Organic Farming, Soil Health, and Farmer Perceptions: A Dynamic Structural Econometric Model

Michael A. Meneses, Miguel I. Gómez, David R. Just,
Ravi Kanbur, David R. Lee, and C.-Y. Cynthia Lin Lawell

Abstract

New insights from soil science show that the use of pesticides can be harmful to beneficial soil microbes that improve agricultural yields. Farmers may not be fully aware of soil microbiomes, however, and as a consequence, may not be making optimal decisions about pesticide use and organic farming adoption. In this paper, we develop and estimate a dynamic structural econometric model to examine whether farmers are aware of and account for soil microbiomes and the feedback between pesticides, soil health, pest resistance, and crop yields when making their decisions about pesticide use and organic farming adoption. Empirical results show that farmers are acting as if the clean soil stock has very little effect on crop yields, when in fact it increases yields. Our structural estimates allow us to simulate a number of key outcomes of interest, including pesticide use and farmer welfare under counterfactual scenarios in which farmers' beliefs about soil-microbe based ecosystem services are brought in line with findings from plant and soil sciences. We find that informing farmers about soil microbiomes decreases pesticide use, increases organic adoption, and increases mean farmer welfare in both the short and long run.

Keywords: organic farming, dynamic structural econometric model, perceptions, soil health
***JEL* Codes:** Q12, Q57, Q24
This draft: August 11, 2025

*Meneses: Yale University; michael.meneses@yale.edu. Gómez: Cornell University; mig7@cornell.edu. Just: Cornell University; drj3@cornell.edu. Kanbur: Cornell University; sk145@cornell.edu. Lee: Cornell University; drl5@cornell.edu. Lin Lawell: Cornell University; clinlawell@cornell.edu.

We thank Shadi Atallah, Gary Baker, Nano Barahona, Marc Bellemare, Luca Bertuzzi, Peter Blair Henry, Apurva Borar, Jacob Bradt, Dylan Brewer, Stephanie Brockmann, Vincenzina Caputo, Clare Casteel, Bingyan Dai, Will Davis, Eric Denemark, Avanti Dey, Victor Simoes Dornelas, Alex Fon, Yurou He, Alan Hinds, John Hubbard, Jenny Ifft, Becca Jablonski, Kelsey Jack, Ted Jaenicke, Jerzy Jaromczyk, Kaleb Javier, Melanie Khamis, Joseph Kuvshinikov, Shanjun Li, Tongzhe Li, Francesco Lissoni, Tong Liu, Trevon Logan, Khyati Malik, Conrad Miller, Rosina Moreno, José Nuño-Ledesma, Matthew Oliver, Yu Qin, Jess Rudder, Todd Schmit, Michelle Segovia, Robert Stavins, Michael Toman, Harold van Es, Joakim Weill, Lawrence White, Casey Wichman, Catherine Wolfram, and Hongyu Zhao for detailed and helpful comments and discussions. We benefited from seminar participants at Cornell University, Georgia Institute of Technology, and Bates White; and from conference participants at the Ph.D. Excellence Initiative Summer Research Workshop at New York University; an Association of Environmental and Resource Economists (AERE) Session at the Eastern Economics Association (EEA) Annual Conference (AERE@EEA); the Heartland Environmental and Resource Economics Workshop; the Institute for Operations Research and the Management Sciences (INFORMS) Annual Meeting; the American Economic Association (AEA) Summer Mentoring Pipeline Conference (SMPC); the Northeastern Agricultural and Resource Economics Association (NAREA) Annual Meeting; the Agricultural and Applied Economics Association (AAEA) Annual Meeting; the California Department of Pesticide Regulation (DPR) Pesticide Use Report (PUR) Analysis Workgroup; the Southern Economic Association (SEA) Annual Meeting; the Southeastern Workshop on Energy & Environmental Economics & Policy (SWEEEP); the Interdisciplinary PhD Workshop in Sustainable Development (IPWSD) at Columbia University; the United States Society for Ecological Economics (USSEE) Conference; the International Business Analytics Conference (IBAC); the Economics of Science and Innovation Summer School at the University of Barcelona; the Ph.D. Applied Macroeconomics guest lecture at Cornell University; and the Voices of Quebec Labrador Foundation panel 'The Rise of Young Professionals'. This research was supported by Cornell Center for Social Sciences Cloud Computing Solutions. We received funding from the U.S. Department of Agriculture (USDA) National Institute of Food and Agriculture (NIFA), the Michael Bruun Family Goldman Sachs Scholarship, Cornell Graduate School Conference Travel Grants, Cornell TREESPEAR Research Grants, and the Cornell DEEP-GREEN-RADAR Khaled H. Kheiravar Memorial Scholarship. Gómez, Just, Kanbur, Lee, and Lin Lawell are Faculty Fellows at the Cornell Atkinson Center for Sustainability. All errors are our own.

1 Introduction

Soil microbes benefit agricultural production and improve agricultural yields by enhancing crop nutrient use, stress tolerance, and pest resistance (Verma et al., 2020; Righini et al., 2022; Thiebaut et al., 2022). Recent insights from soil science show that the use of synthetic pesticides and fertilizers can be harmful to these beneficial soil microbes (Li et al., 2022; Blundell et al., 2020; Lori et al., 2017). Thus, while using pesticides and fertilizers may have the initial effect of increasing crop yields, over time these synthetic compounds exert an indirect negative effect on crop yields through their negative effects on soil health. In contrast, organic farming and other production regimes like regenerative agriculture that reduce dependence on synthetic compounds enhance microbial health and, thus, may lead to higher yields in the long run. These insights have important implications for a farmer's optimal strategy regarding synthetic compound use and organic farming adoption (Meneses et al., 2025).

Farmers may not be fully aware of soil microbiomes and the feedback between synthetic compounds, soil health, pest resistance, and crop yields, however, and as a consequence, may not be making optimal decisions about synthetic compound use and organic farming adoption. Previous theoretical research by Meneses et al. (2025) has shown that not being informed about soil bacteria could change behavior in a way that leads farmers to adopt sub-optimal, and even detrimental management practices. Previous experimental analysis by Murphy et al. (2020) has shown that farmers in developing countries usually do not have sufficient information about their soil nutrient levels to make profit-maximizing decisions about fertilizer usage; and that there can be potentially large net benefits to providing farmers with soil information. Research on farmer perceptions of beneficial soil microbes by Ploll et al. (2022) involving in-depth qualitative interviews and a focus group discussion of farmers in Europe shows that farmer find beneficial soil microbes complex and difficult to observe. Farmers may lack a full understanding of the private benefits of management practices that reduce reliance on synthetic agrichemicals (Miller-Klugesherz and Sanderson, 2023). Bloom, Atallah, and Casteel (2024) conduct a survey to investigate the importance of demographic variables and farm characteristics, compared to farmer beliefs, for adoption of practices that support the microbiome, and find that beliefs are key to understanding farmer motivations and promoting organic farming system sustainability.

In this paper, we develop and estimate a dynamic structural econometric model to examine whether farmers are aware of and account for soil microbiomes and the feedback between pesticides, soil health, pest resistance, and crop yields when making their decisions about pesticide use and organic farming adoption. To do so, we first empirically document the insights from soil science that the use of pesticides may increase contemporaneous yields; and also that, over time, not using pesticides increases yields.

Next, in order to understand the beliefs and perceptions of rice farmers that underlie and rationalize their pesticide use decisions, we develop and estimate a dynamic structural econometric model. The structural parameters we estimate include parameters measuring any misperceptions farmers may have about how enhancing microbial health may affect yields. Our structural parameter estimates

therefore enable us to infer farmers' current understanding of the interrelations between soil microbes, pesticides, and crop yields.

Our structural estimates also allow us to simulate a number of key outcomes of interest, including pesticide use and farmer welfare under counterfactual scenarios in which farmers' beliefs about soil-microbe based ecosystem services are brought in line with findings from plant and soil sciences. In particular, we run counterfactual simulations to compare a base simulation of actual behavior under farmer misperception, with a counterfactual simulation of optimal behavior under full information.

Empirical results show that farmers are acting as if the clean soil stock (as measured by the number of previous consecutive years the farmer has not used pesticide) has very little effect on crop yields, when in fact it increases yields. Results of counterfactual simulations show that informing farmers about soil microbiomes and the feedback between pesticides, soil health, pest resistance, and crop yields will decrease pesticide use and increase organic adoption, will increase farmers' net present value (PDV of the entire stream of per-period profits) in the long run, and increases increase farmers' net present value on average in the short run.

2 Literature Review

Our paper builds on several strands of literature. First, our paper builds on the literature on the relationship between pesticide use and farm production and profit. Chambers, Karagiannis, and Tzouvelekas (2010) shows pesticide use as increasing returns to quasi-fixed factors of production like capital and land. In contrast, Jacquet, Butault, and Guichard (2011) use a mathematical programming model to determine whether pesticide use can be reduced without affecting farmer income and find that a up to a 30 percent reduction is possible.

Second, our paper builds on the economics literature on soil health. There is an abundant literature in the soil sciences showing that soil microbes benefit agricultural production and improve agricultural yields by enhancing crop nutrient use, stress tolerance, and pest resistance (Singh et al., 2016; Lori et al., 2017; Yadav et al., 2017; Yibeltie and Sahile, 2018; Blundell et al., 2020; Kalam et al., 2020; Verma et al., 2020; Righini et al., 2022; Thiebaut et al., 2022). Recent insights from soil science show that the use of synthetic pesticides and fertilizers can be harmful to these beneficial soil microbes (Li et al., 2022; Blundell et al., 2020; Dash et al., 2017; Lori et al., 2017; Newman et al., 2016; Kalia and Gosal, 2011; Lo, 2010; Hussain et al., 2009). In the economics literature, Sexton, Lei, and Zilberman (2007) acknowledge the effect that pesticide use can have on soil health through its impact on soil microbiomes. Jaenicke and Lengnick (1999) estimate a soil-quality index consistent with the notion of technical efficiency. van Kooten, Weisensel, and Chinthammit (1990) use a dynamic model that explicitly includes soil quality in the grower's utility function and the trade-off between soil quality (which may decline due to erosion) and net returns. Meneses et al. (2025) develop a dynamic bioeconomic model of a farmer's decisions regarding the use of synthetic compounds (e.g., synthetic fertilizers and pesticides) and the transition from conventional to organic management, accounting

for the interrelationships among synthetic compound use, soil health, and crop yields.

Our paper also builds on the literature on dynamic structural econometric modeling, including related applications of the seminal dynamic structural econometric model developed by Rust (1987, 1988) to agriculture (Scott, 2013; Carroll et al., 2019; Oliva et al., 2020; Carroll et al., 2025a; Yeh, Gómez, and Lin Lawell, 2025; Carroll et al., 2025b; Sambucci, Lin Lawell, and Lybbert, 2025), natural resource management (Timmins, 2002; Huang and Smith, 2014; **?**; Reeling, Verdier, and Lupi, 2020; Araujo, Costa, and Sant'Anna, 2020; Wu et al., 2025), environmental policy (Fowlie, Reguant, and Ryan, 2016; Cook and Lin Lawell, 2020; Blundell, Gowrisankaran, and Langer, 2020; Feger, Pavanini, and Radulescu, 2020; Langer and Lemoine, 2022; Li, Liu, and Wei, 2022; Gillingham et al., 2022; Gerarden, 2023; Toyama, 2024; Butters, Dorsey, and Gowrisankaran, 2025; Weber, 2022; Bradt, 2024), and agricultural groundwater management (Sears, Lin Lawell, and Walter, 2025; Sears et al., 2025b,a).

3 Empirical Application

3.1 Rice Farmers in California

For the empirical application, we wanted to consider a crop for which a farmer who was fully informed about soil microbiomes and the feedback between pesticides, soil health, pest resistance, and crop yields would plausibly adopt a pesticide use and organic adoption strategy that greatly differs from that of farmer who was unaware of soil microbes.

We choose to focus on rice, as several pesticides commonly used to grow rice have been found to harm plant-growth-promoting soil microbes. As a consequence a rice farmer who is knowledgeable about soil microbes and the interactions between pesticide use, soil health, and crop yields may invest in and maintain a non-zero amount of clean soil stock, while one who is unaware of soil microbes may not.

For example, thiobencarb (a pre-emergence herbicide used to control grasses, sedge, and broadleaf weeds around rice crops) harms nitrogen-fixing cyanobacteria that help maintain soil fertility and support crop yields (Dash et al., 2017). Similarly, bensulfuron methyl has been found to reduce soil microbial biomass in a variety of settings (El-Ghamry et al., 2002; Gigliotti et al., 2002), including in waterlogged rice paddies and paddy soil (Saeki and Toyota, 2004; Xie, Huang, and Liu, 2004; Xie and Abid, 2004). Bispyribac sodium has been found to cause a significant decline in microbial biomass carbon in paddy soil (Kumar et al., 2020). Molinate has been found to harm the nitrogen-fixing cyanobacteria *Anabaena cylindrica* (Galhano et al., 2009) and the cyanobacteria *Nostoc muscorum* (Galhano et al., 2010) in rice fields, and to reduce bacterial diversity (Saison et al., 2009). Lambda cyhalothrin has been found to lower bacterial and fungal count (Abubakar and Faizah, 2025) and harm nitrifying and denitrifying bacteria (Cycoń et al., 2006). Propanil has been found to cause a collapse in aerobic bacteria, including bacteria such as Streptomyces and Acinetobacter (Oanh and Duc, 2021) that promote plant growth (Olanrewaju and Babalola, 2021; Mujumdar et al., 2023).

We focus in particular on rice farmers in California. California is the second largest rice-producing state behind Arkansas, producing about $900 million in production value per year (Smith, 2023). Most California rice is medium-grain japonica rice, which is used in Asian and Mediterranean dishes such as sushi, paella, and risotto (Smith, 2023), and the majority of California rice is grown in the Sacramento Valley, where hot days, cool nights, and clay soil that retains moisture create ideal conditions for growing japonica rice (USA Rice, 2020). Pesticides known to be harmful to soil bacteria were used in at least 88.8% of California farmer-field-years that grew rice between 1990 and 2019 in the California Department of Pesticide Regulation's (CA DPR's) Pesticide Use Reporting (PUR) database (California Department of Pesticide Regulation [CA DPR], 2024; Meneses et al., forthcoming).

3.2 National Organic Program

In the United States, the National Organic Program (NOP), which is directed by the U.S. Department of Agriculture (USDA) Agricultural Marketing Service (AMS) and became effective on February 20, 2001, oversees and enforces the integrity of the rigorous USDA organic standards and the accreditation of organic certifiers (USDA Agricultural Marketing Service, 2000b; Organic Produce Network, 2022). Organic is one of the most heavily regulated and closely monitored food systems in the U.S. Any product labeled as organic must be USDA certified (Organic Produce Network, 2022). The National Organic Program establishes national standards for the production and handling of organically produced products, including a National List of substances approved for and prohibited from use in organic production and handling; as well as requirements for labeling products as organic and containing organic ingredients (USDA Agricultural Marketing Service, 2000b).

The organic production and handling requirements of the National Organic Program include the requirement that the field or farm parcel must have had no prohibited substances applied to it for a period of 3 years immediately preceding harvest of the crop (USDA Agricultural Marketing Service, 2000a).

3.3 Data

We use farmer-field-level pesticide use data from the California Department of Pesticide Regulation (DPR) Pesticide Use Reporting (PUR) database (California Department of Pesticide Regulation [CA DPR], 2024). This data includes information about whether unapproved (or prohibited) pesticides were applied on a given farmer-field in a given year. Thus, in this paper we use the terms 'unapproved pesticide', 'synthetic pesticide', and 'pesticide' synonymously.

For crop yield data, we use county-level rice yield data from the U.S. Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) Quick Stats (U.S. Department of Agriculture [USDA], 2024). We use the yield data to estimate a first-stage production function, and then combine the parameter estimates from our first-stage production function with the more disaggregated farmer-field-level pesticide use data to estimate our dynamic structural econometric model of farmer pesticide

use decision-making. As explained in more detail in Section 4, since yield data is only available at the county level, not the farmer-field level, we estimate the production function using county averages via analytically weighted least squares as our method of estimation because the variables in the regression reflect means rather than individual observations (Davidson and MacKinnon, 2004).

For price data, we obtain data on conventional and organic rice prices from the USDA NASS Organic Production Survey (U.S. Department of Agriculture [USDA], 2007, 2012, 2017), the USDA NASS Certified Organic Survey (U.S. Department of Agriculture [USDA], 2023), the University of California Rice Research and Information Center (UC Rice Research and Information Center, 2023), and the University of California at Davis Cost and Return Studies (Espino et al., 2021). From these sources, we obtain annual data on conventional and organic rice prices for 10 years (2005, 2008-2012, 2014-2016, and 2019). We convert prices to real prices in 2010 USD using the consumer price index from the Federal Reserve Bank of Minneapolis (Federal Reserve Bank of Minneapolis, 2024). Figure 1 plots the annual real conventional rice price. Figure 2 plots the real organic price premium $\frac{P_{org}-P_{con}}{P_{org}}$, as calculated using the real conventional price P_{con} and the real organic price P_{org}.[1]

We focus on farmer-fields in the CA DPR PUR data set that plant rice at least once during the 10 years for which we have conventional and organic rice price data. There are 17,695 such farmer-fields that plant rice, comprising 2,101 farmers[2] and spanning 15 counties. These 15 counties are: Butte, Colusa, Fresno, Glenn, Lassen, Merced, Placer, Sacramento, San Joaquin, Solano, Stanislaus, Sutter, Tehama, Yolo, and Yuba. In the annual farmer-field-level panel data set we use for our dynamic structural econometric model, which spans the 10 years for which we have conventional and organic rice prices (2005, 2008-2012, 2014-2016, and 2019), there are 67,230 farmer-field-year observations.

Among the farmer-fields that plant rice at least once during the 10 years of the annual farmer-field-level panel data set we use for our dynamic structural econometric model, in any given year that farmer-field may or may not plant rice, and may or may not use an (unapproved) pesticide. Thus, the possible actions a_{it} for each rice farmer-field in each year are: (1) plant rice and use pesticide ($a_{it} = \text{RC}$), (2) plant rice and do not use pesticide that year ($a_{it} = \text{RN}$), (3) do not plant rice that year and use pesticide that year ($a_{it} = \text{OC}$), and (4) do not plant rice and do not use pesticide that year ($a_{it} = \text{ON}$). Table 1 presents the distribution of a_{it} in the data.

We also include as a state variable the clean soil stock k_{it} as measured by the number of previous consecutive years the farmer has not used any (unapproved) pesticide. In particular, the clean soil stock k_{it} is equal to 0 if the farmer-field used an unapproved synthetic pesticide the previous year, 1 if the farmer-field did not use an unapproved synthetic pesticide the previous year but used an

[1] As explained in more detail in Section 6, the horizontal grey lines in the figures indicate the cutoffs used to discretize our price variables for use in our dynamic structural econometric model.

[2] There are 2,101 unique grower ID's. This grower ID variable is as close as we can get to observing the unique farmers, though the identifier is not perfect. The California Department of Pesticide Regulation (DPR) does not guarantee that 2,101 unique grower IDs means that 2,101 unique farmers appear in the data. This is because farmers can be assigned new ID's if a county ag commissioner decides to do so, or loses track of old ID numbers, and old ID numbers can in theory be reassigned to new farmers once the previous farmer retires. In the conversations we have had with employees at California's county commissioners offices though, they indicated that cases like this are rare, and not the norm.

unapproved synthetic pesticide two years ago, 2 if the farmer-field did not use an unapproved synthetic pesticide the previous two years but used an unapproved synthetic pesticide three years ago, and 3 if the farmer-field did not use an unapproved synthetic pesticide for the previous three years but used an unapproved synthetic pesticide four years ago, and so on.[3]

Table 2 presents the distribution of the clean soil stock k_{it} in the data. In over 83% of the farmer-field-year observations, the farmer-field used an unapproved synthetic pesticide the previous year, resulting in a clean soil stock k_{it} of 0. Approximately 11% of farmer-field-year observations have clean soil stock k_{it} equal to 1, meaning the farmer-field did not use an unapproved synthetic pesticide the previous year but used an unapproved synthetic pesticide two years ago.

Since farmers who never use pesticides may not be included in the PUR data, it is possible that our data set may under-represent farmers who never use pesticides. To gauge how many farmers who never use pesticides get missed by the PUR data, in Appendix B we compare the data on acres planted among farmer-field-years in our data set with USDA data on rice acres planted and on rice acres harvested for the state of California at the county level over the years 1990-2019 from USDA NASS Quick Stats (U.S. Department of Agriculture [USDA], 2025), and find that the PUR data captures approximately 94.5% of the acres of rice planted in the state of California.

4 Production Function

We first estimate a production function in order to empirically assess whether soil microbes do indeed matter for rice. Our production function also serves as an input into our dynamic structural model.

The production function $q(\cdot)$ for the log of rice yield y_{it} is given by:

$$y_{it} = q(c_{it}, k_{it}; \alpha) = \exp\left(\alpha_0 + \alpha_c c_{it} + \alpha_k k_{it}\right), \tag{1}$$

where y_{it} is the rice yield in hundredweights (CWT); c_{it} is a dummy variable for farmer-field i using an unapproved synthetic pesticide in year t; and k_{it} is the clean soil stock as measured by the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide. In particular, the clean soil stock k_{it} is equal to 0 if the rice farmer-field used an unapproved synthetic pesticide the previous year, 1 if the rice farmer-field did not use an unapproved synthetic pesticide the previous year but used an unapproved synthetic pesticide two years ago, 2 if the rice farmer-field did not use an unapproved synthetic pesticide the previous two years but used an unapproved synthetic pesticide three years ago, and 3 if the rice farmer-field did not use an unapproved synthetic pesticide for the previous three years but used an unapproved synthetic pesticide four years ago, and so on.

Since yield data is only available at the county level, not the farmer-field level, we estimate the production function using county averages, where each observation is a county-year in which data on

[3] Our measure of the clean soil stock is a conservative lower bound measure of the clean soil stock, since it does not account for any years the the farmer-field did not use an unapproved synthetic pesticide prior to the most recent year the the farmer-field used an unapproved synthetic pesticide.

rice yield per farmer-field growing rice data is available. In particular, we regress log rice yield per farmer-field growing rice in a county-year on the fraction of farmer-fields growing rice in a county-year who applied an unapproved synthetic pesticide that year and the average over farmer-fields growing rice in a county-year of the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide. Because the variables in the regression reflect means rather than individual observations, the appropriate method of estimation is analytically weighted least squares (Davidson and MacKinnon, 2004), where the weight is the number of farmer-fields growing rice in the county-year. We therefore use inverse variance weights that weight county-years with more farmer-fields growing rice more heavily.

We estimate two types of specifications of our production function. 'Misperception' specifications do not account for soil microbes, and therefore do not include any regressor relating to the clean soil stock k_{it}. 'Full Information' specifications account for soil microbes, and therefore include a regressor relating the clean soil stock k_{it}, the average over farmer-fields growing rice in a county-year of the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide.

Summary statistics for the annual county-level variables we use to estimate our production function are presented in Table 3. We use data from 1990-2019. We start in 1990 since reporting was not mandatory in 1989. There are 15 counties with data for at least one year over 1990-2019. The county-level yield variable reports rice yields in hundredweights (CWT). In North America, 1 hundredweight equals 100 pounds (U.S. Department of Agriculture [USDA], 1992).

Results for our true 'Full Information' production function estimation are in Table 4. As expected, α_c is significant and positive in the 'Full Information' specifications, which means that, in actuality, conditional on soil microbes, the contemporaneous use of an unapproved synthetic pesticide has a positive effect on yield that year. In addition, consistent with recent insights from soil science that show the importance of soil microbes and a clean soil stock, α_k is significant and positive in the 'Full Information' specifications, which means that, in actuality, yields are higher the higher the clean soil stock (as measured by the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide). Thus, in actuality, while using unapproved synthetic pesticides in any given year improves crop yields contemporaneously during the year of application, having used unapproved synthetic pesticides has a negative effect on crop yields in future years. Moreover, since the magnitude of α_c is approximately 9 to 11 times the magnitude of α_k across the two 'Full Information' specification, this means that the clean soil stock resulting from 9 to 11 consecutive years of not using pesticides has roughly the same impact on yields as does contemporaneous pesticide use; and that after at least approximately 10 or 11 consecutive years of not using pesticides, the clean soil stock improves yields more than contemporaneous pesticide use does. Thus, while pesticide use may initially improve yields in the short run, over time not using pesticides improves yields in the long run.

Results for our 'Misperception' production function estimation are in Table 5. As expected, in

the 'Misperception' specifications, α_c is significant and positive, which means that farmers who do not account for soil microbes perceive that using an unapproved synthetic pesticide that year has a positive effect on yield. A farmer who is unaware of soil microbes misperceives the effect of clean soil stock on log yield to be 0; our parameter estimate for α_c from our 'Misperception' production function is our reduced-form estimate for how a farmer who is unaware of soil microbes misperceives the effect of pesticide use on log yield to be.

For robustness, we estimate several alternative specifications of our 'Full Information' and 'Misperception' production functions in Appendix A. In Appendix A.1, we estimate a linear specification of our production function wherein the dependent variable is yield rather than log yield. In Appendix A.2, despite possible issues with the reliability of the acreage data in the PUR data set (California Department of Pesticide Regulation [CA DPR], 2023), we also run a robustness check in which we estimate our production function using county averages calculated by dividing by acres rather than by farmer-field. In Appendix A.3, we estimate specifications of our true 'Full Information' production function using year effects, and using both year effects and county fixed effects. In Appendix A.4, we estimate our true 'Full Information' production function using only annual county-level data from the 10 years of data in the annual farmer-field-level panel data set we use for our dynamic structural econometric model (2005, 2008-2012, 2014-2016, and 2019); these 10 years of data are the 10 years for which we have data on conventional and organic rice prices. Our empirical findings that, consistent with insights from soil science, the use of pesticides may increase contemporaneous yields in the present but not using pesticides increases yields in the future, are robust to these alternative specifications of the production function.

5 Dynamic Structural Model

To understand the beliefs and perceptions of rice farmers that underlie and rationalize their pesticide use decisions as revealed in the data, we develop and estimate a dynamic structural econometric model.

The vector of structural parameters θ we estimate include parameters that measure any farmer misperception of how the use of pesticides and the stock of clean soil affect yield; and parameters in the cost function. Our structural parameter estimates therefore enable us to infer California rice farmers' current understanding of the interrelations between soil microbes, pesticides, and crop yields.

Each year t, each rice farmer-field i chooses an action $a_{it} \in A$. The possible actions for each rice farmer-field in each year are (1) planting rice and using pesticide ($a_{it} = \text{RC}$), (2) planting rice and not using pesticide that year ($a_{it} = \text{RN}$), (3) not planting rice that year and using pesticide that year ($a_{it} = \text{OC}$), and (4) not planting rice and not using pesticide that year ($a_{it} = \text{ON}$). Table 1 presents the distribution of a_{it} in the data.

The per-period payoff $u(\cdot)$ to a farmer-field from choosing action a_{it} at time t depends on the values of the state variables $\mathbf{s_{it}}$ at time t. The state variables $\mathbf{s_{it}}$ at time t include the (discretized)

conventional rice price $P_{con,t}$, the (discretized) organic rice price $P_{org,t}$, and the clean soil stock k_{it} as measured by the number of previous consecutive years the farmer-field has not used pesticide.

In particular, the clean soil stock k_{it} is equal to 0 if the rice farmer-field used pesticide the previous year, 1 if the rice farmer-field did not use pesticide the previous year but used pesticide two years ago, 2 if the rice farmer-field did not use pesticide the previous two years but used pesticide three years ago, and 3 if the rice farmer-field did not use pesticide for the previous three years but used pesticide four years ago, and so on. Table 2 presents the distribution of the clean soil stock k_{it} in the data. For the structural model, we cap k_{it} at 15. In other words, if the rice farmer-field did not use pesticide for the previous 15 years or more, then $k_{it} = 15$. When $c_{it} = 0$, the yield as given by the first-stage production function in Specification (1) from Table 4 when $k_{it} = 14$ and $k_{it} = 15$ is 39,458.31 CWT and 58,864.88 CWT, respectively, which is below and above the maximum yield in the data of 53,000 CWT (Table 3). Thus, we cap k_{it} at $k_{it} = 15$ so that when a farmer-field choose not to use pesticide for many years, the yield does not continue to increase indefinitely far beyond the range of yield in the data.

The farmer's perceived yield (or quantity) $\breve{q}(c_{it}, \mathbf{s_{it}}; \theta)$ from planting rice is given by:

$$\breve{q}(c_{it}, \mathbf{s_{it}}; \theta) = \exp\left(\ln q(c_{it}, k_{it}; \hat{\alpha}) + \gamma_c c_{it} + \gamma_k k_{it}\right), \tag{2}$$

where c_{it} is a dummy variable for farmer-field i using pesticide in year t, k_{it} is the clean soil stock defined above, $q(c_{it}, k_{it}; \hat{\alpha})$ is the estimated yield as given from the production function estimated in the first stage (given by Specification (1) from Table 4), and $\hat{\alpha}$ are the production function parameter estimates from the first-stage production function estimation in Specification (1) from Table 4. For our structural model, yield (or quantity) and perceived yield are in units of 1 million pounds.[4]

The structural parameters θ to be estimated include γ_c, which measures any misperception of how the use of pesticide c_{it} affects log yield; and γ_k, which measures any misperception of how the clean soil stock k_{it} affects log yield.

We use the term 'misperception in yield' to describe the difference between the perceived yield (or quantity) $\breve{q}(\cdot)$ that rationalizes the data and the estimated 'true' yield $q(\cdot)$. This 'misperception in yield' can encapsulate any of a number of reasons why the yield $\breve{q}(\cdot)$ that rationalizes the decisions made by farmers as revealed in the data may differ from the estimated 'true' yield $q(\cdot)$, including farmers not being fully aware of soil microbes and their interactions with pesticide use and yields, farmers being uncertain about soil microbes and their interactions, and farmers behaving suboptimally. Our structural parameters γ_c and γ_k capture the portion of this 'misperception in yield' that can be explained by the use of pesticide c_{it} and the clean soil stock k_{it}, respectively. Thus, the 'misperceptions' that the structural parameters $\gamma = (\gamma_c, \gamma_k)$ measure encompass a broad set of possible phenomena,

[4]While our county-level yield data and the yield that is logged for the dependent variable in our production function estimation both have yield in hundredweights (CWT), or 100 pounds;(U.S. Department of Agriculture [USDA], 1992), for our dynamic structural model, we divide yield and perceived yield by 10,000 in our structural model. Thus, for our dynamic structural model, yield (or quantity) is in units of 1 million pounds.

including misperception, misunderstanding, lack of information, and uncertainty.

The structural parameters γ_c and γ_k are identified from variation in the use of pesticide c_{it}, the clean soil stock k_{it}, and organic adoption among farmers who plant rice, and variation in the use of pesticide c_{it} and the clean soil stock k_{it} between farmers who do and do not plant rice.

We assume the cost of planting rice is given by the following rice production cost function $\text{cost}(c, q; \theta)$:

$$\text{cost}(c_{it}, \breve{q}_{it}; \theta) = \kappa_c c_{it} + \kappa_1 \breve{q}_{it} + \kappa_{cq} c_{it} \breve{q}_{it} + \kappa_2 \breve{q}_{it}^2 + \kappa_{cq2} c_{it} \breve{q}_{it}^2, \tag{3}$$

where the cost parameters $\kappa = (\kappa_c, \kappa_1, \kappa_{cq}, \kappa_2, \kappa_{cq2})$ are among the structural parameters to be estimated.

If farmer-field i plants rice in year t ($a_{it} \in \{\text{RC}, \text{RN}\}$), then that farmer-field receives the following deterministic payoff $u_0(\cdot)$ in year t:

$$u_0(a_{it} \in \{\text{RC}, \text{RN}\}, \mathbf{s_{it}}; \theta) = P(a_{it}, \mathbf{s_{it}}) \cdot \breve{q}(c_{it}, \mathbf{s_{it}}; \theta) - \text{cost}(c_{it}, \breve{q}(c_{it}, \mathbf{s_{it}}; \theta); \theta), \tag{4}$$

where $P(a_{it}, \mathbf{s_{it}})$ is the relevant rice price. If farmer-field i plants rice and uses pesticide in year t ($a_{it} = \text{RC}$), then that farmer-field is necessarily a conventional farmer-field and the relevant rice price is the conventional rice price: $P(a_{it}, \mathbf{s_{it}}) = P_{con,t}$. If farmer-field i plants rice and does not use pesticide in year t ($a_{it} = \text{RN}$) but the farmer-field used pesticide within the previous three years $k_{it} < 3$, then that farmer-field is a conventional farmer-field and the relevant rice price is the conventional rice price: $P(a_{it}, \mathbf{s_{it}}) = P_{con,t}$. If farmer-field i plants rice and does not use pesticide in year t ($a_{it} = \text{RN}$) and the farmer-field did not use pesticide in any of the previous three years $k_{it} \geq 3$, then that farmer-field is an organic farmer-field and the relevant rice price is the organic rice price: $P(a_{it}, \mathbf{s_{it}}) = P_{org,t}$.

So that we can identify the parameters in the per-period payoff, we normalize the deterministic payoff for a farmer-field who does not plant rice and uses pesticide ($a_{it} = \text{OC}$) to be 0:

$$u_0(a_{it} = \text{OC}, \mathbf{s_{it}}; \theta) = 0. \tag{5}$$

To allow the per-period payoff to possibly differ when the farmer-field did not plant rice and also did not use pesticide, the deterministic per-period payoff for not planting rice that year and not using pesticide ($a_{it} = \text{ON}$) is set equal to a parameter ν to be estimated:

$$u_0(a_{it} = \text{ON}, \mathbf{s_{it}}; \theta) = \nu. \tag{6}$$

In addition to the state variables $\mathbf{s_{it}}$, the per-period payoff $u(\cdot)$ to a farmer-field from choosing action a_{it} at time t also depends on the choice-specific shock $\epsilon_{it}(a_{it})$ at time t. This choice-specific shock $\epsilon_{it}(a_{it})$ is an unobservable state variables that farmer-fields observe (but we do not observe) when they make their decisions regarding whether to plant rice and whether to use pesticides that

year t. There is a choice-specific shock $\epsilon_{it}(a_{it})$ associated with each possible action $a_{it} \in A$. The vector of choice-specific shocks $\epsilon_{it} \equiv \{\epsilon_{it}(a_{it})|a_{it} \in A\}$ is observed by farmer-field i at time t, before farmer-field i makes his time-t action choice, but is never observed by the econometrician.

The per-period payoff to a farmer-field from choosing action a_{it} at time t is given by:

$$u(a_{it}, \mathbf{s_{it}}, \epsilon_{it}; \theta) = u_0(a_{it}, \mathbf{s_{it}}; \theta) + \epsilon_{it}(a_{it}), \tag{7}$$

where $u_0(\cdot)$ is the deterministic component of the per-period payoff.

The structural parameters to be estimated are $\theta = (\gamma, \kappa, \nu)$, where γ are misperception parameters, κ are cost parameters, and ν is the deterministic payoff when a farmer-field does not plant rice and does not use pesticide (this is relative to a deterministic payoff of 0 when a farmer-field does not plant rice and uses pesticide).[5]

We assume the state variables evolve as a finite state first-order Markov process, with a transition density given by $\Pr(\mathbf{s}_{t+1}, \epsilon_{t+1}|\mathbf{s}_t, a_t, \epsilon_t, \theta)$. Since the conventional rice price and organic price premium variables we use are discretized annual averages, we assume that the crop and pesticide use decisions of any one farmer-field would not have a large enough effect to influence crop prices, and therefore that the distribution of discretized annual conventional rice price and organic rice price premium next period does not depend on any single grower's decisions this period; we therefore model rice prices as evolving exogenously. For the clean soil stock k_{it} as measured by the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide, the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide evolves deterministically as a function of this period's value of the clean soil stock k_{it} and this period's pesticide use decision.

We make the following conditional independence assumption on the transition density:

$$\Pr(\mathbf{s}_{t+1}, \epsilon_{t+1}|\mathbf{s}_t, a_t, \epsilon_t, \theta) = \Pr(\epsilon_{t+1}|\mathbf{s}_{t+1}, \theta) \Pr(\mathbf{s}_{t+1}|\mathbf{s}_t, a_t, \theta). \tag{8}$$

We also assume that the choice-specific shocks are distributed multivariate extreme value. A standard assumption in many dynamic structural models, our conditional independence assumption implies that, conditional on the current state variables $\mathbf{s}_{it}$ and the current action a_{it} chosen by the farmer-field, the evolution of the observed state variables $\mathbf{s}_{it}$ does not depend on the particular realization of the idiosyncratic shocks ϵ_{it} to the payoffs of individual farmers from each possible crop and pesticide use choice. For the clean soil stock k_{it} as measured by the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide, the conditional independence assumption makes sense since the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide evolves deterministically as a function of this period's value of the clean soil stock k_{it} and this period's action. For the conventional and organic rice prices, which are stochastic, since there are many growers in California and no grower has a significant market share,

[5] For our dynamic structural model, payoffs (and therefore costs) are in units of ten thousand 2010 USD.

it is reasonable to assume that shocks to any particular individual grower are unlikely to affect what the discretized annual rice prices are at the aggregate level for all growers.

Although some information in the data is lost by discretizing the variables, one advantage of having to use discretized variables in our structural model is that by discretizing the conventionalprice and the organic price premium, our results are robust to the exact value of the conventional price and organic price premium within our broader size bins. As a consequence, it is possible that our discretized conventional price and discretized organic price premium may arguably also capture costs of other inputs and practices that might vary between conventional and organic management, but still lie within the respective broader size bins, thereby enabling us to interpret the discretized organic price premium as a net organic price premium. Likewise, our results are robust to the possibility that high rates of organic adoption may influence prices within our broader size bins. Thus, the discretization of our price variables is not only necessary for our dynamic structural econometric model, but also enables us to best capture additional considerations that might affect the (net) conventional price and the (net) organic price premium within our broader size bins given data availability and computational constraints, and in a manner robust to any additional assumptions or imprecision that may be introduced if we were to instead finely model every last detail of every aspect of conventional and organic rice production for each and every individual rice farmer.

Under the assumptions that the state variables and the choice-specific shocks ϵ_{it} are conditionally independent and that the choice-specific shocks ϵ_{it} are distributed multivariate extreme value, the farmer's value function, which gives the present discounted value of the grower's entire stream of per-period payoffs at the optimum, is given by the following infinite-horizon Bellman equation:

$$V(\mathbf{s_{it}}, \epsilon_{it}; \theta) = \max_{a_{it} \in A} u_0(a_{it}, \mathbf{s_{it}}; \theta) + \epsilon_{it}(a_{it}) + \beta V^c(\mathbf{s_{it}}, a_{it}; \theta), \tag{9}$$

where $V^c(\cdot)$ is the continuation value, which is the expected value of the value function next period conditional on the state variables and action this period:

$$V^c(\mathbf{s_{it}}, a_{it}; \theta) = E[V(\mathbf{s_{it}}, \epsilon_{it}; \theta) | s_{it}, a_{it}] \tag{10}$$

and where β is the annual discount factor. The choice probability is given by:

$$\Pr(a_{it} | \mathbf{s_{it}}, \theta) = \frac{\exp\left(u_0(a_{it}, \mathbf{s_{it}}; \theta) + \beta V^c(\mathbf{s_{it}}, a_{it}; \theta)\right)}{\sum_{\tilde{a}_{it} \in A} \exp\left(u_0(\tilde{a}_{it}, \mathbf{s_{it}}; \theta) + \beta V^c(\mathbf{s_{it}}, \tilde{a}_{it}; \theta)\right)}. \tag{11}$$

After obtaining the model predictions for the choice probabilities as functions of the state variables and the unknown parameters θ, we estimate the parameters θ using the nested fixed point maximum likelihood estimation technique developed by Rust (1987, 1988). The likelihood function is a function of the choice probabilities, and therefore a function of the continuation value $V^c(\cdot)$. For each guess of the parameters θ, we solve for the continuation value $V^c(\cdot)$ by solving for a fixed point, and use the continuation value to solve for the choice probabilities, which we then plug into the likelihood

function. From Blackwell's Theorem, the fixed point is unique. An inner fixed point algorithm to compute the continuation value $V^c(\cdot)$ is nested within an outer optimization algorithm to find the maximizing value of the parameters θ via maximum likelihood estimation (MLE).

Identification of the parameters θ comes from the differences between per-period payoffs across different action choices, which in infinite horizon dynamic discrete choice models are identified when the discount factor β and the distribution of the choice-specific shocks ϵ_{it} are fixed (Abbring, 2010; Magnac and Thesmar, 2002; Rust, 1994). We set our annual discount factor to $\beta = 0.9$.[6] In particular, the parameters θ in our model are identified because we normalize the deterministic payoff $u_0(a_{it} = \text{OC}, \mathbf{s_{it}}, \theta)$ for a farmer-field who does not plant rice and uses unapproved chemical input to be 0; as a consequence, the parameters do not cancel out in the differences between per-period payoffs across different action choices and are therefore identified.

Standard errors are formed by a nonparametric bootstrap. Farmer-fields are randomly drawn from the data set with replacement to generate 100 independent panels each with the same number of farmer-fields as in the original data set. The structural model is run on each of the new panels. The standard errors are then formed by taking the standard deviation of the parameter estimates from each of the panels.

6 Transition Density for Prices

We use two state variables for price: the real conventional price P_{con}; and the real organic price premium $\frac{P_{org}-P_{con}}{P_{org}}$.

Since the conventional and organic rice price variables we use are discretized annual averages, we assume that the crop and pesticide use decisions of any one farmer-field would not have a large enough effect to influence crop prices, and therefore that the distribution of the discretized annual conventional rice price and organic rice price premium next period does not depend on any single grower's decisions this period; we therefore model rice prices as evolving exogenously. Moreover, we assume that the real conventional price and the real organic price premium are each independent and identically distributed with no serial correlation. Furthermore, since we are now using the real organic price premium instead of the real organic price, we assume that the real conventional price and the real organic price premium are distributed independently of one another. We assume no serial correlation for the real conventional price and the real organic price premium both because it appears to be a reasonable approximation given the time series behavior of these variables (see Figures 1 and 2), and also due to data limitations: our annual price data only provides a very limited amount of consecutive observations to credibly estimate a first-order Markov transition matrix for prices.[7]

[6] An annual discount factor of 0.9 is commonly assumed in the literature using dynamic models (see e.g., Ryan (2012); Lin (2013); Sears, Lim, and Lin Lawell (2019); Cook and Lin Lawell (2020)).

[7] Since we only have price data for 10 years of data (2005, 2008-2012, 2014-2016, 2019), we only have 10 observations for estimating the transition density for prices. Since we do not have price data for 2006-2007, if we try to estimate a first-order Markov transition density for prices, we cannot really use information from the 2005 prices either, so any

When discretizing our price variables, we choose the bins for the variables so that we do not have any price tuples that we do not observe in the data. Having tuples we do not observe in the data leads to issues including losing variation/information, which impedes our ability to identify parameters.

We discretize the real conventional price P_{con} into 2 bins, 'low' and 'high', where a real conventional price less than \$19.50 (in 2010 USD per CWT) is in the 'low' bin, and a real conventional price greater than or equal to \$19.50 (in 2010 USD per CWT) is in the 'high' bin. Figure 1 plots the annual real conventional rice price, along with a horizontal dashed grey line indicating the cutoff for the bins. For each discretized bin for conventional prices, we use the average over all annual real conventional price values that fall in that bin for the value of the real conventional price to use for that bin. In particular, real conventional prices in the 'low' bin are assigned a real conventional price value of \$16.4604 (in 2010 USD per CWT), and real conventional prices in the 'high' bin are assigned a real conventional price value of \$23.3469 (in 2010 USD per CWT).

For the distribution for the discretized real conventional price, we use the empirical distribution of the discretized real conventional price. In other words, for each bin, we assume that the probability that the real conventional price in any given year is in that bin is the fraction of years with real price data that have real conventional price in that bin. The empirical distribution of discretized real conventional price has $\Pr(P_{con} = \text{low}) = 0.5$ and $\Pr(P_{con} = \text{high}) = 0.5$.

We discretize the real organic price premium $\frac{P_{org}-P_{con}}{P_{org}}$ into 3 bins: 'low', 'med', and 'high', where a real organic price premium less than \$0.50 (in 2010 USD per CWT) is in the 'low' bin; a real organic price premium greater than or equal to \$1 (in 2010 USD per CWT) is in the 'high' bin; and real organic price premia in between the 2 cutoffs are in the 'med' bin. Figure 2 plots the real organic price premium $\frac{P_{org}-P_{con}}{P_{org}}$, along with horizontal dashed grey lines indicating the cutoffs for the bins. For each discretized bin for the real organic price premium, we use the average over all annual the real organic price premium values that fall in that bin for the value of the real organic price premium to use for that bin. In particular, real organic price premia in the 'low' bin are assigned a real organic price premium value of \$0.2178 (in 2010 USD per CWT); real organic price premia in the 'med' bin are assigned a real organic price premium value of \$0.6186 (in 2010 USD per CWT); and real organic price premia in the 'high' bin are assigned a real organic price premium value of \$1.2326 (in 2010 USD per CWT).

---

first-order Markov transition density we estimate essentially does not include any information for prices for the years 2005-2007. Similarly, since we do not have price data for 2017-2018, if we try to estimate a first-order Markov transition density for prices, we cannot really use information from the 2019 prices either, so any first-order Markov transition density we estimate essentially does not include any information for prices for the years 2017-2019 either. Thus, if we try to estimate a first-order Markov transition density for prices, we are essentially only using data and information from the years 2008-2016 and not using the data and information we have for 2005 and 2019. We are also missing price data for 2013, which further limits the amount of consecutive information we have to credibly estimate a first-order Markov transition matrix for prices. If the organic price premium were to exhibit positive serial correlation, then it is possible that farmers who face a lower organic price premium may be slower to adopt organic (relative to the case in which the organic price premium is not serially correlated), while farmers who face a higher organic price premium may adopt organic more quickly. Nevertheless, based on the time series plots for real conventional price and real organic price premium, the assumption of no serial correlation seems reasonable.

For the distribution for the discretized real organic price premium, we use the empirical distribution of the discretized real organic price premium. In other words, for each bin, we assume that the probability that the real organic price premium in any given year is in that bin is the fraction of years with real price data that have the real organic price premium in that bin. The empirical distribution of discretized real organic price premium has $\Pr(\frac{P_{org}-P_{con}}{P_{org}} = \text{low}) = 0.2$, $\Pr(\frac{P_{org}-P_{con}}{P_{org}} = \text{med}) = 0.6$, and $\Pr(\frac{P_{org}-P_{con}}{P_{org}} = \text{high}) = 0.2$.

7 Results

The structural parameter estimates are presented in Table 6.

Across all specifications, γ_k, which measures any misperception of how log yield is affected by the clean soil stock k_{it} (the number of previous consecutive years (up to 15 years) no unapproved synthetic pesticide was used), is negative. Thus, farmers underestimate how the clean soil stock k_{it}, as measured by the number of previous consecutive years no unapproved synthetic pesticide was used, affects yield. The net perceived effect of clean soil stock k_{it} on log yield is given by the sum of γ_k and the coefficient $\hat{\alpha}_k$ on the clean soil stock k_{it} from our production function estimation. Given $\hat{\alpha}_k = 0.400$ from Specification (1) in Table 4 and $\gamma_k = -0.285$ across all specifications in Table 6, the net perceived effect on log yield of the clean soil stock k_{it}, the number of previous consecutive years (up to 15 years) no unapproved synthetic pesticide was used, is 0.115, which is less than 30% of the actual effect. In other words, farmers are acting as if the clean soil stock k_{it} has very little effect on rice crop yields, when in fact it increases yields.

Across all specifications in Table 6, γ_c, which measures any misperception of how the use of an unapproved synthetic pesticide c_{it} affects log yield, is negative. The net perceived effect of the use of an unapproved synthetic pesticide c_{it} on log yield is given by the sum of γ_c and the coefficient $\hat{\alpha}_c$ on the use of an unapproved synthetic pesticide c_{it} from our production function estimation. Given $\hat{\alpha}_c = 3.757$ from Specification (1) in Table 4 and γ_c ranges from -0.839 to -0.836 in Table 6, the net perceived effect of the use of an unapproved synthetic pesticide c_{it} on log yield ranges from 2.918 to 2.921. This is roughly similar to the coefficient on the use of an unapproved synthetic pesticide c_{it} in the 'Misperception' specification of our production function in Specification (1) of Table 5, which was 3.011. Thus, our structural estimate for γ_c, which measures any misperception of how the use of an unapproved synthetic pesticide c_{it} affects log yield, is consistent with the reduced-form 'Misperception' production function parameter estimates.

To distinguish between misperception and risk aversion, in Appendix C.1 we also consider a specification of our dynamic structural econometric model that incorporates risk aversion by using a constant absolute risk aversion (CARA) utility function for the deterministic payoff (instead of a linear utility function) when a farmer plants rice. We find that farmers have low levels of risk aversion and appear close to risk neutral, and our results regarding farmers' misperception of the effects of clean soil stock and pesticide use on yield are robust to the inclusion of risk aversion. Farmers' choices

regarding pesticide use therefore reflect misperception and cannot be completely explained away by risk aversion.

In Appendix C.2 we also consider a specification of our dynamic structural econometric model that incorporates risk aversion by using a constant relative risk aversion (CRRA) utility function for the deterministic payoff (instead of a linear utility function) when a farmer plants rice. Results suggest that farmers' risk aversion lie somewhere between risk neutral and low risk aversion.

8 Counterfactual Simulations

Our structural estimates allow us to simulate a number of key outcomes of interest, including pesticide use and farmer welfare under counterfactual scenarios in which farmers' beliefs about soil-microbe based ecosystem services are brought in line with findings from plant and soil sciences. In particular, we use the estimated parameters from our dynamic structural econometric model in Table 6 to compare a base simulation of actual behavior under farmer misperception, with a counterfactual simulation of optimal behavior under full information.

We first run a base simulation of actual behavior ('Misperception'), which uses the parameter estimates from our dynamic structural econometric model in Table 6 for all the parameters. For each farmer-field, we start from the actual value of the state (clean soil stock k_{it} and prices) in the first year for which we have data for that farmer-field, and then forward simulate actions and states for each year from that year onwards to the final year of our data set ('short run') and also to 10 years past the the final year of our data set ('long run'). In particular, for each year of our simulation, we draw the action a_{it} for that year from the choice probabilities $\Pr(a_{it}|\mathbf{s_{it}}, \theta)$ evaluated at the state for that year and at the estimated parameters $\hat{\theta}$. Based on that year's action a_{it} and clean soil stock k_{it}, we determine the clean soil stock $k_{i,t+1}$ for next year. For the next year's real conventional price and real organic price premium, we draw from their respective distributions. After one simulation for all farmer-years, we tabulate the actions a_{it} and calculate summary statistics for clean soil stock k_{it}, the number of organic farmers, the mean of the true yield, the mean of the perceived yield, and the PDV of the entire stream of true per-period payoffs (calculated using the true yield). We repeat the simulation 100 times, and average over 100 simulations.

We next run a counterfactual simulation of optimal behavior ('Full Information'), in which we set the misperception parameters γ_k and γ_c both to 0, and then use the parameter estimates from our dynamic structural econometric model in Table 6 for all the remaining parameters. In other words, under full information, rice farmers no longer misperceive the effects of pesticide use and clean soil stock on yield, but instead correctly perceive rice yield as given by the production function in Specification (1) from Table 4. For each farmer-field, we start from the actual value of the state (clean soil stock k_{it} and prices) in the first year for which we have data for that farmer-field, and then forward simulate actions and states for each year from that year onwards to the final year of our data set ('short run') and also to 10 years past the the final year of our data set ('long run').

In particular, for each year of our simulation, we draw the action a_{it} for that year from the choice probabilities $\Pr(a_{it}|\mathbf{s_{it}}, \theta)$ evaluated at the state for that year, $\gamma_k = 0$, $\gamma_c = 0$, and at the estimated parameters $\hat{\theta}$ for the remaining parameters. Based on that year's action a_{it} and clean soil stock k_{it}, we determine the clean soil stock $k_{i,t+1}$ for next year. For the next year's real conventional price and real organic price premium, we draw from their respective distributions. After one simulation for all farmer-years, we tabulate the actions a_{it} and calculate summary statistics for clean soil stock k_{it}, the number of organic farmers, the mean of the true yield, and the PDV of the entire stream of true per-period payoffs (calculated using the true yield). We repeat the simulation 100 times, and average over 100 simulations.

The results for the short run are presented in Tables 7 and the results for the long run are presented in Table 8. In these tables, Specifications (1), (2), and (3) use parameter estimates from Specifications (1), (2), and (3), respectively, of Table 6. We find that the PDV of the entire stream of per-period payoffs is higher under full information than under misperception, both in the short run and the long run, and the margin widens as we move into the long run. Under full information, pesticide use is lower and organic adoption is higher, and this difference is more pronounced in the long run. Under farmer misperception, less than 0.1 percent of farmer-field-years produce organically in both the short and long run. In contrast, under full information, 55 to 58 percent of farmer-field-years produce organically in the short run, and 58 to 62 percent produce organically in the long run. The mean, 75th percentile, and 95th percentile PDV of the entire stream of per-period payoffs are higher under full information than under misperception in both the short run and the long run. While the 5th percentile, 25th percentile, and median (50th percentile) PDV of the entire stream of per-period payoffs are lower under full information than under misperception in the short run, they are all higher under full information than under misperception in the long run.

9 Conclusion

We develop and estimate a dynamic structural econometric model to examine whether farmers are aware of and account for soil microbiomes and the feedback between pesticides, soil health, pest resistance, and crop yields when making their decisions about pesticide use and organic farming adoption.

We first empirically document the insights from soil science that the use of pesticides may increase contemporaneous yields; and also that, over time, not using pesticides increases yields. In particular, our production function parameter estimates show that the clean soil stock resulting from 9 to 11 consecutive years of not using pesticides has roughly the same impact on yields as does contemporaneous pesticide use; and that after at least approximately 10 or 11 consecutive years of not using pesticides, the clean soil stock improves yields more than contemporaneous pesticide use does. Thus, while pesticide use may initially improve yields in the short run, over time not using pesticides improves yields in the long run.

Next, in order to understand the beliefs and perceptions of rice farmers that underlie and rationalize their pesticide use decisions as revealed in the data, we develop and estimate a dynamic structural econometric model. The structural parameters we estimate include parameters measuring any misperceptions farmers may have about how enhancing microbial health may affect yields. Empirical results show that farmers are acting as if the clean soil stock has very little effect on rice crop yields, when in fact it increases yields.

Our structural estimates allow us to simulate a number of key outcomes of interest, including pesticide use and farmer welfare under counterfactual scenarios in which farmer's beliefs about soil-microbe based ecosystem services are brought in line with findings from plant and soil sciences. We find that informing farmers about soil microbiomes and the feedback between pesticides, soil health, pest resistance, and crop yields will decrease pesticide use and increase organic adoption, will increase farmers' net present value (PDV of the entire stream of per-period profits) in the long run, and increases increase farmers' net present value on average in the short run. These results have important implications regarding the possible effects and benefits of extension programs targeting farmers' understanding of soil microbes.

References

Abbring, J. 2010. "Identification of Dynamic Discrete Choice Models." *Annual Review of Economics* 2:367–394.

Abubakar, G., and A. Faizah. 2025. "Impact of lambda cyhalothrin on bacterial and fungal species in soil." *Bulgarian Journal of Soil Science* 10:14–23.

Araujo, R., F.J.M. Costa, and M. Sant'Anna. 2020. "Efficient Forestation in the Brazilian Amazon: Evidence from a Dynamic Model." Working paper, FGV EPGE and University of Delaware.

Bloom, E.H., S.S. Atallah, and C.L. Casteel. 2024. "Motivating organic farmers to adopt practices that support the pest-suppressive microbiome relies on understanding their beliefs." *Renewable Agriculture and Food Systems* 39:e8.

Blundell, R., J.E. Schmidt, A. Igwe, A.L. Cheung, R.L. Vannette, A.C. Gaudin, and C.L. Casteel. 2020. "Organic management promotes natural pest control through altered plant resistance to insects." *Nature Plants* 6:483–491.

Blundell, W., G. Gowrisankaran, and A. Langer. 2020. "Escalation of Scrutiny: The Gains from Dynamic Enforcement of Environmental Regulations." *American Economic Review* 110:2558–2585.

Bradt, J.T. 2024. "A policy by any other name: Unconventional industrial policy in the US residential solar industry." Working paper, University of Texas at Austin.

Butters, R.A., J. Dorsey, and G. Gowrisankaran. 2025. "Soaking Up the Sun: Battery Investment, Renewable Energy, and Market Equilibrium." *Econometrica* 93:891–927.

California Department of Pesticide Regulation [CA DPR]. 2023. "Pesticide Use Annual Report: Data Access, References, and Definitions Guide." [Online]. Accessed 3 May 2025. URL: `https://www.cdpr.ca.gov/wp-content/uploads/2024/07/pur_references_definitions.pdf`.

—. 2024. "Pesticide Use Reporting (PUR) Archive." [Online]. Accessed 1 May 2023. URL: `https://files.cdpr.ca.gov/pub/outgoing/pur_archives/`.

Carroll, C.L., C.A. Carter, R.E. Goodhue, and C.Y.C. Lin Lawell. 2019. "Crop Disease and Agricultural Productivity: Evidence From a Dynamic Structural Model of Verticillium Wilt Management." In W. Schlenker, ed. *Agricultural Productivity and Producer Behavior*. Chicago: University of Chicago Press, pp. 217–249.

—. 2025a. "The economics of decision-making for crop disease control." Working paper, Cornell University.

—. 2025b. "Supply chain externalities and agricultural disease." Working paper, Cornell University.

Chambers, R.G., G. Karagiannis, and V. Tzouvelekas. 2010. "Another look at pesticide productivity and pest damage." *American Journal of Agricultural Economics* 92:1401–1419.

Cook, J.A., and C.Y.C. Lin Lawell. 2020. "Wind Turbine Shutdowns and Upgrades in Denmark: Timing Decisions and the Impact of Government Policy." *Energy Journal* 41:81–118.

Cycoń, M., Z. Piotrowska-Seget, A. Kaczyńska, and J. Kozdrój. 2006. "Microbiological characteristics of a sandy loam soil exposed to tebuconazole and λ-cyhalothrin under laboratory conditions." *Ecotoxicology* 15:639–646.

Dash, N.P., A. Kumar, M.S. Kaushik, G. Abraham, and P.K. Singh. 2017. "Agrochemicals influencing nitrogenase, biomass of N_2-fixing cyanobacteria and yield of rice in wetland cultivation." *Biocatalysis and Agricultural Biotechnology* 9:28–34.

Davidson, R., and J.G. MacKinnon. 2004. *Econometric Theory and Methods*. New York: Oxford University Press.

El-Ghamry, A., J. Xu, C. Huang, and J. Gan. 2002. "Microbial response to bensulfuron-methyl treatment in soil." *Journal of Agricultural and Food Chemistry* 50:136–139.

Espino, L.A., W. Brim-DeForest, M. Leinfelder-Miles, B.A. Linquist, P. Buttner, J. Murdock, D. Stewart, and D.A. Sumner. 2021. "Sample costs to produce rice in the Sacramento Valley - 2021." University of California at Davis Cost and Return Studies. [Online]. Accessed 14 May 2023. URL: `https://coststudyfiles.ucdavis.edu/uploads/pub/2021/11/03/2021ricesacramentovalley.pdf`.

Federal Reserve Bank of Minneapolis. 2024. "Consumer price index, 1913-." [Online]. Accessed 12 Jan 2024. URL: `https://www.minneapolisfed.org/about-us/monetary-policy/inflation-calculator/consumer-price-index-1913`.

Feger, F., N. Pavanini, and D. Radulescu. 2020. "Welfare and Redistribution in Residential Electricity Markets with Solar Power." Working Paper.

Fowlie, M., M. Reguant, and S.P. Ryan. 2016. "Market-Based Emissions Regulation and Industry Dynamics." *Journal of Political Economy* 124:249–302.

Galhano, V., F. Peixoto, J. Gomes-Laranjo, and E. Fernandez-Valiente. 2010. "Comparative toxicity of bentazon and molinate on growth, photosynthetic pigments, photosynthesis, and respiration of the Portuguese ricefield cyanobacterium Nostoc muscorum." *Environmental Toxicology: An International Journal* 25:147–156.

—. 2009. "Differential effects of bentazon and molinate on Anabaena cylindrica, an autochthonous cyanobacterium of Portuguese rice field agro-ecosystems." *Water, Air, and Soil Pollution* 197:211–222.

Gerarden, T. 2023. "Demanding Innovation: The Impact of Consumer Subsidies on Solar Panel Production Costs." *Management Science* 69:7799–7820.

Gigliotti, C., L. Allievi, C. Salardi, F. Ferrari, and A. Farini. 2002. "Microbial response to bensulfuron-methyl treatment in soil." *Journal of Agricultural and Food Chemistry* 50:136–139.

Gillingham, K., F. Iskhakov, A. Munk-Nielsen, J. Rust, and B. Schjerning. 2022. "Equilibrium Trade in Automobiles." *Journal of Political Economy* 130:2534–2593.

Huang, L., and M.D. Smith. 2014. "The Dynamic Efficiency Costs of Common-Pool Resource Exploitation." *American Economic Review* 104:4071–4103.

Hussain, S., T. Siddique, M. Saleem, M. Arshad, and A. Khalid. 2009. "Chapter 5: Impact of Pesticides on Soil Microbial Diversity, Enzymes, and Biochemical Reactions." *Advances in Agronomy* 102:159–200.

Jacquet, F., J.P. Butault, and L. Guichard. 2011. "An economic analysis of the possibility of reducing pesticides in French field crops." *Ecological Economics* 70:1638–1648.

Jaenicke, E., and L. Lengnick. 1999. "A Soil-Quality Index and its Relationship to Efficiency and Productivity Growth Measures: Two Decompositions." *American Journal of Agricultural Economics* 81:881–893.

Kalam, S., A. Basu, I. Ahmad, R.Z. Sayyed, H.A. El-Enshasy, D.J. Dailin, and N.L. Suriani. 2020. "Recent Understanding of Soil Acidobacteria and Their Ecological Significance: A Critical Review." *Frontiers in Microbiology* 11.

Kalia, A., and S. Gosal. 2011. "Effect of pesticide application on soil microorganisms." *Archives of Agronomy and Soil Science* 57:569–596.

Kumar, U., S. Behera, S. Saha, D. Das, P. Guru, M. Kaviraj, S. Munda, T. Adak, and A. Nayak. 2020. "Non-target effect of bispyribac sodium on soil microbial community in paddy soil." *Ecotoxicology and Environmental Safety* 189:110019.

Langer, A., and D. Lemoine. 2022. "Designing Dynamic Subsidies to Spur Adoption of New Technologies." *Journal of the Association of Environmental and Resource Economists* 9:1197–1234.

Li, S., Y. Liu, and C. Wei. 2022. "The Cost of Greening Stimulus: A Dynamic Analysis of Vehicle Scrappage Programs." *International Economic Review* 63:1561–1594.

Li, X., B. Li, L. Chen, J. Liang, R. Huang, X. Tang, X. Zhang, and C. Wang. 2022. "Partial substitution of chemical fertilizer with organic fertilizer over seven years increases yields and restores soil bacterial community diversity in wheat-rice rotation." *European Journal of Agronomy* 133:126445.

Lin, C.Y.C. 2013. "Strategic Decision-Making with Information and Extraction Externalities: A Structural Model of the Multi-Stage Investment Timing Game in Offshore Petroleum Production." *Review of Economics and Statistics* 95:1601–1621.

Lo, C.C. 2010. "Effect of pesticides on soil microbial community." *Journal of Environmental Science and Health Part B* 45:348–359.

Lori, M., S. Symnaczik, P. Mäder, G. De Deyn, and A. Gattinger. 2017. "Organic farming enhances soil microbial abundance and activity-A meta-analysis and meta-regression." *PLOS ONE* 12:e0180442.

Magnac, T., and D. Thesmar. 2002. "Identifying Dynamic Discrete Decision Processes." *Econometrica* 70:801–816.

Meneses, M.A., C.L. Casteel, M.I. Gómez, D.R. Just, R. Kanbur, D.R. Lee, and C.Y.C. Lin Lawell. 2025. "Organic Farming Transitions: A Dynamic Bioeconomic Model." Working paper, Cornell University.

Meneses, M.A., M.I. Gómez, D.R. Just, R. Kanbur, D.R. Lee, and C.Y.C. Lin Lawell. forthcoming. "How to Address Barriers to the Adoption of Agricultural Carbon Sequestration." In A. Dinar and R. Mendelsohn, eds. *Handbook on Climate Change Impacts, Mitigation, and Adaptation in Agriculture*. Cheltenham, United Kingdom: Edward Elgar Publishing.

Miller-Klugesherz, J.A., and M.R. Sanderson. 2023. "Good for the soil, but good for the farmer? Addiction and recovery in transitions to regenerative agriculture." *Journal of Rural Studies* 103:103123.

Mujumdar, S., J. Bhoyar, A. Akkar, S. Hundekar, N. Agnihotri, P. Jaybhay, and S. Bhuyan. 2023. "Acinetobacter: A versatile plant growth-promoting rhizobacteria (PGPR)." In P. Swapnil, M. Meena, Harish, A. Marwal, S. Vijayalakshmi, and A. Zehra, eds. *Plant-Microbe Interaction - Recent Advances in Molecular and Biochemical Approaches*. Academic Press, pp. 327–362.

Murphy, D.M., D. Roobroeck, D.R. Lee, and J. Thies. 2020. "Underground Knowledge: Estimating the Impacts of Soil Information Transfers Through Experimental Auctions." *American Journal of Agricultural Economics* 102:1468–1493.

Newman, M.M., N. Hoilett, N. Lorenz, R. Dick, M. Liles, C. Ramsier, and J. Kloepper. 2016. "Glyphosate effects on soil rhizosphere-associated bacterial communities." *Science of the Total Environment* 543:155–160.

Oanh, N., and H. Duc. 2021. "Anaerobic degradation of propanil in soil and sediment using mixed bacterial culture." *Current Microbiology* 78:1499–1508.

Olanrewaju, O., and O. Babalola. 2021. "Streptomyces: implications and interactions in plant growth promotion." *Applied Microbiology and Biotechnology* 103:1179–1188.

Oliva, P., B.K. Jack, S. Bell, E. Mettetal, and C. Severen. 2020. "Technology Adoption Under Uncertainty: Take-Up and Subsequent Investment in Zambia." *Review of Economics and Statistics* 102:617–632.

Organic Produce Network. 2022. "Organic Certification Requirements." [Online]. Accessed 19 September 2022. URL: `https://www.organicproducenetwork.com/article-education/6/organic-certification-requirements-and-the-usda-organic-standards`.

Ploll, U., M. Arato, J. Börner, and M. Hartmann. 2022. "Sustainable Innovations: A Qualitative Study on Farmers' Perceptions Driving the Diffusion of Beneficial Soil Microbes in Germany and the UK." *Sustainability* 14.

Reeling, C., V. Verdier, and F. Lupi. 2020. "Valuing goods allocated via dynamic lottery." *Journal of the Association of Environmental and Resource Economists* 7:721–749.

Righini, H., O. Francioso, A. Martel Quintana, and R. Roberti. 2022. "Cyanobacteria: A Natural Source for Controlling Agricultural Plant Diseases Caused by Fungi and Oomycetes and Improving Plant Growth." *Horticulturae* 8.

Rust, J. 1988. "Maximum Likelihood Estimation of Discrete Control Processes." *SIAM Journal on Control and Optimization* 26:1006–1024.

—. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica: Journal of the Econometric Society* 55:999–1033.

—. 1994. "Structural Estimation of Markov Decision Processes." In R. Engle and D. McFadden, eds. *Handbook of Econometrics*. Amsterdam: North-Holland, vol. 4, pp. 3081–3143.

Ryan, S.P. 2012. "The Costs of Environmental Regulation in a Concentrated Industry." *Econometrica* 80:1019–1061.

Saeki, M., and K. Toyota. 2004. "Effect of bensulfuron-methyl (a sulfonylurea herbicide) on the soil bacterial community of a paddy soil microcosm." *Biology and Fertility of Soils* 40:110–118.

Saison, C., N. Waller, A. Kumar, and R. Kookana. 2009. "Effects of thiobencarb in combinations with molinate and chlorpyrifos on selected soil microbial processes." *Environmental Science and Health Part B* 44:226–234.

Sambucci, O., C.Y.C. Lin Lawell, and T.J. Lybbert. 2025. "Pesticide Spraying and Disease Forecasts: A Dynamic Structural Econometric Model of Grape Growers in California." Working paper, Cornell University.

Scott, P.T. 2013. "Dynamic Discrete Choice Estimation of Agricultural Land Use." Working paper, New York University.

Sears, L., D. Lim, and C.Y.C. Lin Lawell. 2019. "Spatial groundwater management: A dynamic game framework and application to California." *Water Economics and Policy* 5:1850019.

Sears, L., C.Y.C. Lin Lawell, G. Torres, and M.T. Walter. 2025a. "Adjudicated Groundwater Property Rights: A Structural Model of the Dynamic Game Among Groundwater Users in California." Working paper, Cornell University.

—. 2025b. "Moment-based Markov Equilibrium Estimation of High-Dimension Dynamic Games: An Application to Groundwater Management in California." Working paper, Cornell University.

Sears, L., C.Y.C. Lin Lawell, and M.T. Walter. 2025. "Groundwater under open access: A structural model of the dynamic common pool extraction game." Working paper, Cornell University.

Sexton, S.E., Z. Lei, and D. Zilberman. 2007. "The economics of pesticides and pest control." *International Review of Environmental and Resource Economics* 1:271–326.

Singh, J.S., A. Kumar, A.N. Rai, and D.P. Singh. 2016. "Cyanobacteria: A Precious Bio-resource in Agriculture, Ecosystem, and Environmental Sustainability." *Frontiers in Microbiology* 7.

Smith, A. 2023. "California Rice is Back." Ag Data News, 23 Aug 2023. URL: `https://agdatanews.substack.com/p/california-rice-is-back`.

Taboga, M. 2021. "Likelihood ratio test." Lectures on Probability Theory and Mathematical Statistics. Kindle Direct Publishing. Online appendix. URL: `https://www.statlect.com/fundamentals-of-statistics/likelihood-ratio-test`.

Thiebaut, F., M. Urquiaga, A. Rosman, M. da Silva, and A. Hemerly. 2022. "The Impact of Non-Nodulating Diazotrophic Bacteria in Agriculture: Understanding the Molecular Mechanisms That Benefit Crops." *International Journal of Molecular Sciences* 23:11301.

Timmins, C. 2002. "Measuring the Dynamic Efficiency Costs of Regulators' Preferences: Municipal Water Utilities in the Arid West." *Econometrica* 70:603–629.

Toyama, Y. 2024. "Dynamic Incentives and Permit Market Equilibrium in Cap-and-Trade Regulation." *American Economic Journal: Microeconomics* 16:374â423.

UC Rice Research and Information Center. 2023. "Rice." University of California at Davis. [Online]. Accessed 14 May 2023. URL: `https://agronomy-rice.ucdavis.edu/`.

U.S. Department of Agriculture [USDA]. 2023. "Certified organic survey." National Agricultural Statistics Service (NASS). [Online]. Accessed 14 May 2023. URL: `https://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Organic_Production/`.

—. 2024. "Data and Statistics." National Agricultural Statistics Service (NASS). [Online]. Accessed 3 June 2024. URL: `https://www.nass.usda.gov/Data_and_Statistics/index.php`.

—. 2007. "Organic Production Survey." National Agricultural Statistics Service (NASS). [Online]. Accessed 14 May 2023. URL: https://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Organic_Production/.

—. 2012. "Organic Production Survey." National Agricultural Statistics Service (NASS). [Online]. Accessed 14 May 2023. URL: https://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Organic_Production/.

—. 2017. "Organic Production Survey." National Agricultural Statistics Service (NASS). [Online]. Accessed 14 May 2023. URL: https://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Organic_Production/.

—. 2025. "Quick Stats." National Agricultural Statistics Service (NASS). [Online]. Accessed 16 May 2025. URL: https://quickstats.nass.usda.gov/.

—. 1992. "Weights, Measures, and Conversion Factors for Agricultural Commodities and Their Products." Economic Research Service in cooperation with the Agricultural Marketing Service, the Agricultural Research Service, and the National Agricultural Statistics Service, U.S. Department of Agriculture. Agricultural Handbook No. 697. URL: https://www.ers.usda.gov/webdocs/publications/41880/33132_ah697_002.pdf?v=0.

USA Rice. 2020. "Where Rice Grows: California." URL: https://www.usarice.com/thinkrice/discover-us-rice/where-rice-grows/state/california.

USDA Agricultural Marketing Service. 2000a. "National Organic Program." 7 CFR 205. 21 December 2000.A Rule by the Agriculture Department, and the Agricultural Marketing Service. URL: https://www.ecfr.gov/current/title-7/subtitle-B/chapter-I/subchapter-M/part-205.

—. 2000b. "National Organic Program: A Rule by the Agricultural Marketing Service on 12/21/2000." 65 FR 80637. 21 December 2000. URL: https://www.federalregister.gov/documents/2000/12/21/00-32257/national-organic-program.

van Kooten, G.C., W.P. Weisensel, and D. Chinthammit. 1990. "Valuing Tradeoffs between Net Returns and Stewardship Practices: The Case of Soil Conservation in Saskatchewan." *American Journal of Agricultural Economics* 72:104–113.

Verma, R., H. Annapragada, N. Katiyar, N. Shrutika, K. Das, and S. Murugesan. 2020. "Chapter 4 - Rhizobium." In N. Amaresan, M. Senthil Kumar, K. Annapurna, K. Kumar, and A. Sankaranarayanan, eds. *Beneficial Microbes in Agro-Ecology*. Academic Press, pp. 37–54.

Weber, P. 2022. "Dynamic responses to carbon pricing in the electricity sector." Working paper, University of California at Berkeley.

Wu, T., D.R. Just, C.Y.C. Lin Lawell, A. Ortiz-Bobea, and J. Zhao. 2025. "Optimal forest management for interdependent products: A nested stochastic dynamic bioeconomic model and application to bamboo." Working paper, Cornell University.

Xie, M.L.C.H.W.L., X.M., and S. Abid. 2004. "Effects of pesticides on soil biochemical characteristics of a paddy soil." *Journal of Environmental Sciences* 16:252–255.

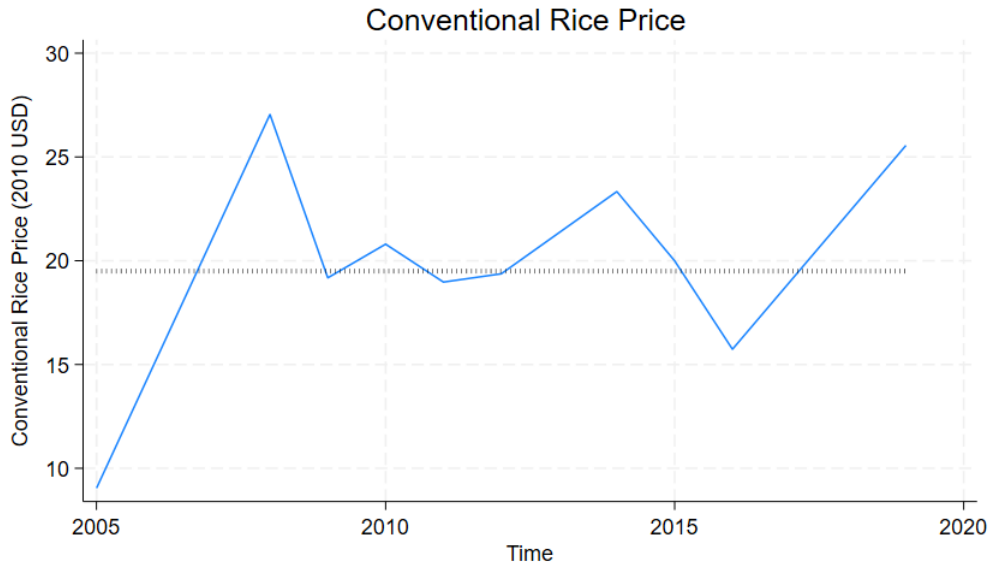
Xie, X.a.M.L., C. Huang, and W. Liu. 2004. "Effects of bensulfuron-methyl on soil microbial activity and biochemical characteristics in paddy." *Chinese Journal of Rice Science* 18:67.

Yadav, S., S. Rai, R. Rai, A. Shankar, S. Singh, and L.C. Rai. 2017. *Cyanobacteria: Role in Agriculture, Environmental Sustainability, Biotechnological Potential and Agroecological Impact*, Singapore: Springer Singapore. pp. 257–277.

Yeh, D.A., M.I. Gómez, and C.Y.C. Lin Lawell. 2025. "Sustainable pest management, beliefs, and behavior: A dynamic bioeconomic analysis." Working paper, Cornell University.

Yibeltie, G., and S. Sahile. 2018. "The Role of Cyanobacteria on Agriculture." *Journal of Natural Sciences Research* 8.

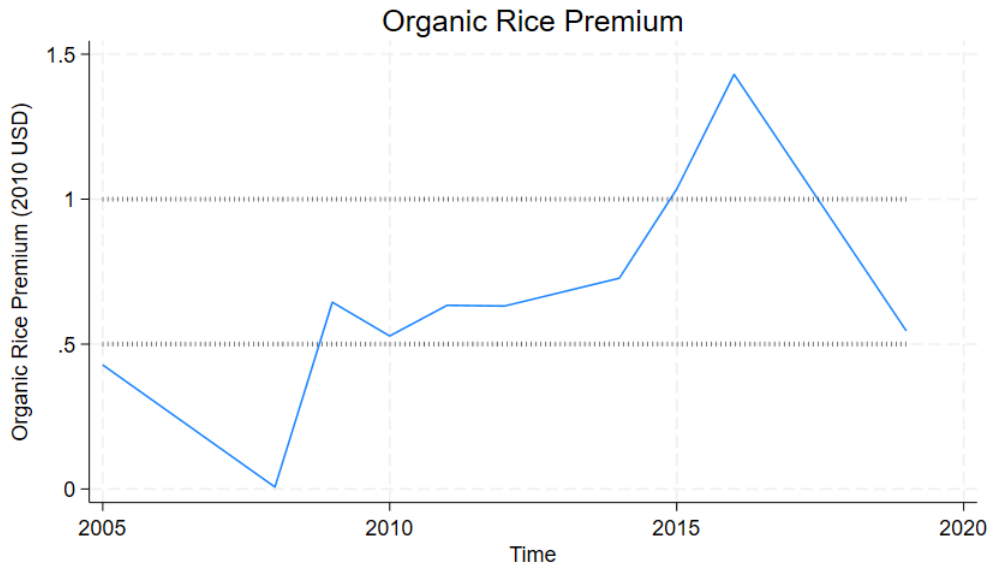
Figure 1: Real Conventional Rice Price



Notes: Price is in 2010 USD per CWT. Horizontal dashed grey line indicates cutoff used to discretize real conventional rice price.

Data Sources: USDA NASS Organic Production Survey (U.S. Department of Agriculture [USDA], 2007, 2012, 2017), USDA NASS Certified Organic Survey (U.S. Department of Agriculture [USDA], 2023), University of California Rice Research and Information Center (UC Rice Research and Information Center, 2023), and University of California at Davis Cost and Return Studies (Espino et al., 2021). We convert prices to real prices in 2010 USD using the consumer price index from the Federal Reserve Bank of Minneapolis (Federal Reserve Bank of Minneapolis, 2024).

Figure 2: Real Organic Rice Price Premium



Notes: The real organic price premium is given by $\frac{P_{org}-P_{con}}{P_{org}}$, where P_{con} is the real conventional price and P_{org} is the real organic price. Price premium is in 2010 USD per CWT. Horizontal dashed grey lines indicate cutoffs used to discretize real organic rice price premium.

Data Sources: USDA NASS Organic Production Survey (U.S. Department of Agriculture [USDA], 2007, 2012, 2017), USDA NASS Certified Organic Survey (U.S. Department of Agriculture [USDA], 2023), University of California Rice Research and Information Center (UC Rice Research and Information Center, 2023), and University of California at Davis Cost and Return Studies (Espino et al., 2021). We convert prices to real prices in 2010 USD using the consumer price index from the Federal Reserve Bank of Minneapolis (Federal Reserve Bank of Minneapolis, 2024).

Table 1: Distribution of Action a_{it}

Action a_{it}		**Frequency**	**Percent**
$a_{it} = \text{RC}$	Plant rice and use pesticide that year	61,890	92.06
$a_{it} = \text{RN}$	Plant rice and do not use pesticide that year	1,355	2.02
$a_{it} = \text{OC}$	Do not plant rice that year and use pesticide that year	3,931	5.85
$a_{it} = \text{ON}$	Do not plant rice and do not use pesticide that year	54	0.08

Notes: Each observation is a farmer-field-year. There are 67,230 farmer-field-year observations. There are 17,695 farmer-fields in the data set, comprising 2,101 farmers and spanning 15 counties. By 'pesticide', we mean 'unapproved synthetic pesticide'.

Data Source: CA DPR PUR (California Department of Pesticide Regulation [CA DPR], 2024).

Table 2: Distribution of Clean Soil Stock k_{it}

Clean Soil Stock k_{it}	**Frequency**	**Percent**
0	55,973	83.26
1	7,452	11.08
2	1,508	2.24
3	912	1.36
4	317	0.47
5	204	0.30
6	152	0.23
7	125	0.19
8	97	0.14
9	80	0.12
10	67	0.10
11	62	0.09
12	63	0.09
13	48	0.07
14	42	0.06
15	34	0.05
16	23	0.03
17	16	0.02
18	19	0.02
19	11	0.02
20	9	0.01
21	5	0.01
22	5	0.01
23	4	0.01
24	1	0.00
25	0	0.00
26	1	0.00

Notes: Clean soil stock k_{it} is measured by the number of previous consecutive years the farmer has not used any unapproved synthetic pesticide. Each observation is a farmer-field-year. There are 67,230 farmer-field-year observations. There are 17,695 farmer-fields in the data set, comprising 2,101 farmers and spanning 15 counties.

Data Source: CA DPR PUR (California Department of Pesticide Regulation [CA DPR], 2024).

Table 3: Summary Statistics for Production Function Estimation, 1990-2019

	Mean	**Std. Dev.**	**Min**	**Max**	**# Obs**
Number of farmer-fields growing rice in county-year	553.50	591.91	1	2,522	325
Rice yield (CWT) per farmer-field growing rice in county-year	6,731	3,661	1,145	53,000	325
Fraction of farmer-fields growing rice in county-year who applied pesticide that year	0.9580	0.1269	0	1	325
Average over farmer-fields growing rice in county-year of # previous consecutive years no pesticide was used	0.3116	0.9024	0	10.2500	325

Notes: Each observation is a county-year. We use data from 1990-2019. There are 15 counties with data for at least one year over 1990-2019. In North America, 1 hundredweight equals 100 pounds (U.S. Department of Agriculture [USDA], 1992). By 'pesticide', we mean 'unapproved synthetic pesticide'.

Data Sources: CA DPR PUR (California Department of Pesticide Regulation [CA DPR], 2024), USDA NASS Quick Stats (U.S. Department of Agriculture [USDA], 2024).

Table 4: True 'Full Information' Production Function

		Dependent variable is: Log of rice yield (CWT)	
Parameter	Regressor	(1)	(2)
α_c	Use pesticide	3.757^{***}	4.032^{***}
		(0.503)	(0.444)
α_k	# Previous consecutive years no pesticide was used	0.400^{***}	0.367^{***}
		(0.0675)	(0.0541)
α_0	Constant	4.983^{***}	4.280^{***}
		(0.496)	(0.469)
County Fixed Effects		N	Y
# Observations		325	325
# Counties		15	15
R-squared		0.185	0.521

Notes: Standard errors in parentheses. Each observation is a county-year. We use data from 1990-2019. Inverse variance weights that weight county-years with more farmer-fields growing rice more heavily are used. By 'pesticide', we mean 'unapproved synthetic pesticide'. Since yield data is only available at the county level, not the farmer-field level, we estimate the production function using county averages, where each observation is a county-year in which data on rice yield per farmer-yield growing rice data is available. In particular, we regress log rice yield per farmer-field growing rice in a county-year on the fraction of farmer-fields growing rice in a county-year who applied an unapproved synthetic pesticide that year and the average over farmer-fields growing rice in a county-year of the number of previous consecutive years no unapproved synthetic pesticide was used. 'Full Information' specifications account for soil microbes, and therefore include as a regressor the average over farmer-fields growing rice in a county-year of the number of previous consecutive years no unapproved synthetic pesticide was used. In North America, 1 hundredweight equals 100 pounds (U.S. Department of Agriculture [USDA], 1992). Significance codes: *** $p<0.001$, ** $p<0.01$, * $p<0.05$

Table 5: 'Misperception' Production Function

Dependent variable is: Log of rice yield (CWT)			
Parameter	Regressor	(1)	(2)
α_c	Use pesticide	3.012^{***}	3.473^{***}
		(0.513)	(0.467)
α_0	Constant	5.789^{***}	4.885^{***}
		(0.502)	(0.493)
County Fixed Effects		N	Y
# Observations		325	325
# Counties		15	15
R-squared		0.097	0.449

Notes: Standard errors in parentheses. Each observation is a county-year. We use data from 1990-2019. Inverse variance weights that weight county-years with more farmer-fields growing rice more heavily are used. By 'pesticide', we mean 'unapproved synthetic pesticide'. Since yield data is only available at the county level, not the farmer-field level, we estimate the production function using county averages, where each observation is a county-year in which data on rice yield per farmer-yield growing rice data is available. In particular, for the 'Misperception' specifications, which do not account for soil microbes, we regress log rice yield per farmer-field growing rice in a county-year on the fraction of farmer-fields growing rice in a county-year who applied an unapproved synthetic pesticide that year. In North America, 1 hundredweight equals 100 pounds (U.S. Department of Agriculture [USDA], 1992). Significance codes: *** p<0.001, ** p<0.01, * p<0.05

Table 6: Dynamic Structural Parameter Estimates

		(1)	(2)	(3)
Parameters in the perceived yield $\breve{q}(\cdot)$				
γ_c	Misperception of how log yield is affected by pesticide use c_{it}	-0.839***	-0.836***	-0.836***
		(0.0118)	(0.0139)	(0.0136)
γ_k	Misperception of how log yield is affected by clean soil stock k_{it}	-0.285***	-0.285***	-0.285***
		(0.0042)	(0.0043)	(0.0043)
Coefficients in rice production cost on:				
κ_c	Use pesticide c_{it}	-0.025***	-0.042***	-0.048***
		(0.0007)	(0.0012)	(0.0015)
κ_1	Perceived yield $\breve{q}_{it}$	0.070***	0.065***	0.063***
		(0.0024)	(0.0024)	(0.0026)
κ_{cq}	Perceived yield $\breve{q}_{it}$ X Use pesticide c_{it}		0.069***	0.067***
			(0.0022)	(0.0023)
κ_2	Perceived yield squared $\breve{q}_{it}^2$	0.0863***	0.0846***	0.08543***
		(0.0011)	(0.0012)	(0.0014)
κ_{cq2}	Perceived yield squared $\breve{q}_{it}^2$ X Use pesticide c_{it}			0.08535***
				(0.0014)
Other parameters				
ν	Deterministic payoff when farmer does not plant rice and does not use pesticide	-10.00***	-10.00***	-10.00***
		(0.00000)	(0.00000)	(0.00001)
Log likelihood		-7,184	-7,185	-7,186
# Observations		67,230	67,230	67,230
# Farmer-fields		17,695	17,695	17,695
# Counties		15	15	15

Notes: Standard errors in parentheses. Each observation is a farmer-field-year. Clean soil stock k_{it} is measured by the number of previous consecutive years (up to 15 years) no unapproved synthetic pesticide was used. Yield is in units of 1 million pounds. Payoffs and costs are in units of ten thousand 2010 USD. Significance codes: *** p<0.001, ** p<0.01, * p<0.05

Table 7: Counterfactual Simulations: Short Run

		Misperception			**Full Information**		
		(1)	(2)	(3)	(1)	(2)	(3)
Distribution of a_{it} (% of farmer-field-years)							
a_{it} = RC	Plant rice and use pesticide that year	97.7	97.8	97.8	5.7	6.2	4.7
a_{it} = RN	Plant rice and do not use pesticide that year	1.6	1.6	1.6	94.3	93.8	95.3
a_{it} = OC	Do not plant rice that year and use pesticide that year	0.7	0.7	0.7	0.0	0.0	0.0
a_{it} = ON	Do not plant rice and do not use pesticide that year	0.0	0.0	0.0	0.0	0.0	0.0
Organic production (% of farmer-field-years)		0.06	0.06	0.06	56.12	54.57	58.15
Clean soil stock k_{it} (mean over farmer-field-years)		0.05	0.05	0.05	4.15	3.85	4.36
PDV of entire stream of per-period payoffs							
Mean over farmer-fields (million 2010 USD)		0.73	0.73	0.51	1.43	1.28	1.36
(Std. Dev. over farmer-fields in parentheses (million 2010 USD))		(0.89)	(0.89)	(3.19)	(2.32)	(1.72)	(2.34)
5th %-ile over farmer-fields (million 2010 USD)		0.17	0.17	0.14	0.01	0.01	0.01
25th %-ile over farmer-fields (million 2010 USD)		0.52	0.52	0.52	0.07	0.07	0.08
50th %-ile over farmer-fields (million 2010 USD)		0.72	0.72	0.71	0.19	0.19	0.19
75th %-ile over farmer-fields (million 2010 USD)		0.86	0.86	0.85	2.07	2.07	1.91
95th %-ile over farmer-fields (million 2010 USD)		1.00	0.99	0.99	3.29	3.22	3.50
Total PDV of entire stream of per-period payoffs (billion 2010 USD)		18.4	18.5	12.8	36.3	32.5	34.4

Notes: Table presents averages over 100 simulations. The 'short run' simulations simulate each farmer-field from the actual value of the state (clean soil stock k_{it} and prices) in the first year for which we have data for that farmer-field, to the final year of our data set, year 2019. The 'Misperception' scenario is a base simulation of actual behavior, and uses the parameter estimates from our dynamic structural econometric model in Table 6 for all the parameters; Specifications (1), (2), and (3) use parameter estimates from Specifications (1), (2), and (3), respectively, of Table 6. The 'Full Information' scenario is a counterfactual simulation of optimal behavior, in which we set the misperception parameters γ_k and γ_c both to 0, and then use the parameter estimates from our dynamic structural econometric model in Table 6 for all the remaining parameters; Specifications (1), (2), and (3) use parameter estimates from Specifications (1), (2), and (3), respectively, of Table 6. The PDV of entire stream of per-period payoffs is relative to the per-period payoff from the outside option of not planting rice and using pesticide that year (a_{it} = OC), which is normalized to 0.

Table 8: Counterfactual Simulations: Long Run

		Misperception			**Full Information**		
		(1)	(2)	(3)	(1)	(2)	(3)
Distribution of a_{it} (% of farmer-field-years)							
a_{it} = RC	Plant rice and use pesticide that year	97.8	97.8	97.8	8.2	8.6	6.8
a_{it} = RN	Plant rice and do not use pesticide that year	1.6	1.6	1.6	91.8	91.4	93.2
a_{it} = OC	Do not plant rice that year and use pesticide that year	0.7	0.7	0.6	0.0	0.0	0.0
a_{it} = ON	Do not plant rice and do not use pesticide that year	0.0	0.0	0.0	0.0	0.0	0.0
Organic production (% of farmer-field-years)		0.04	0.04	0.04	59.19	57.57	62.02
Clean soil stock k_{it} (mean over farmer-field-years)		0.04	0.04	0.04	4.77	4.40	5.10
PDV of entire stream of per-period payoffs							
Mean over farmer-fields (million 2010 USD)		0.92	0.93	0.70	2.68	2.45	2.47
(Std. Dev. over farmer-fields in parentheses (million 2010 USD))		(0.88)	(0.88)	(3.19)	(2.42)	(1.45)	(2.58)
5th %-ile over farmer-fields (million 2010 USD)		0.67	0.67	0. 67	1.78	1.77	0.29
25th %-ile over farmer-fields (million 2010 USD)		0.75	0.75	0.75	1.88	1.88	1.87
50th %-ile over farmer-fields (million 2010 USD)		0.86	0.86	0.85	2.08	2.08	1.93
75th %-ile over farmer-fields (million 2010 USD)		1.00	1.00	0.99	2.66	2.66	2.53
95th %-ile over farmer-fields (million 2010 USD)		1.15	1.14	1.14	4.04	3.84	4.41
Total PDV of entire stream of per-period payoffs (billion 2010 USD)		23.4	23.5	17.8	67.7	62.0	62.7

Notes: Table presents averages over 100 simulations. The 'long run' simulations simulate each farmer-field from the actual value of the state (clean soil stock k_{it} and prices) in the first year for which we have data for that farmer-field, to 10 years past the the final year of our data set (i.e., to year 2029). The 'Misperception' scenario is a base simulation of actual behavior, and uses the parameter estimates our dynamic structural econometric model in Table 6 for all the parameters; Specifications (1), (2), and (3) use parameter estimates from Specifications (1), (2), and (3), respectively, of Table 6. The 'Full Information' scenario is a counterfactual simulation of optimal behavior, in which we set the misperception parameters γ_k and γ_c both to 0, and then use the parameter estimates from our dynamic structural econometric model in Table 6 for all the remaining parameters; Specifications (1), (2), and (3) use parameter estimates from Specifications (1), (2), and (3), respectively, of Table 6. The PDV of entire stream of per-period payoffs is relative to the per-period payoff from the outside option of not planting rice and using pesticide that year (a_{it} = OC), which is normalized to 0.

Appendix

A Production Function Robustness Checks

A.1 Linear Production Function

For robustness, we also estimate a linear specification of our production function wherein the dependent variable is yield rather than log yield. The linear specification of our production function $q(\cdot)$ for rice yield y_{it} is given by:

$$y_{it} = q(c_{it}, k_{it}; \alpha) = \alpha_0 + \alpha_c c_{it} + \alpha_k k_{it}, \tag{A.1}$$

where y_{it} is the rice yield in hundredweights (CWT); c_{it} is a dummy variable for farmer-field i using an unapproved synthetic pesticide in year t; and k_{it} is the clean soil stock as measured by the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide. In particular, the clean soil stock k_{it} is equal to 0 if the rice farmer-field used an unapproved synthetic pesticide the previous year, 1 if the rice farmer-field did not use an unapproved synthetic pesticide the previous year but used an unapproved synthetic pesticide two years ago, 2 if the rice farmer-field did not use an unapproved synthetic pesticide the previous two years but used an unapproved synthetic pesticide three years ago, and 3 if the rice farmer-field did not use an unapproved synthetic pesticide for the previous three years but used an unapproved synthetic pesticide four years ago, and so on.

Since yield data is only available at the county level, not the farmer-field level, we estimate the production function using county averages, where each observation is a county-year in which data on rice yield per farmer-field growing rice data is available. In particular, we regress rice yield per farmer-field growing rice in a county-year on the fraction of farmer-fields growing rice in a county-year who applied an unapproved synthetic pesticide that year and the average over farmer-fields growing rice in a county-year of the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide. Because the variables in the regression reflect means rather than individual observations, the appropriate method of estimation is analytically weighted least squares (Davidson and MacKinnon, 2004), where the weight is the number of farmer-fields growing rice in the county-year. We therefore use inverse variance weights that weight county-years with more farmer-fields growing rice more heavily.

We estimate two types of specifications of our production function. 'Misperception' specifications do not account for soil microbes, and therefore do not include any regressor relating to the clean soil stock k_{it}. 'Full Information' specifications account for soil microbes, and therefore include a regressor relating the clean soil stock k_{it}, the average over farmer-fields growing rice in a county-year of the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide.

Results for our true 'Full Information' production function estimation are in Table A.1. As expected, α_c is significant and positive in the 'Full Information' specifications, which means that, in actuality, when accounting for soil microbes, the contemporaneous use of an unapproved synthetic

pesticide has a positive effect on yield that year. In addition, consistent with recent insights from soil science that show the important of soil microbes and a clean soil stock, α_k is significant and positive in the 'Full Information' specifications, which means that yields are higher the more previous consecutive years the farmer-field has not used any unapproved synthetic pesticide.

Results for our 'Misperception' production function estimation are in Table A.2. As expected, in the 'Misperception' specifications, α_c is significant and positive, which means that farmers who do not account for soil microbes perceive that using an unapproved synthetic pesticide that year has a positive effect on yield.

Table A.1: Robustness: True 'Full Information' Production Function, Linear Specification

		Dependent variable is: Rice yield (CWT)	
Parameter	Regressor	(1)	(2)
α_c	Use pesticide	17,257***	18,513***
		(3,144)	(2,684)
α_k	# Previous consecutive years no pesticide was used	2,026***	1,805***
		(421.6)	(327.2)
α_0	Constant	-10,779***	-14,043***
		(3,098)	(0.469)
County Fixed Effects		N	Y
# Observations		325	325
# Counties		15	15
R-squared		0.117	0.513

Notes: Standard errors in parentheses. Each observation is a county-year. We use data from 1990-2019. Inverse variance weights that weight county-years with more farmer-fields growing rice more heavily are used. By 'pesticide', we mean 'unapproved synthetic pesticide'. Since yield data is only available at the county level, not the farmer-field level, we estimate the production function using county averages, where each observation is a county-year in which data on rice yield per farmer-yield growing rice data is available. In particular, we regress rice yield per farmer-field growing rice in a county-year on the fraction of farmer-fields growing rice in a county-year who applied an unapproved synthetic pesticide that year and the average over farmer-fields growing rice in a county-year of the number of previous consecutive years no unapproved synthetic pesticide was used. 'Full Information' specifications account for soil microbes, and therefore include as a regressor the average over farmer-fields growing rice in a county-year of the number of previous consecutive years no unapproved synthetic pesticide was used. In North America, 1 hundredweight equals 100 pounds (U.S. Department of Agriculture [USDA], 1992). Significance codes: *** $p<0.001$, ** $p<0.01$, * $p<0.05$

Table A.2: Robustness: 'Misperception' Production Function, Linear Specification

Dependent variable is: Rice yield (CWT)			
Parameter	Regressor	(1)	(2)
α_c	Use pesticide	13,484***	15,761***
		(3,147)	(2,760)
α_0	Constant	-6,694**	-11,062***
		(3,080)	(2,911)
County Fixed Effects		N	Y
# Observations		325	325
# Counties		15	15
R-squared		0.054	0.465

Notes: Standard errors in parentheses. Each observation is a county-year. We use data from 1990-2019. Inverse variance weights that weight county-years with more farmer-fields growing rice more heavily are used. By 'pesticide', we mean 'unapproved synthetic pesticide'. Since yield data is only available at the county level, not the farmer-field level, we estimate the production function using county averages, where each observation is a county-year in which data on rice yield per farmer-yield growing rice data is available. In particular, for the 'Misperception' specifications, which do not account for soil microbes, we regress rice yield per farmer-field growing rice in a county-year on the fraction of farmer-fields growing rice in a county-year who applied an unapproved synthetic pesticide that year. In North America, 1 hundredweight equals 100 pounds (U.S. Department of Agriculture [USDA], 1992). Significance codes: *** p<0.001, ** p<0.01, * p<0.05

A.2 Acreage

Summary statistics for reported acres and reported acres growing rice for the farmer-field-years in our data set are presented in Table A.3.

The measure of acres planted reported in the PUR data may not be entirely reliable since farmers sometimes report actual acres planted but other times instead report the entire area of the larger plot of land in which the planting occurred (California Department of Pesticide Regulation [CA DPR], 2023).

Nevertheless, for robustness, we also run a robustness check in which we estimate our production function using county averages calculated by dividing by acres rather than by farmer-field. In particular, we regress rice yield per acre growing rice in a county-year (calculated as the rice yield in that county-year, divided by total number of acres of all farmer-fields in that county growing rice in that year) on the average over acres growing rice of the dummy for applying an unapproved synthetic pesticide that year (calculated as the number of all farmer-fields in that county both growing rice and applying an unapproved synthetic pesticide that year, divided by total number of acres of all farmer-fields in that county growing rice in that year) and the average over acres growing rice in a county-year of the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide (calculated as sum over farmer-fields growing rice in that county-year of the number of previous consecutive years the farmer-field has not used any unapproved synthetic pesticide, divided by total number of acres of all farmer-fields in that county growing rice in that year). Because the variables in the regression reflect means rather than individual observations, the appropriate method of estimation is analytically weighted least squares (Davidson and MacKinnon, 2004), where the weight is the total number of acres of all farmer-fields in that county growing rice in that year. We therefore use inverse variance weights that weight county-years with more acres planted to rice more heavily.

We use a conservative version of our county average variables in which we drop county-year observations that include a farmer-field-year for which the number of acres planted to rice in that farmer-field-year in that county in that year is missing. The counties appearing in the conservative estimation are identical to the counties appearing in the standard estimation. The conservative estimation loses disproportionately more observations from Yuba County. Summary statistics for the annual county-level variables we use to estimate our production function are presented in Table A.4.

Results for our true 'Full Information' production function estimation are in Table **??**. As expected, α_c is significant and positive in the 'Full Information' specifications, which means that, in actuality, when accounting for soil microbes, the contemporaneous use of an unapproved synthetic pesticide has a positive effect on yield that year. In addition, consistent with recent insights from soil science that show the important of soil microbes and a clean soil stock, α_k is significant and positive in the 'Full Information' specifications, which means that yields are higher the more previous consecutive years the farmer-field has not used any unapproved synthetic pesticide.

Results for our 'Misperception' production function estimation are in Table **??**. As expected, in

the 'Misperception' specifications, α_c is significant and positive, which means that farmers who do not account for soil microbes perceive that using an unapproved synthetic pesticide that year has a positive effect on yield.

Table A.3: Summary Statistics for Acreage

	Mean	**Std. Dev.**	**Min**	**Max**	**# Obs**
Acres	82.05	57.22	1.00	721.25	67,299
Acres growing rice	81.63	56.94	1.00	700.00	63,244

Notes: Each observation is a farmer-field-year.

Data Source: CA DPR PUR (California Department of Pesticide Regulation [CA DPR], 2024).

Table A.4: Summary Statistics for Production Function by Acres, 1990-2019

	Mean	**Std. Dev.**	**Min**	**Max**	**# Obs**
Number of acres growing rice in county-year	44,617	46,391	49.7	174,900	303
Rice yield (CWT) per acre growing rice in county-year	78.61	60.08	13.42	1,066	303
Average over acres growing rice in county-year of applied pesticide that year	0.0115	0.0032	0	0.0240	303
Average over acres growing rice in county-year of # previous consecutive years no pesticide was used	0.0038	0.0102	0	0.1145	303

Notes: Each observation is a county-year. We use data from 1990-2019. There are 15 counties with data for at least one year over 1990-2019. In North America, 1 hundredweight equals 100 pounds (U.S. Department of Agriculture [USDA], 1992). By 'pesticide', we mean 'unapproved synthetic pesticide'.

Data Sources: CA DPR PUR (California Department of Pesticide Regulation [CA DPR], 2024), USDA NASS Quick Stats (U.S. Department of Agriculture [USDA], 2024).

Table A.5: Robustness: True 'Full Information' Production Function by Acre

		Dependent variable is: Rice yield (CWT)	
Parameter	Regressor	(1)	(2)
α_c	Use pesticide	1,552***	3,487***
		(304.2)	(408.4)
α_k	# Previous consecutive years no pesticide was used	675.0**	731.0***
		(264.7)	(219.7)
α_0	Constant	59.93***	12.26
		(3.60)	(9.12)
County Fixed Effects		N	Y
# Observations		303	303
# Counties		15	15
R-squared		0.107	0.452

Notes: Standard errors in parentheses. Each observation is a county-year. We use data from 1990-2019. Inverse variance weights that weight county-years with more acres growing rice more heavily are used. By 'pesticide', we mean 'unapproved synthetic pesticide'. Since yield data is only available at the county level, not the farmer-field level, we estimate the production function using county averages, where each observation is a county-year in which data on rice yield per acre growing rice data is available. In particular, we regress rice yield per acre growing rice in a county-year on the average over acres growing rice in a county-year of the dummy for applying an unapproved synthetic pesticide that year and the average over acres growing rice in a county-year of the number of previous consecutive years no unapproved synthetic pesticide was used. 'Full Information' specifications account for soil microbes, and therefore include as a regressor the average over farmer-fields growing rice in a county-year of the number of previous consecutive years no unapproved synthetic pesticide was used. In North America, 1 hundredweight equals 100 pounds (U.S. Department of Agriculture [USDA], 1992). Significance codes: *** p<0.001, ** p<0.01, * p<0.05

Table A.6: Robustness: 'Misperception' Production Function by Acre

		Dependent variable is: Rice yield (CWT)	
Parameter	Regressor	(1)	(2)
α_c	Use pesticide	1,643***	3,615***
		(304.9)	(413.7)
α_0	Constant	60.60***	11.96
		(3.62)	(9.277)
County Fixed Effects		N	Y
# Observations		303	303
# Counties		15	15
R-squared		0.088	0.431

Notes: Standard errors in parentheses. Each observation is a county-year. We use data from 1990-2019. Inverse variance weights that weight county-years with more acres growing rice more heavily are used. By 'pesticide', we mean 'unapproved synthetic pesticide'. Since yield data is only available at the county level, not the farmer-field level, we estimate the production function using county averages, where each observation is a county-year in which data on rice yield per farmer-yield growing rice data is available. In particular, for the 'Misperception' specifications, which do not account for soil microbes, we regress rice yield per acre growing rice in a county-year on the average over acres growing rice in a county-year of the dummy for applying an unapproved synthetic pesticide that year. In North America, 1 hundredweight equals 100 pounds (U.S. Department of Agriculture [USDA], 1992). Significance codes: *** p<0.001, ** p<0.01, * p<0.05

A.3 Year Effects

For robustness, we also estimate a specification of our true 'Full Information' production function using year effects, and a specification of our true 'Full Information' production function using both year effects and county fixed effects.

Results for our true 'Full Information' production function estimation are in Table A.7. Consistent with recent insights from soil science that show the important of soil microbes and a clean soil stock, α_k is significant and positive in the 'Full Information' specifications, which means that yields are higher the more previous consecutive years the farmer-field has not used any unapproved synthetic pesticide.

While the coefficient α_c on the contemporaneous use of an unapproved synthetic pesticide is positive, it is not statistically significant at a 5% level. Thus, in order for to obtain a lower bound conservative measure of the benefits of forgoing pesticide use for future benefits via clean soil stock, we opt not to use the specifications with year effects for our dynamic structural econometric model.

Table A.7: Robustness: True 'Full Information' Production Function with Year Effects

		Dependent variable is: Log of rice yield (CWT)	
Parameter	Regressor	(1)	(2)
α_c	Use pesticide	0.973	0.079
		(0.607)	(0.443)
α_k	# Previous consecutive years no pesticide was used	0.254^{***}	0.145^{***}
		(0.119)	(0.0825)
α_0	Constant	7.063^{***}	7.498^{***}
		(0.558)	(0.426)
County Fixed Effects		N	Y
Year Effects		Y	Y
# Observations		325	325
# Counties		15	15
R-squared		0.458	0.794

Notes: Standard errors in parentheses. Each observation is a county-year. We use data from 1990-2019. Inverse variance weights that weight county-years with more farmer-fields growing rice more heavily are used. By 'pesticide', we mean 'unapproved synthetic pesticide'. Since yield data is only available at the county level, not the farmer-field level, we estimate the production function using county averages, where each observation is a county-year in which data on rice yield per farmer-yield growing rice data is available. In particular, we regress log rice yield per farmer-field growing rice in a county-year on the fraction of farmer-fields growing rice in a county-year who applied an unapproved synthetic pesticide that year and the average over farmer-fields growing rice in a county-year of the number of previous consecutive years no unapproved synthetic pesticide was used. 'Full Information' specifications account for soil microbes, and therefore include as a regressor the average over farmer-fields growing rice in a county-year of the number of previous consecutive years no unapproved synthetic pesticide was used. In North America, 1 hundredweight equals 100 pounds (U.S. Department of Agriculture [USDA], 1992). Significance codes: *** $p<0.001$, ** $p<0.01$, * $p<0.05$

A.4 Subset of Years Used in Dynamic Structural Econometric Model

For robustness, we also estimate our true 'Full Information' production function using only annual county-level data from the 10 years of data in the annual farmer-field-level panel data set we use for our dynamic structural econometric model (2005, 2008-2012, 2014-2016, and 2019); these 10 years of data are the 10 years for which we have data on conventional and organic rice prices.

Results for our true 'Full Information' production function estimation are in Table A.8. These regressions are likely under-powered, owing to the smaller sample size that we are left with when we restrict the estimation to only include observations from years 2005, 2008-2012, 2014-2016, and 2019. Nevertheless, consistent with recent insights from soil science that show the important of soil microbes and a clean soil stock, α_k is positive in all 4 specifications, and statistically significant in the year effects specification (Specification (3)), which means that yields are higher the more previous consecutive years the farmer-field has not used any unapproved synthetic pesticide.

Table A.8: Robustness: True 'Full Information' Production Function, Subset of Years in Dynamic Structural Econometric Model

		Dependent variable is: Log of rice yield (CWT)			
Parameter	Regressor	(1)	(2)	(3)	(4)
α_c	Use pesticide	0.230	-0.144	1.721	0.561
		(1.154)	(1.355)	(1.393)	(1.356)
α_k	# Previous consecutive years no pesticide was used	0.0919	0.0171	0.394**	0.0661
		(0.0917)	(0.0532)	(0.197)	(0.108)
α_0	Constant	8.549***	8.928***	7.040***	8.259***
		(1.143)	(1.337)	(1.379)	(1.329)
County Fixed Effects		N	Y	N	Y
Year Effects		N	N	Y	Y
# Observations		91	91	91	91
# Counties		13	13	13	13
R-squared		0.012	0.726	0.145	0.819

Notes: Standard errors in parentheses. Each observation is a county-year. We use data from the 10 years of data in the annual farmer-field-level panel data set we use for our dynamic structural econometric model (2005, 2008-2012, 2014-2016, and 2019); these 10 years of data are the 10 years for which we have data on conventional and organic rice prices. Inverse variance weights that weight county-years with more farmer-fields growing rice more heavily are used. By 'pesticide', we mean 'unapproved synthetic pesticide'. Since yield data is only available at the county level, not the farmer-field level, we estimate the production function using county averages, where each observation is a county-year in which data on rice yield per farmer-yield growing rice data is available. In particular, we regress log rice yield per farmer-field growing rice in a county-year on the fraction of farmer-fields growing rice in a county-year who applied an unapproved synthetic pesticide that year and the average over farmer-fields growing rice in a county-year of the number of previous consecutive years no unapproved synthetic pesticide was used. 'Full Information' specifications account for soil microbes, and therefore include as a regressor the average over farmer-fields growing rice in a county-year of the number of previous consecutive years no unapproved synthetic pesticide was used. In North America, 1 hundredweight equals 100 pounds (U.S. Department of Agriculture [USDA], 1992). Significance codes: *** p<0.001, ** p<0.01, * p<0.05

B PUR Data Coverage

Since farmers who never use pesticides may not be included in the PUR data, it is possible that our data set may under-represent farmers who never use pesticides.

To gauge how many farmers who never use pesticides get missed by the PUR data, we compare the data on acres planted among farmer-field-years in our data set with USDA Survey data on rice acres planted for the state of California at the county level over the years 1990-2019 from USDA NASS Quick Stats (U.S. Department of Agriculture [USDA], 2025), and with USDA Census data on rice acres harvested at the county level for the years 1997, 2002, 2007, 2012, and 2017 (which are the years for which USDA Census data is available during the 1990-2019 period of our data set) from USDA NASS Quick Stats (U.S. Department of Agriculture [USDA], 2025).

The measure of acres planted reported in the PUR data may not be entirely reliable since farmers sometimes report actual acres planted but other times instead report the entire area of the larger plot of land in which the planting occurred (California Department of Pesticide Regulation [CA DPR], 2023).

We use a conservative version of our county average variables in which we drop county-year observations that include a farmer-field-year for which the number of acres planted to rice in that farmer-field-year in that county in that year is missing.

Table B.1 compares the rice acreage data in the PUR data and in the USDA data. Panel A compares summary statistics for rice acres using only county-year observations that have both USDA acres planted data and conservative PUR acreage data. The conservative measure of acres planted reported in PUR are similar to the acres planted and acres harvested reported by the USDA. The maximum value of the PUR measure exceeds the maximum value of each of the USDA measures, however, consistent with the notion that farmers in PUR are sometimes reporting the area of the entire property in which the specific parcel of land where rice was planted is located, instead of just the area of the specific parcel of land where rice was planted. On the other hand, the minimum value of the PUR measure is lower than the minium value of each of the USDA measures, which is consistent with the possibility that the DPR data may not include farmers who never use pesticides.

Panel B compares summary statistics for rice acres using only county-year observations that have both USDA acres harvested data and conservative PUR acreage data. As before, the maximum value of the PUR measure exceeds the maximum value of each of the USDA measures, which again may potentially because farmers in PUR have been known to sometimes report the entire acreage of their property, instead of just the acreage of the plot of land where their crop is being grown. Also as before, the minimum value of the PUR measure is lower than the minium value of each of the USDA measures, which again is consistent with the possibility that the DPR data may not include farmers who never use pesticides.

Panel C compares summary statistics for rice acres using only county-year observations with both USDA acres planted data and conservative PUR acreage data, and for which the reported PUR acres

does not exceed USDA acres planted. The minimum value of the PUR measure is lower than the minium value of each of the USDA measures, which again is consistent with the possibility that the DPR data may not include farmers who never use pesticides. Nevertheless, if we were to believe that the reported acres planted in PUR for this subset of observation are accurate, then the results in Panel C suggest that the PUR data captures approximately 94.5% of the acres of rice planted in the state of California.

Table B.1: Acreage in PUR vs. USDA data, 1990-2019

	Mean	Std. Dev.	Min	Max	# Obs
A: County-year observations with both USDA acres planted data and conservative PUR acreage data					
Number of acres growing rice in county-year (CA DPR PUR)	44,401	46,629	49.7	174,900	294
Number of rice acres planted in county-year (USDA Survey)	44,719	47,528	500	164,000	294
Number of rice acres harvested in county-year (USDA Census)	55,379	50,108	915	153,224	38
B: County-year observations with both USDA acres harvested data and conservative PUR acreage data					
Number of acres growing rice in county-year (CA DPR PUR)	48,499	48,745	466	174,900	47
Number of rice acres planted in county-year (USDA Survey)	55,424	50,348	800	157,000	38
Number of rice acres harvested in county-year (USDA Census)	47,918	48,678	219	153,224	47
C: County-year observations with both USDA acres planted data and conservative PUR acreage data, and for which reported PUR acres does not exceed USDA acres planted					
Number of acres growing rice in county-year (CA DPR PUR)	57,249	49,862	49.7	148,823	136
Number of rice acres planted in county-year (USDA Survey)	60,579	52,702	500	164,000	136
Number of rice acres harvested in county-year (USDA Census)	55,057	48,989	915	147,817	23

Notes: Each observation is a county-year. We use data from 1990-2019. Panel A compares summary statistics for rice acres using only county-year observations that have both USDA acres planted data and conservative PUR acreage data. Panel B compares summary statistics for rice acres using only county-year observations that have both USDA acres harvested data and conservative PUR acreage data. Panel C compares summary statistics for rice acres using only county-year observations with both USDA acres planted data and conservative PUR acreage data, and for which the reported PUR acres does not exceed USDA acres planted.

Data Sources: CA DPR PUR (California Department of Pesticide Regulation [CA DPR], 2024), USDA NASS Quick Stats (U.S. Department of Agriculture [USDA], 2025).

C Risk Aversion

C.1 Constant Absolute Risk Aversion (CARA)

To distinguish between misperception and risk aversion, we also consider a specification of our dynamic structural econometric model that incorporates risk aversion by using a constant absolute risk aversion (CARA) utility function for the deterministic payoff (instead of a linear utility function) when a farmer plants rice.

In particular, if farmer-field i plants rice in year t ($a_{it} \in \{\text{RC}, \text{RN}\}$), then that farmer-field receives the following deterministic payoff $u_0(\cdot)$ in year t:

$$u_0(a_{it} \in \{\text{RC}, \text{RN}\}, \mathbf{s_{it}}; \theta) = \begin{cases} \frac{1-\exp(-\sigma \cdot (P(a_{it}, \mathbf{s_{it}}) \cdot \breve{q}(c_{it}, \mathbf{s_{it}}; \theta) - \text{cost}(c_{it}, \breve{q}(c_{it}, \mathbf{s_{it}}; \theta); \theta)))}{\sigma} & \sigma \neq 0 \\ P(a_{it}, \mathbf{s_{it}}) \cdot \breve{q}(c_{it}, \mathbf{s_{it}}; \theta) - \text{cost}(c_{it}, \breve{q}(c_{it}, \mathbf{s_{it}}; \theta); \theta) & \sigma = 0, \end{cases} \tag{C.1}$$

where σ is the coefficient of absolute relative risk aversion, measures the degree of risk preference, and is an additional structural parameter to be estimated. When $\sigma > 0$, the farmer is risk averse; when $\sigma < 0$ the farmer is risk seeking. When $\sigma = 0$, the farmer is risk neutral, and the deterministic payoff when a farmer-field plants rice corresponds to the linear payoff function in equation (4) from our base case specification.

As before, so that we can identify the parameters in the per-period payoff, we normalize the deterministic payoff for a farmer-field who does not plant rice and uses pesticide ($a_{it} = \text{OC}$) to be 0:

$$u_0(a_{it} = \text{OC}, \mathbf{s_{it}}; \theta) = 0. \tag{C.2}$$

Also as before, to allow the per-period payoff to possibly differ when the farmer-field did not plant rice and also did not use pesticide, the deterministic per-period payoff for not planting rice that year and not using pesticide ($a_{it} = \text{ON}$) is set equal to a parameter ν to be estimated:

$$u_0(a_{it} = \text{ON}, \mathbf{s_{it}}; \theta) = \nu. \tag{C.3}$$

For the cost function, we use the cost function in Specification (1) of Table 6 since this specification is favored by pairwise likelihood ratio tests (Taboga, 2021).

For this CARA version of the structural model, we focus on distinguishing between misperception and risk aversion, and therefore focus on estimating the structural parameters $\theta = (\gamma, \sigma)$, where γ are misperception parameters and σ is the coefficient of absolute relative risk aversion. We hold the cost parameters κ and the deterministic payoff ν when a farmer-field does not plant rice and does not use pesticide fixed at the point estimates from Specification (1) of Table 6.

As seen in Table C.1, farmers are close to risk neutral ($\hat{\sigma} = 0.02$), and, even when allowing for risk aversion, farmers still underestimate the the beneficial effects of clean soil stock on yield ($\hat{\gamma}_k < 0$). Both the misperception parameters γ_c and γ_k are close to the values we estimated in Table 6 in the

absence of risk aversion, but in both cases the absolute value of the misperception parameters under risk aversion are somewhat smaller than they are in the absence of risk aversion.

Thus, farmers have low levels of risk aversion and appear close to risk neutral, and our results regarding farmers' misperception of the effects of clean soil stock and pesticide use on yield are robust to the inclusion of risk aversion. Farmers' choices regarding pesticide use therefore reflect misperception and cannot be completely explained away by risk aversion.

Table C.1: Dynamic Structural Model with Risk Aversion: CARA

		(1)
Parameters in the perceived yield $\breve{q}(\cdot)$		
γ_c	Misperception of how log yield is affected by pesticide use c_{it}	-0.653***
		(0.009)
γ_k	Misperception of how log yield is affected by clean soil stock k_{it}	-0.199***
		(0.001)
Parameter measuring risk aversion		
σ	Coefficient of absolute relative risk aversion	0.0238***
		(0.0009)
Log likelihood		-6,376
# Observations		67,230
# Farmer-fields		17,695
# Counties		15

Notes: Standard errors in parentheses. Each observation is a farmer-field-year. Clean soil stock k_{it} is measured by the number of previous consecutive years (up to 15 years) no unapproved synthetic pesticide was used. Yield is in units of 1 million pounds. Payoffs and costs are in units of ten thousand 2010 USD. κ and ν are held fixed at point estimates from Specification (1) of Table 6. Significance codes: *** p<0.001, ** p<0.01, * p<0.05

C.2 Constant Relative Risk Aversion (CRRA)

To distinguish between misperception and risk aversion, we also consider a specification of our dynamic structural econometric model that incorporates risk aversion by using a constant relative risk aversion (CRRA) utility function for the deterministic payoff (instead of a linear utility function) when a farmer plants rice.

In particular, if farmer-field i plants rice in year t ($a_{it} \in \{\text{RC}, \text{RN}\}$), then that farmer-field receives the following deterministic payoff $u_0(\cdot)$ in year t:

$$u_0(a_{it} \in \{\text{RC}, \text{RN}\}, \mathbf{s_{it}}; \theta) = \frac{(P(a_{it}, \mathbf{s_{it}}) \cdot \breve{q}(c_{it}, \mathbf{s_{it}}; \theta) - \text{cost}(c_{it}, \breve{q}(c_{it}, \mathbf{s_{it}}; \theta); \theta))^{1-\eta}}{1-\eta}, \tag{C.4}$$

where η is the coefficient of constant relative risk aversion and is an additional structural parameter to be estimated. When $\eta = 0$, the farmer is risk neutral, and the deterministic payoff when a farmer-field plants rice corresponds to the linear payoff function in equation (4) from our base case specification.

As before, so that we can identify the parameters in the per-period payoff, we normalize the deterministic payoff for a farmer-field who does not plant rice and uses pesticide ($a_{it} = \text{OC}$) to be 0:

$$u_0(a_{it} = \text{OC}, \mathbf{s_{it}}; \theta) = 0. \tag{C.5}$$

Also as before, to allow the per-period payoff to possibly differ when the farmer-field did not plant rice and also did not use pesticide, the deterministic per-period payoff for not planting rice that year and not using pesticide ($a_{it} = \text{ON}$) is set equal to a parameter ν to be estimated:

$$u_0(a_{it} = \text{ON}, \mathbf{s_{it}}; \theta) = \nu. \tag{C.6}$$

For the cost function, we use the cost function in Specification (1) of Table 6 since this specification is favored by pairwise likelihood ratio tests (Taboga, 2021).

For this CRRA version of the structural model, the structural parameters to be estimated are $\theta = (\gamma, \kappa, \nu, \eta)$, where γ are misperception parameters, κ are cost parameters, ν is the deterministic payoff when a farmer-field does not plant rice and does not use pesticide (this is relative to a deterministic payoff of 0 when a farmer-field does not plant rice and uses pesticide)[8], and η is the coefficient of constant relative risk aversion.

For a first-cut at assessing whether farmers are risk averse, we calculate the log likelihood for different values of the coefficient of constant relative risk aversion η from 0 (risk neutral) to 1.5 (high risk aversion), using the point estimates from Specification (1) of Table 6 for all other structural parameters γ, κ, and ν. As seen in Table C.2, these findings suggest that farmer's choices are best explained by risk preferences that lie somewhere between risk neutral and low risk aversion. Among the risk aversion parameter values for which we have calculated log likelihoods, the value of

[8]For our dynamic structural model, payoffs (and therefore costs) are in units of ten thousand 2010 USD.

the constant relative risk aversion η yielding the largest log likelihood, $\eta = 0.3$ corresponds to rice producers exhibiting low risk aversion.

We conduct a likelihood ratio test (Taboga, 2021) to compare the model with the best-fit η (i.e., $\eta = 3$) with the base case model without risk aversion (i.e., $\eta = 0$). This is a conservative version of the likelihood ratio test, since the log likelihood from estimating all the parameters in the dynamic structural model with risk aversion would be weakly higher. The test statistic D is distributed chi-squared with 1 degree of freedom, and is equal to 388. As the test statistic D is greater than the critical value 3.841, the data reject the constrained model in which there is no risk aversion since the unconstrained model with risk aversion produces a statistically significant improvement in the ability of the model to fit data at a 5% level.

Thus, our results suggest that farmers' risk aversion lie somewhere between risk neutral and low risk aversion.

Table C.2: Dynamic Structural Model with Risk Aversion: CRRA

Coefficient of constant relative risk aversion		Log Likelihood
$\eta = 0$	Risk neutral	-7,184
$\eta = 0.3$	Very low risk aversion	-6,990
$\eta = 0.5$	Low risk aversion	-7,021
$\eta = 1$	Moderate risk aversion	-17,234
$\eta = 1.5$	High risk aversion	-80,997

Notes: Table presents the log likelihood for different values of the coefficient of constant relative risk aversion η from 0 (risk neutral) to 1.5 (high risk aversion), using the point estimates from Specification (1) of Table 6 for all other structural parameters γ, κ, and ν.