lookahead Policy

Horizon	Optimal policy		One-period lookahead policy		p-value
T (in years)	expected profits	standard deviation	expected profits	standard deviation	
5	1287.6	393.81	1287.4	394.19	$0.06^{(b)}$
10	2555.5	704.23	2555.2	704.73	$0.26^{(b)}$
15	3856.6	953.75	3855.5	954.16	$< 0.01^{(a)}$
20	5140.4	1137.9	5139.1	1138.5	$< 0.01^{(a)}$
30	7721.5	1450.4	7720.2	1451.5	$0.049^{(b)}$
40	10307	1704.7	10306	1706.3	$0.432^{(b)}$
50	12870	1923.2	12868	1924.2	$0.053^{(b)}$
100	25712	2788.2	25707	2789.7	$< 0.01^{(a)}$

Notes. (a) Statistically significant at a significance level of 1% (i.e., the difference between the values of the optimal policy and the one-period lookahead policy is statistically significant), which means that the one-period lookahead policy is only near optimal (not optimal) as seen by the relative value of the one-period lookahead policy versus that of the optimal policy. (b) Statistically insignificant at a significance level of 1% (i.e., the difference between the values of the optimal policy and the one-period lookahead policy is *not* statistically significant), which means that the one-period lookahead policy is optimal for such instances.

policy performs better than the always rotate (monoculture) policy; (ii) it is statistically significant that the one-period lookahead policy outperforms all other heuristic policies; (iii) the one-period lookahead policy always provides a near-optimal performance, and sometimes even an optimal performance, a performance that is statistically insignificantly different from that of the optimal policy.

References

Burk, F. 1987. The geometric, logarithmic, and arithmetic mean inequality. *The American Mathematical Monthly*, **94** (6), 527–528.

Cain, M. 1994. The moment-generating function of the minimum of bivariate normal random variables. *The American Statistician*, **48** (2), 124–125.

Devalkar, S. K., R. Anupindi, A. Sinha. 2011. Integrated optimization of procurement, processing, and trade of commodities. *Oper. Res.*, **59** (6), 1369–1381.