1 Modeling Rice Cultivation

In this section we specify and solve our theoretical model of crop cultivation.

With crop cultivation, each sequential stage can be thought of as a separate production subprocess with its own production function. We construct a three-stage rice production function. Within each stage, several operations can be performed simultaneously. Output from the previous stage is an initial condition for the next stage production subprocess. Input decisions are made at the start of each stage, with updated expectations based on history at that point in time. We are able to describe crop cultivation process as a system of equations, with one equation specifying final output as a function of all inputs from all three production stages, and three sets of equations describing input demands in each of the three production stages. As a result, we can perform qualitative analysis of the effects of the magnitude and the timing of production shocks, and in particular weather shocks, on optimal farmer's behavior, crop development, and final yields. We now derive this system of equations.

Let i index the three production stages and let vector $x_i = (x_{i1}\ x_{i2}\ \ldots\ x_{iN_i})'$ denote N_i inputs in stage i. Let ε_i be production shock realized during stage i, with ε_0 describing weather in pre-planting months. Let y_i be the deterministic part of stage i output, with y_0 describing initial conditions of production such as plot characteristics. We assume that crop cultivation process is CES across stages and Cobb-Douglas within stages, with constant returns to scale in both instances. Output in stage i is $f_i(y_{i-1}, x_i, \varepsilon_{i-1}, \varepsilon_i) = y_i \exp(\varepsilon_i)$ for $i = 1, 2, 3$, where f_i is stage i-specific stochastic production process and y_i is stage i-specific CES production function[1]:

where does this equation come from.

$$f_i(y_{i-1}, x_i, \varepsilon_{i-1}, \varepsilon_i) = y_i \exp(\varepsilon_i)\,, \text{ where}$$

share parameter

substitution parameter.

$$y_i(y_{i-1}, x_i, \varepsilon_{i-1}) = A_i \left(\theta_i \left(y_{i-1} \exp(\varepsilon_{i-1})\right)^{\gamma_i} + (1-\theta_i) \left(B_i \prod_{n=1}^{N_i} x_{in}^{\alpha_{in}} \right)^{\gamma_i} \right)^{1/\gamma_i}. \qquad (1)$$

factor productivity

output factor

C-D formula.

For all $i = 1, 2, 3$, $\gamma_i \leq 1$, and A_i and B_i are stage-specific productivity parameters. The elasticity of substitution between production stages i and $i-1$ is equal to $1/(1-\gamma_i)$ and increases in γ_i: $\frac{\partial}{\partial \gamma_i}\left(\frac{1}{1-\gamma_i}\right) = \frac{1}{(1-\gamma_i)^2} > 0$.[2] The order of events in each stage i is as follows. Input decisions x_i are made based on the history of production shocks and intermediate outputs realized in previous stages, and before stage i shocks are realized. Next, production takes place and inputs x_i are used at the same time as current stage production shocks, ε_i, are realized. At the end of the stage, the current stage output, y_i, is observed.

[1] Values of inputs, outputs, and production shocks are plot-specific. Plot indexing is omitted for simplicity of presentation.

[2] Special limit cases are Cobb-Douglas specification with $\gamma_i \to 0$, linear with $\gamma_i = 1$, and Leontieff with $\gamma_i \to -\infty$.

Let p denote the final output price and $w_i = (w_{i1}\ w_{i2}\ ...\ w_{iN_i})$ denote a vector of stage i input prices. At each stage i, expectations are taken conditional on all information available to the farmer at that point, which includes all realized production shocks, intermediate outputs and factor prices from previous stages, and all current stage prices: $I_i = \left\{\{y_j, \varepsilon_i\}_{j=0}^{i-1}, \{w_j\}_{j=1}^{i}\right\}$. Stage 3 information set also includes the final output price, p.

We first solve the farmer's profit maximization problem, and then extend our analysis and maximize farmer's utility to account for farmer's risk aversion. At the beginning of stage 3, the farmer chooses profit-maximizing levels of stage 3 inputs, x_3. At that point, only stage 3 production shock is not yet realized, and the farmer's information set includes realizations of stage 2 output and all production shocks that occurred in earlier stages: $I_3 = \left\{\{y_j, \varepsilon_i\}_{j=0}^{2}, \{w_j\}_{j=1}^{3}, p\right\}$. The farmer solves

(e) total revenue - total cost.

$$\max_{\{x_{3n}\}_{n=1}^{N_3}} \underbrace{pA_3\left(\theta_3\left(y_2 \exp\left(\varepsilon_2\right)\right)^{\gamma_3} + (1-\theta_3)\left(B_3\prod_{n=1}^{N_3} x_{3n}^{\alpha_{3n}}\right)^{\gamma_3}\right)^{1/\gamma_3}}_{y_3} E_3\left[\exp\left(\varepsilon_3\right)\right] - \sum_{n=1}^{N_3} w_{3n}x_{3n},$$

with the expectation of stage 3 production shock, $E_3\left[\exp\left(\varepsilon_3\right)\right]$, conditional on the information set I_3. The first order conditions are

is this real price?

$$E_3\left[\exp\left(\varepsilon_3\right)\right]\frac{\partial y_3}{\partial x_{3n}} = \frac{w_{3n}}{p}, \forall n \in \{1, ..., N_3\} \tag{2}$$

and have the standard interpretation that, at the optimal level, input's marginal product is equal to its real price.

From the first order conditions it follows that $\alpha_{3j}w_{3k}x_{3k} = \alpha_{3k}w_{3j}x_{3j}$ for all $j, k \in \{1, ..., N_3\}$. This lets us express all stage 3 inputs in terms of one stage 3 input, say, x_{31}, as $x_{3k} = \frac{\alpha_{3k}w_{31}}{\alpha_{31}w_{3k}}x_{31}$ for all $k \in \{1, ..., N_3\}$. Solving first for the optimal x_{31}, we can then solve for optimal levels of all stage 3 inputs:

C-D coeff.

$$x_{3k} = \underbrace{\left(\frac{\theta_3}{(1-\theta_3)\, g_3}\right)^{1/\gamma_3}\frac{y_2 \exp\left(\varepsilon_2\right)}{B_3\tilde{\lambda}_3}}_{\text{same } \forall k \in \{1,...,N_3\}} \underbrace{\frac{\alpha_{3k}}{\tilde{w}_{3k}}}_{\text{varies with } k} \quad \forall k \in \{1, ..., N_3\} \tag{3}$$

we use data to estimate structural parameters (coefficients).

$\forall k \in \{1, ..., N_3\}$, where $\tilde{w}_{3k} = \frac{w_{3k}}{p}$, $\tilde{\lambda}_3 = \prod_{n=1}^{N_3}\left(\frac{\alpha_{3n}}{\tilde{w}_{3n}}\right)^{\alpha_{3n}}$,

real price

$$g_3\left(\varepsilon_3, \tilde{w}_3\right) = (1-\theta_3)^{\frac{1}{\gamma_3 - 1}}\left(A_3B_3\tilde{\lambda}_3E_3\left[\exp\left(\varepsilon_3\right)\right]\right)^{\frac{\gamma_3}{\gamma_3-1}} - 1, \tag{4}$$

and $\tilde{w}_i = (\tilde{w}_{i1}\ \tilde{w}_{i2}\ ...\ \tilde{w}_{iN_i})\ \forall i = 1, 2, 3$. As in the standard production literature, the only source of variation across optimal levels of different stage 3 inputs is the ratio of each input's Cobb-Douglas coefficient to its real price.

At the beginning of stage 2, the farmer chooses profit-maximizing levels of stage 2 inputs, x_2, given realized stage 1 output and taking into account his anticipated stage 3 input demand. At this point, the farmer's information set is $I_2 = \left\{ \{y_j, \varepsilon_i\}_{j=0}^{1}, \{w_j\}_{j=1}^{2} \right\}$. The farmer solves

$$\max_{\{x_{2n}\}_{n=1}^{N_2}} E_2[p]\, E_2[y_3 \exp(\varepsilon_3)] - \sum_{n=1}^{N_2} w_{2n} x_{2n} - E_2\left[\sum_{n=1}^{N_3} w_{3n} x_{3n}\right],$$

where expectations are conditional on the information set I_2, y_3 is given in equation (1) and is a function of expected stage 3 inputs demands in equation (3). Substituting equation (3) for stage 3 input demands $\{x_{3n}\}_{n=1}^{N_3}$, we can eliminate stage 3 inputs from the expression for stage 3 production costs, $\sum_{n=1}^{N_3} w_{3n} x_{3n}$, and deterministic stage 3 output, y_3:

$$\sum_{n=1}^{N_3} w_{3n} x_{3n} = \left(\frac{\theta_3}{(1-\theta_3)\, g_3}\right)^{1/\gamma_3} \frac{y_2 \exp(\varepsilon_2)}{B_3 \tilde{\lambda}_3},$$

and

$$y_3 = A_3 \theta_3^{1/\gamma_3} y_2 \left(1 + \frac{1}{g_3}\right)^{1/\gamma_3} \exp(\varepsilon_2). \tag{5}$$

The first order conditions for $x_{2k}, k = 1, ..., N_2$, are

$$\underbrace{A_3 \theta_3^{1/\gamma_3} \frac{\partial y_2}{\partial x_{2k}} E_2\left[\left(\frac{1+g_3}{g_3}\right)^{1/\gamma_3} \exp(\varepsilon_2 + \varepsilon_3)\right]}_{\text{total marginal product}} =$$
$$= \underbrace{\underbrace{\frac{w_{2k}}{E_2[p]}}_{\text{current MC}} + \underbrace{\frac{1}{B_3}\left(\frac{\theta_3}{1-\theta_3}\right)^{1/\gamma_3} \frac{\partial y_2}{\partial x_{2k}} E_2\left[\frac{\exp(\varepsilon_2)}{\lambda_3 p g_3}\right]}_{\text{future MC}}}_{\text{total marginal cost (MC)}}. \tag{6}$$

The optimal stage 2 input levels are

$$x_{2k} = \underbrace{\left(\frac{\theta_2}{(1-\theta_2)\, g_2}\right)^{1/\gamma_2} \frac{y_1 \exp(\varepsilon_1)}{B_2 \tilde{\lambda}_2}}_{\text{same } \forall k \in \{1,...,N_2\}} \underbrace{\frac{\alpha_{2k}}{\tilde{w}_{2k}}}_{\text{varies with } k} \quad \forall k \in \{1, ..., N_2\}, \tag{7}$$

where $\tilde{w}_{2k} = \frac{w_{2k}}{E_2[p]}$, $\tilde{\lambda}_2 = \prod_{n=1}^{N_2} \left(\frac{\alpha_{2n}}{\tilde{w}_{2n}}\right)^{\alpha_{2n}}$,

$$g_2\left(\{\varepsilon_j, \tilde{w}_j\}_{j=2}^3\right) = (1-\theta_2)^{\frac{1}{\gamma_2-1}} \left(A_2 B_2 \tilde{\lambda}_2 E_2 \left[P_3 \exp(\varepsilon_2)\right]\right)^{\frac{\gamma_2}{\gamma_2-1}} - 1, \tag{8}$$

$$P_3(\varepsilon_3, \tilde{w}_3) = \theta_3^{1/\gamma_3} Q_3^{\frac{\gamma_3-1}{\gamma_3}}, \tag{9}$$

and

$$Q_3(\varepsilon_3, \tilde{w}_3) = \left[A_3 \exp(\varepsilon_3)\right]^{\frac{\gamma_3}{\gamma_3-1}} - \left[(1-\theta_3)\left(B_3 \tilde{\lambda}_3\right)^{\gamma_3}\right]^{\frac{1}{1-\gamma_3}}. \tag{10}$$

Note the structural similarity of stage 3 input demands in equation (3) and stage 2 input demands in equation (7). The only difference is the extra term inside the expectation, $P_3(\varepsilon_3, \tilde{w}_3)$, which is a function of information revealed only in stage 3 - stage 3 production shock and prices - and of stage 3 production parameters.

The recursive nature of the model lets us solve the farmer's stage 1 optimization problem following the same steps as we used to optimize his stage 2 input demand. At the beginning of stage 1 the farmer observes only pre-cultivation, initial plot conditions and production shocks that prevailed since the end of the previous cultivation cycle, as well as stage 1 factor prices: $I_1 = \{y_0, \varepsilon_0, w_1\}$. Optimal stage 1 input levels are:

$$x_{1k} = \underbrace{\left(\frac{\theta_1}{(1-\theta_1)\, g_1}\right)^{1/\gamma_1} \frac{y_0 \exp(\varepsilon_0)}{B_1 \tilde{\lambda}_1}}_{\text{same } \forall k \in \{1,\ldots,N_1\}} \underbrace{\frac{\alpha_{1k}}{\tilde{w}_{1k}}}_{\text{varies with } k}, \tag{11}$$

where $\tilde{w}_{1k} = \frac{w_{1k}}{E_1[p]}$, $\tilde{\lambda}_1 = \prod_{n=1}^{N_1} \left(\frac{\alpha_{1n}}{\tilde{w}_{1n}}\right)^{\alpha_{1n}}$,

$$g_1\left(\{\varepsilon_j, \tilde{w}_j\}_{j=2}^3\right) = (1-\theta_1)^{\frac{1}{\gamma_1-1}} \left(A_1 B_1 \tilde{\lambda}_1 E_1 \left[P_2 \exp(\varepsilon_1)\right]\right)^{\frac{\gamma_1}{\gamma_1-1}} - 1, \tag{12}$$

$$P_2\left(\{\varepsilon_j, \tilde{w}_j\}_{j=2}^3\right) = \theta_2^{1/\gamma_2} Q_2^{\frac{\gamma_2-1}{\gamma_2}}, \text{ and}$$

$$Q_2\left(\{\varepsilon_j, \tilde{w}_j\}_{j=2}^3\right) = \left[A_2 P_3 \exp(\varepsilon_2)\right]^{\frac{\gamma_2}{\gamma_2-1}} - \left[(1-\theta_2)\left(B_2 \tilde{\lambda}_2\right)^{\gamma_2}\right]^{\frac{1}{1-\gamma_2}}. \tag{13}$$

Note the structural similarity of input demand equations for stages 1 (equation 11), 2 (equation 7), and 3 (equation 3). The only structural difference between stages 1 and 3 is the extra term, $P_2\left(\{\varepsilon_j, \tilde{w}_j\}_{j=2}^3\right)$, in the expectation inside $g_1\left(\{\varepsilon_j, \tilde{w}_j\}_{j=2}^3\right)$. P_2 is a function of information revealed in stages 2 and 3 - production shocks and prices for stages 2 and 3 - and of production parameters for stages 2 and 3. The only difference between stages 1 and 2 is the extra term,

$P_3(\varepsilon_3, \tilde{w}_3)$, in the expectation inside $Q_2\left(\{\varepsilon_j, \tilde{w}_j\}_{j=2}^3\right)$.

1.1 Risk Aversion

We next demonstrate that the farmer's risk aversion doesn't change any of the qualitative results from our profit maximization analysis. We use a very general utility maximization setup. Weather-induced uncertainty is associated with crop cultivation, when all production choices are made. The post-cultivation period is uncertainty-free. Post-cultivation consumption depends on the farmer's choices during the crop cultivation, both directly through any savings made during cultivation, and indirectly though the input choices, which in turn affect the final yield. The farmer has to make input choices that maximize his combined utility both during crop cultivation and after. We assume no credit constraints.

Let t index calendar years. As before, let stage 0 denote the interim period between the end of last year's harvest and the beginning of this year's crop cultivation. That is, post-cultivation period for year t crop is stage 0 of year $t+1$. The farmer's optimization problem is

$$\max_{\left\{\{c_{it}\}_{i=0}^3\right\}_{t=0}^\infty} E\left[\sum_{t=0}^\infty \beta^t \left(\underbrace{u(c_{1t}) + u(c_{2t}) + u(c_{3t})}_{\text{crop cultivation}} + \delta \underbrace{u(c_{0,t+1})}_{\text{post-cultivation}}\right)\right]$$

subject to the following budget constraints:

$$\begin{array}{rl}
\text{stage 1 of year } t \text{ cultivation:} & R_{0t}s_{0t} \geq c_{1t} + \sum_{n=1}^{N_1} w_{1nt}x_{1nt} + s_{1t} \\
\text{stage 2 of year } t \text{ cultivation:} & R_{1t}s_{1t} \geq c_{2t} + \sum_{n=1}^{N_2} w_{2nt}x_{2nt} + s_{2t} \\
\text{stage 3 of year } t \text{ cultivation:} & R_{2t}s_{2t} \geq c_{3t} + \sum_{n=1}^{N_3} w_{3nt}x_{3nt} + s_{3t} \\
\begin{array}{c}\text{year } t \text{ post-cultivation period} \\ (\text{stage 0 of year } t+1 \text{ cultivation})\end{array} : & R_{3t}s_{3t} + p_t y_{3t}\exp(\varepsilon_{3t}) \geq c_{0,t+1} + s_{0,t+1}.
\end{array} \tag{14}$$

The farmer's utility function $u(.)$ is concave, $\beta \in (0,1)$ is the discount rate across cultivation cycles, $\delta \in (0,1)$ is the discount rate between consumption during and post-cultivation, $R_{it} \geq 1$ and s_{it} denote, respectively, the real interest rate and savings in stage i of year t, and w_{int} and x_{int} denote, respectively, the unit price and amount of input n used in production stage i in year t.[3] The final yield in year t is $y_{3t}\exp(\varepsilon_{3t}) = f\left(y_{0t}, \{x_{it}\}_{i=1}^3, \{\varepsilon_{it}\}_{i=0}^3\right)$, where production function $f(.)$ is given in equation (1). The expectation $E[.]$ is based on the most recent information set the farmer has access to. We augment our original structure of gradual information release throughout the crop cultivation cycle by assuming that at the beginning of each stage i in year

[3] In vector form, $x_{it} = (x_{i1t}\ x_{i2t}\ \ldots\ x_{iN_it})$ and $w_{it} = (w_{i1t}\ w_{i2t}\ \ldots\ w_{iN_it})$ $\forall i = 1,2,3$.

t the farmer learns the interest rate for that period, R_{it}. As always, the farmer retains all previously revealed information:

$$I_{1t} = \{I_{0,t-1}, y_{0t}, \varepsilon_{0t}, w_{1t}, R_{1t}\},$$
$$I_{2t} = \{I_{1t}, y_{1t}, \varepsilon_{1t}, w_{2t}, R_{2t}\},$$
$$I_{3t} = \{I_{2t}, y_{2t}, \varepsilon_{2t}, w_{3t}, R_{3t}, p_t\},$$
$$I_{0,t+1} = \{I_{3t}, y_{3t}, \varepsilon_{3t}, R_{0,t+1}\}.$$

Let μ_{it} denote the Lagrangian multiplier corresponding to stage i, year t budget constraint. The FOCs are:

?

wrt consumption c_{it}:

$$u'(c_{it}) = \mu_{it} \text{ for } i = 1,2,3, \quad \delta u'(c_{0t}) = \mu_{0t} \text{ for } i = 0; \tag{15}$$

wrt savings s_{it}:

$$\mu_{it} = E_{it}[\mu_{i+1,t}] R_{it} \text{ for } i = 0,1,2, \quad \mu_{3t} = E_{3t}[\mu_{0,t+1}] R_{3t} \text{ for } i = 3; \tag{16}$$

wrt stage 1 production inputs x_{1kt}, $k \in \{1, ..., N_1\}$:

$$\mu_{1t} w_{1kt} + E_{1t}\left[\mu_{2t} \sum_{n=1}^{N_2} w_{2nt} \frac{\partial x_{2nt}}{\partial x_{1kt}} + \mu_{3t} \sum_{n=1}^{N_3} w_{3nt} \frac{\partial x_{3nt}}{\partial x_{1kt}}\right] = \tag{17}$$
$$= E_{1t}\left[\mu_{0,t+1} p_t \frac{\partial y_{3t}}{\partial x_{1kt}} \exp(\varepsilon_{3t})\right];$$

wrt stage 2 production inputs x_{2kt}, $k \in \{1, ..., N_2\}$:

$$\mu_{2t} w_{2kt} + E_{2t}\left[\mu_{3t} \sum_{n=1}^{N_3} w_{3nt} \frac{\partial x_{3nt}}{\partial x_{2kt}}\right] = E_{2t}\left[\mu_{0,t+1} p_t \frac{\partial y_{3t}}{\partial x_{2kt}} \exp(\varepsilon_{3t})\right]; \tag{18}$$

and wrt stage 3 production inputs x_{3kt}, $k \in \{1, ..., N_3\}$:

$$\mu_{3t} w_{3kt} = p_t \frac{\partial y_{3t}}{\partial x_{3kt}} E_{3t}[\mu_{0,t+1} \exp(\varepsilon_{3t})]. \tag{19}$$

Note that the FOCs for stage i inputs, x_{ikt}, take into account the effect of the farmer's input choice in stage i on his future input choices in subsequent stages $j \in \{i+1, ..., 3\}$. From the consumption and savings FOCs in equations (15) and (16) we know that, at the optimum,

$$R_{it} = \frac{\mu_{it}}{E_{it}\left[\mu_{i+1,t}\right]} = \frac{u'\left(c_{it}\right)}{E_{it}\left[u'\left(c_{i+1,t}\right)\right]} \text{ for } i = 0, 1, 2, \tag{20}$$

$$\text{and } R_{3t} = \frac{\mu_{3t}}{E_{3t}\left[\mu_{0,t+1}\right]} = \frac{u'\left(c_{3t}\right)}{E_{3t}\left[\delta u'\left(c_{0,t+1}\right)\right]}.$$

Using this result, we can rewrite the expectation inside the FOC for stage 3 input in equation (19) as

$$E_{3t}\left[\mu_{0,t+1}\exp\left(\varepsilon_{3t}\right)\right] = E_{3t}\left[E_{3t}\left[\mu_{0,t+1}\right]\exp\left(\varepsilon_{3t}\right)\right] = E_{3t}\left[\frac{\mu_{3t}}{R_{3t}}\exp\left(\varepsilon_{3t}\right)\right] = \frac{\mu_{3t}}{R_{3t}}E_{3t}\left[\exp\left(\varepsilon_{3t}\right)\right].$$

Substituting this expression for the expectation back into equation (19) and rearranging the terms, we can write the FOC for stage 3 input as

$$\frac{R_{3t}w_{3kt}}{p_t} = \frac{\partial y_{3t}}{\partial x_{3kt}}E_{3t}\left[\exp\left(\varepsilon_{3t}\right)\right] \ \forall k \in \{1, ..., N_3\}. \tag{21}$$

Compare this FOC with respect to the amount of stage 3 input $k \in \{1, ..., N_3\}$ in year t, x_{3kt}, which accounts for the possibility that the farmer is risk-averse, to the FOC with respect to stage 3 input from stage-by-stage profit maximization in equation (2). The only difference between the risk aversion and the profit maximization FOCs is that factor prices are now multiplied by stage 3 real interest rate, R_{3t}, so that the real marginal cost is expressed in stage 3 period terms. Consequently, the profit-maximizing stage 3 input choices in equation (3) are also utility maximizing when we replace real factor prices $\tilde{w}_{3k} = w_{3k}/E_{3t}[p]$ with stage 3 period real factor prices defined as $\hat{w}_{3kt} = R_{3t}w_{3kt}/E_{3t}[p_t]$, and, correspondingly, replace $\tilde{\lambda}_3 = \prod_{n=1}^{N_3}\left(\frac{\alpha_{3n}}{\tilde{w}_{3n}}\right)^{\alpha_{3n}}$ with $\hat{\lambda}_{3t} = \prod_{n=1}^{N_3}\left(\frac{\alpha_{3n}}{\hat{w}_{3nt}}\right)^{\alpha_{3n}} = \prod_{n=1}^{N_3}\left(\frac{\alpha_{3n}}{R_{3t}w_{3nt}/p_t}\right)^{\alpha_{3n}}$.

Similarly, by repeatedly substituting the ratios in equation (20) into the FOCs for stage 2 inputs in equation (18), we derive the following utility-maximizing stage 2 input demands:

$$x_{2kt} = \underbrace{\left(\frac{\theta_2}{\left(1-\theta_2\right)\hat{g}_{2t}}\right)^{1/\gamma_2}\frac{y_{1t}\exp\left(\varepsilon_{1t}\right)}{B_2\hat{\lambda}_{2t}}}_{\text{same } \forall k\in\{1,...,N_2\}} \underbrace{\frac{\alpha_{2k}}{\hat{w}_{2kt}}}_{\text{varies with } k} \ \forall k \in \{1, ..., N_2\}, \tag{22}$$

where $\hat{w}_{2kt} = \frac{R_{2t}w_{2kt}}{E_2\left[p_t\right]}$, and $\hat{\lambda}_{2t} = \prod_{n=1}^{N_2}\left(\frac{\alpha_{2n}}{\hat{w}_{2nt}}\right)^{\alpha_{2n}} = \prod_{n=1}^{N_2}\left(\frac{\alpha_{2n}}{R_{2t}w_{2nt}/p_t}\right)^{\alpha_{2n}}$. There are two differences between the profit-maximizing stage 2 input demand in equation (7) and the utility-maximizing stage 2 input demand in equation (22). First, once again real factor costs are expressed in stage 2 period terms, so that function Q_3 in equation (10) is expressed in terms of $\hat{w}_{3t} = (\hat{w}_{31t}\ \hat{w}_{32t}\ \dots\ \hat{w}_{3N_3t})$: $Q_{3t} = Q_{3t}(\varepsilon_{3t}, \hat{w}_{3t})$. Second, we define function $\hat{P}_{3t}$ that parallels the function P_3 in equation (9) as

$$\hat{P}_{3t}(\varepsilon_{3t}, \hat{w}_{3t}, R_{3t}) = R_{3t}^{-1} P_{3t}(\varepsilon_{3t}, \hat{w}_{3t}) = \theta_3^{1/\gamma_3} R_{3t}^{-1} Q_3^{\frac{\gamma_3 - 1}{\gamma_3}}(\varepsilon_{3t}, \hat{w}_{3t}), \tag{23}$$

and, consequently, redefine function g_2 in equation (8) as

$$\hat{g}_{2t}\left(\{\varepsilon_{jt}, \hat{w}_{jt}\}_{j=2}^{3}, R_{3t}\right) = (1-\theta_2)^{\frac{1}{\gamma_2 - 1}} \left(A_2 B_2 \hat{\lambda}_{2t} E_{2t}\left[\hat{P}_{3t} \exp(\varepsilon_{2t})\right]\right)^{\frac{\gamma_2}{\gamma_2 - 1}} - 1. \tag{24}$$

Compared to the profit-maximizing specification of function P_3 in equation (9), the above utility-maximizing equation for $\hat{P}_{3t}$ has an additional term, R_{3t}^{-1}.

2 Estimation

2.1 Outline of Our Approach

Together, equations (1), (3), (7), and (11) describe the whole cultivation process (I'll refer to these equations as our 'system' from now on). We get measurements of the final yield, $y_3 \exp(\varepsilon_3)$, all inputs (x_1, x_2, x_3), initial cultivation conditions (y_0), all weather shocks $(\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3)$, and all prices ($p$ and all w_i and R_{it}) from the data. We estimate the model with these data, and thus acquire estimates of the model's coefficients, aka structural parameters. In practice, ε_i and y_0 are not scalars but linear combination of multiple measurements; we estimate these linear coefficients together with the structural parameters. indirect inference

Our system of equations is highly nonlinear. To estimate its structural parameters, we use an indirect inference method. It uses an auxiliary log-linear model, which we estimate as SUR. We first choose "true" values of structural parameters, and plug them into our structural system to construct "true" or "actual" or "real" data. We then pretend we don't know these "true" values of structural parameters, and make an initial guess of their values. We plug this initial guess into our structural system to construct the initial set of synthetic data. To improve the accuracy of estimation, we construct not one but M synthetic data - they all use the same initial guess of the structural parameters but different draws of error terms used to construct them. We then estimate the SUR auxiliary model with both true and M synthetic datasets, and compare the estimates of the auxiliary parameters obtained with true data to the average of the estimates obtained with M synthetic datasets. We use a Wald metric for this comparison.

Set up:

OLS

survey data vs. constructed real data.
↳ choose real parameters.
to test our model.

each synthetic data draws a sample like data points lying around the regression line.

β is structural parameters

δ is auxiliary parameters

Log-linear auxiliary model:

$$
y \left\{
\begin{aligned}
\ln y_3 &= f_y(\ln y_0, \ln x_1, \ln x_2, \ln x_3, \ln \varepsilon_0, \ln \varepsilon_1, \ln \varepsilon_2, \ln \varepsilon_3) \\
\ln x_3 &= f_3(\ln y_0, \ln x_1, \ln x_2, \ln w_3, \ln p, \ln \varepsilon_0, \ln \varepsilon_1, \ln \varepsilon_2) \\
\ln x_2 &= f_2(\ln y_0, \ln x_1, \ln w_3, \ln w_2, \ln p, \ln \varepsilon_0, \ln \varepsilon_1) \\
\ln x_1 &= f_1(\ln y_0, \ln w_3, \ln w_2, \ln w_1, \ln p, \ln \varepsilon_0)
\end{aligned}
\right\} x
$$

result term results from descrippen

For now, assume that the error terms e_i in auxiliary model are normal

M sequences of random draws of structural error term $u^{(m)}$, $m \in \{1, \dots, M\}$. These draws $u^{(m)}$ are independent across m

$\hat{\delta}$ is the estimate of auxiliary parameters obtained with real data

$\hat{\delta}_n^{(m)}$ is the estimate of auxiliary parameters obtained with simulated data that was constructed using guess b_n of structural parameters β and sequence m of error terms $u^{(m)}$. Thus, $\hat{\delta}_n^{(m)}$ is a function of guess b_n

synthetic data sets $\rightarrow \bar{\delta}$

$\hat{\beta} = \arg\min_{b_n} \left(\hat{\delta} - \bar{\delta}_n\right)' W \left(\hat{\delta} - \bar{\delta}_n\right)$, where $\bar{\delta}_n = \frac{1}{M}\sum_{m=1}^{M} \hat{\delta}_n^{(m)}$. The expression being optimized, $\left(\hat{\delta} - \bar{\delta}_n\right)' W \left(\hat{\delta} - \bar{\delta}_n\right)$, is the Wald metric

Wald metric

$\rightarrow$ minimize it, then we can find b_n (guess of structural parameters)

W^{-1} is the estimate of the variance-covariance matrix of $\hat{\delta}$ (using real data)

$W^{-1} = V_0 + \frac{1}{M^2}\sum_{m=1}^{M} V_m$, where V_0 is the variance-covariance matrix of auxiliary coefficients obtained with real data, and V_m is the variance-covariance matrix of auxiliary coefficients obtained with m^{th} synthetic data

2.1.1 First Main Challenge: Optimization Process

The idea behind this estimation method is that the Wald metric evaluated at the true values of structural parameters is equal to zero. Thus, our goal is to find a value of structural parameters that minimizes of the Wald metric. This optimization process is highly computationally involved, because each evaluation requires reconstruction of M synthetic datasets with a new (updated) guess of structural parameters, the re-estimation of the SUR system with the updated M synthetic datasets, and because the Wald metric requires computation of the estimate of the variance-covariance matrix for each of these M SUR estimations (as well as that for the true data, but everything involving true data is computed only once). Because the Wald metric is highly non-linear, the optimization routine can get stuck at a local optimum and not be able to (or take too long) to climb out of it.

The goal is to implement this optimization routine in a computationally efficient and precise manner. We achieve computational efficiency by figuring out, for each step involved in recomputing and optimizing the Wald metric:

1. how to make it as mathematically straightforward as possible, in a way that translates into minimal computational demands when implemented in a code

2. how to code it in the most efficient way

Another component to achieving computational efficiency is figuring out the best method to implement the optimization of the Wald metric, in a way that is computationally quick, thorough, and doesn't get trapped at local optima.

Precision refers to the magnitude of the Wald metric computed at our estimate of the structural parameters. At the true value of structural parameters, the Wald metric is zero. In practice, we run the estimation routine until the magnitude of the Wald metric is below the tolerance level we specify. The smaller is the magnitude of this tolerance, the higher is the precision of the estimation routine. Of course, in practice the size of the tolerance depends in large part on how computationally efficient the estimation code is.

2.1.2 Second Main Challenge: Grid-Like Check of Initial Guesses

Because the structural system is highly non-linear, we need to ensure that our estimation results do not depend on our initial guess. To do that, we perform multiple Monte Carlo runs, with each run using a different initial guess. The goal is to do a systematic grid-like check of all plausible values of structural parameters as an initial guess. This, however, is highly computationally challenging, due to the large number of structural parameters, and large theoretically possible intervals for many of them.

These are the theoretical restrictions on the structural parameters:

we want to use data to estimate them

Boundaries	# Parameters	Meaning
$\gamma_i \leq 1$	# of stages (3 so far)	substitution across stages
$\theta_i \in (0, 1)$	# of stages (3 so far)	stage-specific CES shares
$\alpha_{in} \in (0, 1)$, $\sum_{n=1}^{N_i} \alpha_{in} = 1$	total # inputs, i.e. $\sum_{i=1}^{3} N_i$	stage- and input-specific CD shares
$A_i > 0, > 1$	# of stages (3 so far)	stage-specific CES productivity
$B_i > 0, > 1$	# of stages (3 so far)	stage-specific CD productivity
$(-\infty, +\infty)$	all linear coefficients inside $\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3$, and y_0, and those used to construct measures of expectations	

The goal is to design and implement a grid-like search that is thorough and robust on the one hand, yet computationally feasible on the other hand. The need to execute multiple Monte

Carlo runs of the estimation routine underlines the importance of its computational feasibility discussed in section 2.1.1

2.2 Code Flexibility

The code for the algorithm should be sufficiently flexible so that the following changes could be (relatively) easily implemented:

1. Adding a production stage
2. Changing the length of ε_i, x_i (and therefore w_i) and y_0 vectors
3. Changing the way we compute expectations