Evaluating Optimal Farm Management of Phosphorus Fertilizer Inputs with Partial Observability of Legacy Soil Stocks*

Chanheung Cho Zachary S. Brown David M. Kling Luke Gatiboni Justin S. Baker

Abstract

Decades of intensive fertilizer application have led to the accumulation of phosphorus (P) in soils across US cropland. This over-application can have negative consequences for water quality, but a portion of the accumulated P in soils can serve as a substitute for increasingly costly fertilizer applications. We investigate whether it is economical for farmers to utilize bioavailable legacy soil P stocks (by reducing P fertilizer use) when they are imperfectly observed and soil sampling is costly. Using legacy P measurements from corn field trials spanning over a decade in eastern North Carolina, we develop a dynamic model of this decision problem, which is a natural fit for the 'partial-observability Markov decision process' (POMDP) framework. In a novel contribution to the POMDP literature, we analyze how agent preferences over risk and intertemporal substitution affect optimal monitoring and resource use by incorporating an Epstein-Zin preference structure. Using contemporary computational methods for analyzing POMDPs, simulation results indicate that more risk-averse agents optimally apply more fertilizer in a range of bioavailable legacy P stocks, reflecting fertilizer's dual role as both an immediate input to crop production and a forward-looking investment in the soil P capital stock. Results of sensitivity analyses suggest that optimizing farmers are relatively unresponsive to sustained increases in stochastic P fertilizer prices or to decreased monitoring costs. We discuss the implications of these findings for potential interventions aimed at reducing environmental externalities of P fertilizer by providing better and cheaper information to farmers on their legacy P soil stocks.

JEL Codes: Q15, Q24, C61, C63
Keywords: Legacy Phosphorus, Risk Preference, State Uncertainty, Epstein-Zin Preference, Partial-Observability Markov Decision Process

*Cho: North Carolina State University (ccho5@ncsu.edu); Brown: North Carolina State University (zsbrown2@ncsu.edu); Kling: Oregon State University (David.Kling@oregonstate.edu); Gatiboni: North Carolina State University (luke_gatiboni@ncsu.edu); Baker: North Carolina State University (jsbaker4@ncsu.edu). The authors thank participants at various conferences and seminars for useful comments and discussions.
Financial Support from the National Science Foundation (NSF) Award CBET-2019435: Science and Technologies for Phosphorus Sustainability.

1 Introduction

Economically efficient management of the agricultural nutrient phosphorus (P) is a critical global challenge to ensure sustainable food production and environmental quality protection. P is imbalanced in the global food system, and some regions lack sufficient access to synthetic or organic P fertilizers that could boost yields and rural incomes, leaving producers to rely on limited P stocks in nutrient-deficient soils (Zou et al. 2022). In the United States (and other advanced economies) the main social challenge in P management is the excessive application of P fertilizer on farmlands, which eutrophies and degrades surface water. In addition, there are concerns that the overuse of P fertilizers in advanced economies depletes mineral stocks and increases prices. However, P fertilizer consumption by US farmers has remained relatively stable over the last few decades and has evidently responded only temporarily to recent and persistent price increases, suggesting relatively price inelastic demand for P in US cropping systems (Denbaly and Vroomen 1993, Sud 2020, Kile et al. 2025).

Notably, unlike nitrogen fertilizer, P fertilizer application residuals after crop uptake accumulate in soils. This accumulating soil stock of P—referred to in this manuscript as 'legacy P'—can be stored in non-bioavailable reserves, taken up by future crop plantings, or mobilized by subsequent precipitation events, flowing into water bodies. A significant amount of agricultural land in the US has accumulated legacy P stocks over decades of continuous cultivation and application of P from synthetic and organic sources (for example, annually, $>$ 1,000 tonnes of P have been accumulated in the agricultural region of Vermont) (Wironen et al. 2018, Ringeval et al. 2018). Phosphorus runoff into surface water bodies catalyzes eutrophication, which can lead to hypoxic 'dead zones' and greenhouse gas (GHG) emissions (Arrow et al. 2018, Conley et al. 2009, Iho and Laukkanen 2012, Rabotyagov et al. 2014b, Beaulieu et al. 2019, Paudel and Crago 2020, Downing et al. 2021). Downing et al. (2021) estimate a substantial economic cost associated with GHG emissions from eutrophication in freshwater systems globally.

Various policies have been proposed to mitigate environmental issues arising from the overuse of P fertilizers, including the Numeric Nutrient Criteria under the Clean Water Acts and Binational Phosphorus Reduction Strategy in Lake Erie (US EPA 1995, Lake Erie LaMP 2011), with one notable proposal focusing on incentivizing farmers to substitute legacy, soil-bound P stocks for P fertilizer and to reduce P fertilizer applications (Sattari et al. 2012, USDA 2020). Properly managed, bioavailable legacy P stocks can substitute for P fertilizer, reducing costs and environmental impacts

from intensive crop operations (Sattari et al. 2012). However, this proposal raises the question of why farmers, in many cases, do not currently utilize legacy P stocks, given their accumulation over time and the potential cost savings for farmers from doing so? This puzzle is more pronounced in areas with publicly available information on soil P content provided by state extension services.

This paper studies this question using a model that incorporates biophysical crop production and legacy P stock dynamics for a representative agricultural system with dynamic farm-scale management incentives and behavioral factors to simulate P stock dynamics in a setting of imperfect information on legacy P bioavailability, market uncertainty, and risk aversion. The complexity of managing legacy P stocks poses significant challenges for farmers and the economic benefits of different strategies recommended by agricultural extension are uncertain. Recent analysis suggests that farmers may not fully account for these residual P stocks in their P fertilizer application decisions due to a lack of high-quality information and the inherent uncertainty about the quantity and bioavailability of legacy P stocks across their farmland. When accounting for farmer risk aversion, the uncertainty surrounding legacy P could contribute to its under-utilization. This paper explores how these factors affect the intertemporal dynamics of legacy P stocks and utilization, and examines whether improved access to enhanced (and higher cost) monitoring of legacy P stocks could reduce P fertilizer application.

To address the management of legacy P accumulated in soil and its losses to surface water, previous studies have analyzed the optimization of fertilizers in farmland along with P control or conservation policies. Schnitkey and Miranda (1993) analyze the optimal steady-state application of fertilizer under various policy settings which limit soil P levels. Goetz and Zilberman (2000) examine intertemporal and spatially optimal application of mineral fertilizer levels given P concentrations in bodies of water associated with agricultural land for optimal lake restoration policy. Innes (2000) explains that environmental impact of nutrient runoff from livestock production can be mitigated by regulating facility size, implementing waste policies based on cleanup costs, and combining fertilizer taxes with subsidies for manure spreading equipment. Lötjönen et al. (2020) provide a theoretical spatial modeling framework to study climate and water policies for P mineral and manure fertilizer use in dairy farm management. While the models in these studies account for optimal fertilizer usage decisions to manage P accumulation in soils and to reduce P loss to the surface water, they do not incorporate the observational uncertainty related to legacy P, and thus cannot answer questions about soil testing which we address here.

Farmers in the US do typically have some baseline information about soil P, as US farmers

commonly employ standard soil sampling, provided by state agencies or extension services and by private soil testing service laboratories at nominal fees. These tests can help gauge legacy P availability, among other soil health metrics. Soil tests are usually conducted at a few spots within fields, offering preliminary insight into soil P content, and serving as noisy indicators of the actual bioavailable legacy P stock across a field (Austin et al. 2020). While more comprehensive sampling options exist (e.g. point sampling, cell sampling, and zone sampling), offering clearer information on legacy P heterogeneity across a field, they come at a higher cost, presenting a trade-off between accuracy and expense (Austin et al. 2020, Gatiboni et al. 2022).

Economically, this situation can be described as one in which the agent—here, the farmer—optimizes their utilization of an uncertain resource stock—here, legacy P—in which they may dynamically update their beliefs about these fluctuating stocks based on costly monitoring. Generically, this situation represents a common class of problems in the resource management literature, referred to as a 'partial-observability Markov decision process' or POMDP (Clark 2010, Fackler and Pacifici 2014, Fackler 2014). Previous applications of POMDP models and extensions in resource management have included invasive species control (Haight and Polasky 2010, Rout et al. 2014, Kling et al. 2017), forestry (Sloggy et al. 2020), environmental conservation (White 2005), erosion prevention (Tomberlin and Ish 2007), and infectious diseases (Chadès et al. 2011).

While POMDP methods have been applied in related renewable resource contexts, such as invasive species control and forest management, their application to farm production economics remains limited, particularly in settings involving partially observed and depletable inputs. This reflects one contribution of this paper. Previous agricultural economics studies have addressed the partial observability and monitoring problem using more heuristic optimization methods that separate inference about unobserved state variables from the optimization. For example, Fan et al. (2020) employ such an approach using state-space models to analyze efficient monitoring of an agricultural pest, but they specifically note the theoretical superiority of a POMDP approach for their application were it not for the computational difficulty of applying these methods in their study.

Additionally, as far as we are aware, agent risk preferences have not previously been included in POMDP applications, at least in agricultural or resource economics. It is natural to conjecture that risk aversion could strongly affect demand for and enhanced monitoring of soil P stocks, and the associated demand for applied P fertilizer. Our analysis of that general conjecture represents another contribution. Because standard discounted expected utility in dynamic economic models

conflates preference parameters for risk aversion and intertemporal substitution, we employ a widely used recursive utility Epstein-Zin specification to disentangle these effects in our analysis (Epstein and Zin 1991).

We develop our model's empirical foundation through econometric analysis of North Carolina field data on legacy P abundance, stock accumulation, fertilizer application, and yield response in a corn-farming context spanning over a decade. We also account for stochastic crop and P fertilizer price dynamics, which we jointly estimate using publicly available USDA data. This is important because both prices are volatile, and P fertilizer—by accumulating in soils—can act as a form of capital investment whose cost is uncertain. Such price stochasticity introduces an option value from delaying investment, even for risk-neutral farmers, as shown in classical investment under uncertainty models (McDonald and Siegel 1986, Pindyck 1993, Dixit and Pindyck 1994). Accounting for these price dynamics requires extending our POMDP framework into what are known as a 'mixed-observability Markov decision processes' or MOMDPs (Chades et al. 2012, Sloggy et al. 2020). Their inclusion increases the robustness of our analysis, given that previous studies show that stochastic price dynamics have important effects on other dynamic farm resource management problems, such as crop rotation and cover crop planting (Livingston et al. 2015, Chen 2022).

Including all the elements described above is a significant computational challenge. The problem is made more computationally complex due to the continuous legacy P state variable. In particular, continuous-state POMDPs involve stochastic dynamic programming, in which the agents possess belief states that specify their current subjective probability distributions about imperfectly observed biophysical states, with these beliefs states updated via Bayes' Rule. Using probability distributions as state variables leads to a high-dimensional state space that imposes considerable challenges for numerical computation. To address these challenges, we closely follow recently-applied density projection methods (e.g. Zhou et al. 2010, Springborn and Sanchirico 2013, MacLachlan et al. 2017, Kling et al. 2017, Sloggy et al. 2020) which use a low-dimension approximation of the belief state, while avoiding some of the restrictions and potential pitfalls of prior methods (e.g. restricting model structure to permit the use of conjugate priors or discretization of the unobserved state). We also use an econometric approach in estimating price dynamics that aids numerical tractability in the MOMDP optimization while remaining empirically grounded: We first econometrically estimate a Markov-switching Vector Autoregressive (MSVAR) model for the price dynamics (supported by statistical tests), and then in the dynamic programming impose a

conditional, within-regime equilibrium assumption to maintain computational tractability.

We find that optimizing farmers in the model generally employ more intensive and costly soil sampling only at low levels of estimated legacy P stocks, and typically only early in the simulation horizon. Higher risk aversion generally decreases the reliance on legacy P stocks in favor of fertilizer application. Meanwhile, farmer preferences for profit-smoothing over time do not appreciably affect optimal fertilizer use or monitoring. Furthermore, sensitivity analysis with much higher fertilizer prices (e.g., from a sustained global market disruption or a fertilizer tax) or much lower monitoring costs (e.g., from a subsidy for more intensive soil testing) do not induce much substitution from fertilizer to legacy P use. These results suggest limited effectiveness of proposed policies to promote enhanced soil P testing (either through cost-subsidization or other means) to decrease P fertilizer applications.

This paper's sections proceed as follows. First, a model of legacy P dynamics and crop production is described, capturing both the accumulation and bioavailability of legacy P. Next, the economic and management problems are discussed, describing how farmers can evaluate the recursive expected utility of their controls, P fertilizer application, and soil sampling in the face of stochastic prices and the unobservable state of legacy P. Then, the methodological framework and model specification are presented, including price dynamics and the density projection approach for managing Bayesian belief updating. The application of this model to the corn market provides a practical example of how it can be used to guide decision-making in agriculture. We then present results of the model across a range of scenarios and sensitivity analyses, and conclude with a discussion of generalizable lessons from these analyses, their limitations, and considerations for future research.

2 Model Description and Computational Methods

This section outlines the key components of the model: the dynamics of legacy P in soil, the information farmers receive through soil sampling, their fertilizer and sampling decisions, and the economic optimization framework that guides these choices under uncertainty.

2.1 A Model of Stochastic Legacy Phosphorus Dynamics

We start a deterministic model of legacy P dynamics from Iho and Laukkanen (2012), which we enhance by incorporating realistic stochasticity. The average soil-accumulated legacy P stock per

hectare is given by L_t, with its dynamics specified in the following recursive equation:

$$L_{t+1} = \rho_t L_t + (\gamma_1 + \gamma_2 L_t) \underbrace{\left[F_t - \overbrace{(\gamma_3 \log(L_t) + \gamma_4)}^{\text{Concentration on Yield}} \; Y(L_t, F_t, \varepsilon_t^Y) \right]}_{\text{Legacy P Surplus}} \tag{1}$$

where ρ_t is a 'carry-over' parameter of legacy P, F_t represents the amount of P fertilizer input, and $Y(L_t, F_t, \varepsilon_t^Y)$ is a crop yield function at time t, including a stochastic yield shocks ε_t^Y. The term $(\gamma_3 \log(L_t) + \gamma_4)$ defines the legacy P concentration of the crop yield, which increases logarithmically with L_t. As L_t increases, this concentration on yields also rises, though at a diminishing rate. The term $(\gamma_1 + \gamma_2 L_t)$ explains the change in next-period legacy P stocks brought about by a current net surplus or deficit in legacy P balance (Ekholm et al. 2005). The parameter values of γ are summarized in Table 4.

While there are several empirically-grounded ways to introduce stochastic legacy P dynamics in this model, we focus on stochastic transport into the environment, owing to precipitation and other environmental factors. In a deterministic model, a carry-over parameter $\rho_t < 1$ implies a decay of soil P stock on farmland in the absence of further fertilizer inputs, e.g. loss of soil-bound P to surface water systems (Ekholm et al. 2005, Iho and Laukkanen 2012). We introduce stochasticity into legacy P dynamics by specifying this carry-over parameter as a random variable:

$$\rho_t = \exp\left[\left(\mu_\rho - \frac{\sigma_\rho^2(L_t)}{2} + \sigma_\rho(L_t) W_t \right) \right], \qquad \text{with } W_t \sim \mathcal{N}(0,1), \tag{2}$$

where μ_ρ is the log-mean of ρ_t (so that $\mathbb{E}[\rho_t] = \exp \mu_\rho$) and $\sigma_\rho(L_t)$ is the standard deviation of log ρ_t, with $\sigma_\rho(L_t)$ specified as potentially a function current stocks L_t. We assume $\mu_\rho < 0$, so that the legacy P stock available for crop uptake stochastically diminishes without added P fertilizer F_t.

Note the log-normal distribution of ρ_t means that a fixed standard deviation σ_ρ would result in the conditional variance of the annual change in legacy P stocks growing without bound as L_t grows (i.e. $\lim_{L_t \to \infty} \text{Var}(L_{t+1}|L_t) = \infty$), which is not biophysically realistic. Following previous studies that have dealt with similar issues (Loury 1978, Gilbert 1979, Melbourne and Hastings 2008, Sims et al. 2017, Sloggy et al. 2020), we therefore specify the log standard deviation as a decreasing function of the stock. Specifically, in our main specification, we assume that the portion of the stock carried over to the next period ($\rho_t L_t$) has a fixed variance ς^2, invariant with the current stock level L_t. This assumption implies the log standard deviation function takes the

form $\sigma_\rho(L_t) = \sqrt{\ln(1+\varsigma^2 \exp(-2\mu_\rho)/L_t^2)}$.[1] We investigate the importance of this assumption by also considering a fixed log standard deviation ($\sigma_\rho(L_t) = \bar{\sigma}$) in the Appendix.

2.2 Soil Sampling and Partial Observability

Legacy P is not perfectly observed, but farmers in the model receive information through soil sampling. We consider two kinds of soil sampling: standard sampling (*ss*) and point sampling (*ps*). Standard sampling, typically provided by state agencies or extension services at nominal fees, involves collecting samples from a few spots within fields. These tests offer a noisy indicator of the actual bioavailable legacy P stock across a whole field (Austin et al. 2020). Point sampling, on the other hand, involves collecting multiple samples at specific grid points or random locations within grid cells, providing more precise information on legacy P bioavailability, but at a higher cost (Austin et al. 2020, Gatiboni et al. 2022).[2]

To specify this observation process in our model, we denote the current soil sample test result as O_t. As a noisy measure of legacy P across a hectare of farmland, we assume a multiplicative test error λ which is zero-truncated and normally distributed with variance σ_s^2 determined by the type of sampling $s \in \{ss, ps\}$.

$$O_t^s = \lambda_t^s L_t, \qquad \text{where } \lambda_t^s \sim \mathcal{N}_{(0,\infty)}(1, \sigma_s^2). \tag{3}$$

The additional information gained from point relative to standard sampling is captured by the assumption that $\sigma_{ss} > \sigma_{ps}$. In principle, truncation also implies that for a small enough (but positive) level of legacy P relative to the test error variance σ_s^2, the soil test result may measure soil P, which is physically possible though highly uncommon (suggesting generally good test accuracy).

The farmer's beliefs about the distribution of legacy P are denoted by the pdf $b_t(L_t)$, representing a subjective probability distribution over the unobserved L_t, conditional upon the history of control choices and resulting observations (Porta et al. 2006). Bayesian updating of these beliefs combines each period's prior beliefs regarding L_t, with projected dynamics for L_{t+1}, along with

[1] In reality, yield shocks ε_t^y and shocks to the legacy P carryover process ρ_t may share common sources -such as rainfall or temperature- suggesting potential correlation. However, to maintain tractability in belief updating and value function computation, we assume that these stochastic components are independent.

[2] We exclude the no sampling case in our main analysis because in practice, commercial farmers in the US almost always conduct at least standard sampling, which is offered by state agencies for a nominal fee. This was confirmed in a more elaborate version of the model, which allowed for a no-sampling option: When the sampling cost is negligible, then intuitively the farmer would always acquire the low cost information.

new information O^s_{t+1}, via the following:

$$b_{t+1}(L_{t+1}) \propto p(O^s_{t+1} \mid L_{t+1}, s_t) \int p(L_{t+1} \mid L_t, F_t) b_t(L_t) dL_t \tag{4}$$

with a given $b_0(L_0)$ specifying the prior beliefs about initial stocks and where $p(O^s_{t+1}|L_{t+1}, s_t)$ is the conditional pdf of the observation. The Markovian properties ensure that the next-period beliefs only depend on the current beliefs, controls, and information gained in the current period. Figure 1 illustrates how the farmer updates their belief state b_t based on their soil sampling decision s_t and resulting soil test result O^s_t.[3]

2.3 Economics and Management

Annual payoffs in the model are evaluated as the profit determined by the crop yield function $Y(L_t, F_t, \varepsilon^Y_t)$ and stochastic prices:

$$\pi(L_t, F_t, P^Y_{t+1}, P^F_t, s_t, \varepsilon^Y_t) = P^Y_{t+1} Y_t(L_t, F_t, \varepsilon^Y_t) - P^F_t F_t - c_s s_t, \tag{5}$$

where P^Y_{t+1} and P^F_t are prices for the crop and P fertilizer, respectively, ε^Y_t is an exogenous yield shock term, and c_s is a soil sampling cost (with $c_{ss} < c_{ps}$), and $s_t \in \{ss, ps\}$ reflects the soil sampling decision at time t. Fertilizer application decisions are based on the observed fertilizer price P^F_t at the time of application, whereas the crop price P^Y_{t+1} is only realized at the end of the season and not yet observed at the time of applying fertilizer (Chen, 2022). This means that the decision to apply fertilizer is informed by the current fertilizer price and the last harvest's crop price. Dynamics for the prices $\boldsymbol{P}_t = [P^Y_t, P^F_t]$ are assumed to be determined by a joint Markov process, such that $\boldsymbol{P}_{t+1} = G(\boldsymbol{P}_t, \boldsymbol{\epsilon}_t)$ where $G(\cdot)$ is a transition function and $\boldsymbol{\epsilon}_t$ is a vector of price shocks driven by macroeconomic conditions or short-term exogenous shocks.[4] We discuss the specific structure used for these dynamics below in econometric estimation.

A risk-neutral farmer agent with no preference for profit-smoothing over time and a fixed discount rate would seek to maximize the expected net present value (ENPV) of their profits. For

[3] In principle, in addition to soil test results, farmers could infer the adequacy of their soil P stocks through observed yields (e.g. by observing yields when no fertilizer is applied). Modeling belief-updating with this additional information source is significantly more complicated. However, we did undertake this effort, the results of which - shown in the Appendix - suggest that at least in our application such a yield signal provides very little information relative to soil tests. We thus exclude this additional complication from the main model and results presented here.

[4] Here, $G(\cdot)$ generically denotes the transition function mapping current prices and shocks to next-period prices, with $f(\cdot)$ as the implicit corresponding pdf of next-period prices conditional on current prices.

an agent with an infinite time horizon (or a stochastic time horizon with a constant hazard rate of termination), the Bellman equation characterizes the maximal ENPV as a function $V(\boldsymbol{S}_t)$ of the observed state variables collected in $\boldsymbol{S}_t \equiv [b_t(\cdot), P_t^Y, P_t^F]$, and can be written as follows:

$$V(\boldsymbol{S}_t) = \max_{F,s} \Pi(\boldsymbol{S}_t, F_t, s_t) + \beta \mathbb{E}\{V(\boldsymbol{S}_{t+1}) \mid \boldsymbol{S}_t, F_t, s_t\}, \tag{6}$$

where β is the discount factor and $\Pi(\boldsymbol{S}_t, F_t, s_t)$ are expected end-of-season profits given currently observed states and actions:

$$\Pi(\boldsymbol{S}_t, F_t, s_t) \equiv \iiint \pi(L_t, F_t, P_{t+1}^Y, P_t^F, s_t) f(P_{t+1}^Y \mid P_t^Y, P_t^F) b_t(L_t) g(\varepsilon_t^Y) d\varepsilon_t^Y dP_{t+1}^Y dL_t \tag{7}$$

where $f(P_{t+1}^Y | P_t^Y, P_t^F)$ is the conditional pdf of crop price P_{t+1}^Y at the upcoming harvest, given the last observed harvest price P_t^Y and the current fertilizer price P_t^F.

In this paper, we are interested in studying how risk and intertemporal preferences affect optimal monitoring of the unobserved state of legacy P, L_t. To do so, we generalize the above Bellman equation via the commonly used Epstein-Zin recursive preference structure. Originally developed in the macro-finance literature to allow nontrivial risk premiums in empirically-defensible capital asset pricing models (Epstein and Zin 1989), this preference structure has since been applied in dynamic agricultural production-inventory models (e.g. Lybbert and McPeak 2012), valuation of ecological insurance (Augeraud-Véron et al. 2019), and in integrated assessment models for evaluating the economic damages from climate change (Cai and Lontzek 2019). The key advantage of Epstein-Zin preferences is that they disentangle risk aversion from preferences for intertemporal smoothing, which are conflated in expected discounted utility models. For our purposes, this allows us isolate how risk aversion versus intertemporal smoothing preferences affect optimizing agents' demand for monitoring.

The Bellman equation for the recursive expected utility function, given Epstein-Zin preferences, is as follows:

$$V_{EZ}(\boldsymbol{S}_t) = \max_{F,s} \left[(1-\beta) \Pi_{EZ}(\boldsymbol{S}_t, F_t, s_t)^{1-\psi^{-1}} + \beta \mathbb{E}\{V_{EZ}(\boldsymbol{S}_{t+1})^{1-\eta} \mid \boldsymbol{S}_t, F_t, s_t\}^{\frac{1-\psi^{-1}}{1-\eta}} \right]^{\frac{1}{1-\psi^{-1}}} \tag{8}$$

where $\Pi_{EZ}(\boldsymbol{S}_t, F_t, s_t)$ is the certainty-equivalent expected utility of end-of-season profits:

$$\Pi_{EZ}(\boldsymbol{S}_t, F_t, s_t) \equiv \left(\iiint \pi(L_t, F_t, P_{t+1}^Y, P_t^F, s_t)^{1-\eta} f(P_{t+1}^Y \mid P_t^Y, P_t^F) b_t(L_t) g(\varepsilon_t^Y) d\varepsilon_t^Y dP_{t+1}^Y dL_t \right)^{\frac{1}{1-\eta}} \tag{9}$$

and where η and ψ indicate, respectively, the coefficient of relative risk aversion (RA) and the elasticity of intertemporal substitution (EIS): Higher η and ψ correspond respectively to greater risk aversion and a weaker preference for intertemporal smoothing, with $\eta = \psi^{-1}$ reducing EZ preferences to the expected discounted utility preference structure and $\eta = \psi^{-1} = 0$ ($\psi = \infty$) reducing the EZ Bellman equation to Bellman eq. (6) for an agent who simply seeks to maximize their ENPV of profits.[5]

2.4 Computational Methods: Density Projection and Particle Filtering

One of the key computational challenges in continuous-state POMDPs is that the space of possible belief states $b_t(\cdot)$—the set of all continuous probability distributions over the unobserved state L_t—is infinite-dimensional. Exact belief updating requires integrating over both the transition dynamics and the observation process, but the primary computational obstacle is the curse of dimensionality arising from the infinite-dimensional belief space. To address this challenge, we adopt an approximation method known as *density projection*, which summarizes the belief distribution using a parametric family (in our case, the log-normal distribution) and updates the parameters over time using particle filtering combined with projection. This approach avoids the need for restrictive assumptions about problem structure and allows us to flexibly incorporate nonlinear transition dynamics and non-Gaussian noise, which are common in real-world resource management problems.

Our algorithm for implementing density projection using the particle filtering technique, detailed in the Appendix, closely follows that used by Sloggy et al. (2020) in another MOMDP application in resource economics.[6] In summary, we first specify a parametric distribution family for prior beliefs $b_t(L_t)$- here, a log-normal distribution, parameterized by its arithmetic mean μ^L and coefficient of variation v^L. The method then takes the pdfs for these prior beliefs, the conditional likelihood of the observations $p(O^s_{t+1} \mid L_{t+1}, s_t)$, and the transition dynamics $p(L_{t+1} \mid L_t, F_t)$, and uses particle filtering with Bayes' rule in eq. (4) to simulate draws from the posterior updated beliefs $b_{t+1}(L_{t+1})$. This posterior belief pdf is no longer log-normal; however, density projection is used to fit an approximating log-normal distribution to the posterior draws, by minimizing a measure of distance between the approximating pdf and the true posterior captured in the draws from the

[5] In this case, the value function in the EZ Bellman equation, $V_{EZ}(\boldsymbol{S})$ is simply a rescaling of the risk neutral value function by $V_{EZ}(\boldsymbol{S}) = (1-\beta)V(\boldsymbol{S})$.

[6] Point-based solvers with discretized state spaces (e.g., Memarzadeh et al. (2019)) are powerful but require specifying state bins a priori. Given limited knowledge of legacy P dynamics, we retain a continuous state space to better capture the underlying process dynamics, the trade-off being that we approximate the belief state.

particle filter. Density projection uses the Kullback-Liebler divergence as the distance measure between the approximating and posterior pdfs (Zhou et al., 2010). This results in the approximating distribution's distance-minimizing parameters effectively being maximum-likelihood estimates, treating the particle filter draws as observations. This procedure ensures that belief updating involves only the mean and coefficient of variation of the log-normal distribution, instead of requiring the full specification of an infinite-dimensional nonparametric density.

This density projection procedure is integrated into computation of the dynamic programming solutions, by first discretizing the belief state parameters and actions $(\mu_t^L, \nu_t^L, F_t, s_t)$ and then calculating the discretized transition probabilities for the next-period belief parameters $(\mu_{t+1}^L, \nu_{t+1}^L)$. These transition probabilities are pre-computed, before solving the infinite-horizon Bellman equation using standard value- or policy-iteration algorithms for discrete-state dynamic programming (see Appendix).

3 Application to Eastern North Carolina Corn Farming and Econometric Estimation

We apply the model in the previous section to a representative corn production system in eastern North Carolina, using long-term field trial data containing soil P measurements from this region. While a strict interpretation of our results would suggest they only apply to this specific context, we show in the Appendix (Section D) that it shares many broad features of a generic US corn farming operation. We therefore argue that the results from our analysis—particularly the qualitative patterns and results from sensitivity analysis—provide broadly applicable insights. This section describes the econometric estimation of model parameters for this context. The first subsection describes estimation of the yield function, and the second describes the joint estimation of US corn and P fertilizer price dynamics.

3.1 Production function estimation and model parameterization

To estimate the yield response function, we analyze data covering 5 years of field experiments (2010, 2012, 2014, 2021, and 2022) at the North Carolina Cooperative Extension Tidewater Research Station on the coastal plain. These experiments measured legacy P bioavailability using the Mehlich 3 method reported in milligrams per cubic decimeter of soil (mg/dm^3), P fertilizer application (kg/ha), and corn yield (kg/ha). Further details on these data are provided by (Morales et al. 2023). Table 1 provides the summary statistics from the field trial data. While the distribution of all

Table 1: Summary Statistics of North Carolina Tidewater Data

Variable	Obs	Mean	Median	IQR	SD	Min	Max
Legacy P (mg/dm^3)	139	63.986	46	37–66.75	50	28	279
P application (kg/ha)	139	47.036	22	11–67	53.948	0	168
Corn yield (kg/ha)	139	4751.9	4442	2266.7–6517.9	2950.3	131	13712

Notes: Interquartile Range (IQR) is a measure of statistical dispersion, being equal to the difference between the 75th and 25th percentiles. It represents the range within which the central 50% of the data lie.

measurements is skewed to the right, legacy P (mg/dm^3) measurements are particularly so, a fact we take into account in regression analysis.

To permit flexible estimation of the substitutability between legacy P and fertilizer applications, we use a slight modification of a translog production function, as follows:

$$\ln(Y_{i,t}) = \beta_0 + \beta_1 F_{i,t} + \beta_2 \ln(L_{i,t}) + \beta_3 F_{i,t}^2 + \beta_4 F_{i,t} \ln(L_{i,t}) + \omega_t + \varepsilon_{i,t}^Y, \tag{10}$$

where i denotes experiment plot, ω_t is a year fixed effect, and $\varepsilon_{i,t}^Y$ is a time-varying error component. Also, note that a $\ln(L_{i,t})^2$ term is excluded from the above specification, as would be included in a full translog specification; this is due to (a) a limited ability to identify nonlinear yield effects of $L_{i,t}$ in our setting where we have a limited number of years in the dataset and (b) related concerns about outliers particularly regarding $L_{i,t}$ and significant skew that still remains in these measurements even after taking logs. We estimate this function using OLS.

The results in Table 2 present the regression results. The coefficients on $F_{i,t}$ and $F_{i,t}^2$ together imply an initially positive yield response to P fertilizer inputs, with diminishing marginal productivity from higher fertilizer levels, eventually resulting in a saturation point beyond which further P fertilizer inputs decrease yields. The positive coefficient on $\ln(L_{i,t})$ confirms that higher legacy P improves corn yields; however, because this coefficient is less than unity, it also implies diminishing marginal productivity in legacy P. Finally, the negative interaction term between $F_{i,t}$ and $L_{i,t}$ implies that P fertilizer and legacy P are indeed substitutes. This interaction term also implies that higher legacy P leads to a lower yield saturation point from P fertilizer inputs. For a visualization of this production function, see Figure F1 in the Appendix.

Table 2: Corn yield estimation

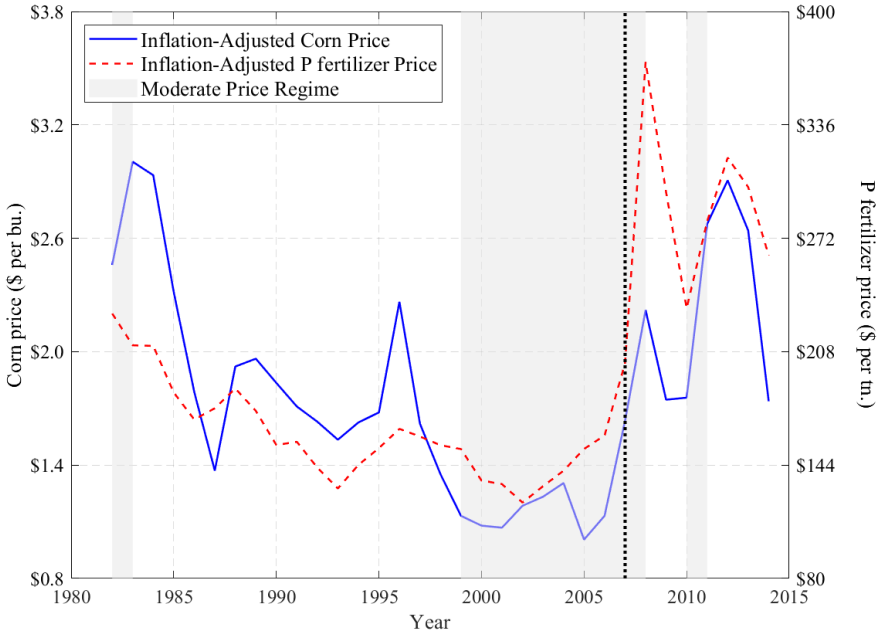
Log Corn Yield (kg/ha)	
$F_{i,t}$	0.043***
	(0.011)
$\ln(L_{i,t})$	0.876*
	(0.481)
$F_{i,t}^2$	-0.00006*
	(0.00003)
$F_{i,t} \times \ln(L_{i,t})$	-0.007*
	(0.003)
Constant	4.910***
	(1.736)
Time Fixed Effect	Yes
Observations	139
Adjusted R-squared	0.5103

Notes: The standard errors are adjusted using the jackknife method with clusters defined by soil sampling plots. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. See eq. (10) for regression specification.

3.2 Corn and phosphorus fertilizer prices

To estimate a dynamic model for corn and fertilizer prices, we analyze USDA time series on corn and P fertilizer prices from 1982 through 2014 (Figure 1). We use P fertilizer (44%-46% phosphate) price data from the USDA "Fertilizer Use and Price" report and corn price data from the USDA's "U.S. Bioenergy Table" (USDA 2024a, USDA 2024b), spanning 33 years (1982-2014). For both empirical reasons and to facilitate MOMDP numerical implementation, we estimate price dynamics using a 'Markov switching vector autoregression' (MSVAR). This econometric method generalizes a standard multivariate time-series vector autoregression by allowing for probabilistic regime-switching in the regression intercepts and coefficients, in order to accommodate qualitative changes observed in the nature of the price dynamics (Hamilton 1989). In our application, use of MSVAR is empirically motivated by observing an abrupt and sustained change in corn and P fertilizer price patterns after ca. 2007, as seen in Fig. 1. Before 2007, the inflation-adjusted prices of both corn and P fertilizer show a clear decreasing trend, whereas after 2007, corn and P fertilizer prices beginning to rise significantly. This rise aligns with the global increase in commodity prices more broadly.

Figure 1: Inflation-adjusted corn and phosphorus fertilizer prices



Notes: Inflation-adjusted prices are adjusted using the Consumer Price Index (CPI) for all urban consumers (index 1983=100), with data sourced from the Federal Reserve Bank of Minneapolis (2024.04). The vertical line marks 2007, where dynamics appear to qualitatively change. The Moderate Price Regime, shown in gray, was estimated based on the Markov-switching Vector Autoregressive model.

We thus estimate a log-linear MSVAR specification of the following form:

$$\ln P_{t+1} = \boldsymbol{\mu}_{(r_{t+1})} + \boldsymbol{\Phi}_{(r_{t+1})} \ln P_t + \boldsymbol{\epsilon}_{t+1}, \quad \text{where} \quad \varepsilon_{t+1} \sim \mathcal{N}(0, \boldsymbol{\Sigma}), \tag{11}$$

where $\ln P_{t+1}$ is the vector of corn and P fertilizer prices, $\ln P_{t+1} \in \mathbb{R}^2$ and $\boldsymbol{\Phi}_{(r_{t+1})}$ represents the autoregressive coefficients that vary depending on the regime r_{t+1}. $\boldsymbol{\mu}_{(r_{t+1})}$ contains the regime-specific regression intercepts, and $\boldsymbol{\Sigma}$ is the covariance matrix of the error terms. The probability of switching to a given regime i next period given a current regime j is $p_{ij} = \Pr(r_{t+1} = i \mid r_t = j)$, which are also estimated (Hamilton 1989). We allow for two price regimes in the model, $r_t \in \{\text{moderate, high}\}$, based on visual inspection of the data.

To estimate the MSVAR, we employ a Bayesian approach used by Osmundsen et al. (2021). This approach uses a Hamiltonian Monte Carlo (HMC) method implemented via the Stan software. Unlike traditional Gibbs sampling, HMC leverages gradient information to explore the posterior distribution efficiently, particularly in high-dimensional parameter spaces.

The likelihood of the model is constructed conditional on the latent state sequence, and prior

Table 3: Markov-switching Vector Autoregressive and transition probabilities

	Corn ($\ln(P^Y_{t+1})$)		Phosphorus fertilizer ($\ln(P^F_{t+1})$)	
	Moderate	High	Moderate	High
$\ln(P^F_t)$	0.105***	0.096***	0.472***	0.469***
	(0.009)	(0.009)	(0.011)	(0.010)
$\ln(P^Y_t)$	0.344***	0.346***	-0.013	-0.008
	(0.010)	(0.010)	(0.009)	(0.009)
$\mu_{(S_t)}$	0.748***	1.274***	0.834***	1.321***
	(0.053)	(0.051)	(0.053)	(0.052)
Variance (Σ_{11})	0.072	Variance (Σ_{22})	0.059	
	(0.0002)		(0.0002)	
Covariance ($\Sigma_{12} = \Sigma_{21}$)	0.0165			
	(0.0001)			
	Transition Probabilities			
	Moderate (t)		High (t)	
Moderate ($t+1$)	0.735 (0.003)		0.264 (0.003)	
High ($t+1$)	0.265 (0.003)		0.736 (0.003)	

Notes: Standard errors are in parentheses. In the estimation, constant variance of residual $\mathbf{\Sigma} = \mathbf{\Sigma}_{(i)} = \mathbf{\Sigma}_{(j)}$ is assumed for $r_t \in \{i, j\}$, $i \neq j$. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Regime values of corn and phosphorus fertilizer prices, P^Y and P^F, for the moderate and high regimes, are the average values of the regime estimated from the Markov-switching Vector Autoregressive model results. Specifically, the moderate and high regime average prices for corn and phosphorus fertilizer are $P^Y_{\text{Moderate}} = \$1.432$, $P^Y_{\text{High}} = \$2.027$ per bu., and $P^F_{\text{Moderate}} = \$180.530$, $P^F_{\text{High}} = \$193.775$ per tn.

distributions are specified for the model parameters, including the autoregressive coefficients, intercepts, and covariance matrices. Specifically, we use the priors on the intercept $\boldsymbol{\mu}_{(r_{t+1})} \sim \mathcal{N}(\mu_\mu, \sigma^2_\mu)$, priors on the autoregressive coefficients $\mathbf{\Phi}_{(r_{t+1})} \sim \mathcal{N}(\mu_\Phi, \sigma^2_\Phi)$, and priors on the covariance matrix $\mathbf{\Sigma} \sim \text{Wishart}(I, \nu)$, where I is the identity matrix and ν is the degree of freedom.[7]

MSVAR results are presented in Table 3. The results show that in both regimes, the next-year corn price is significantly influenced by both the current corn price and the current P fertilizer price, as indicated by the significant coefficients for $\ln(P^Y_t)$ and $\ln(P^F_t)$. However, the next year's P fertilizer price is only influenced by its own current price, with no significant effect from the current corn price.

The asymmetry in price dynamics can be attributed to the differing market structures and

[7]The mean (μ_μ, μ_Φ) and variance (σ^2_μ, σ^2_Φ) of the prior distributions are derived from the estimation results of the Markov Switching Dynamic Regression (MSDR) model, which independently estimates the price process for the Markov-switching regimes (see Appendix).

roles of corn and P fertilizer. Corn, as a staple commodity, is more sensitive to input costs such as fertilizer, which directly affects production costs and, consequently, market prices. In both regimes, the significant effect of P fertilizer prices on future corn prices reflects the pass-through of input cost changes to agricultural output prices. On the other hand, the P fertilizer market is largely driven by supply-side factors, such as production costs and global demand for fertilizers, rather than fluctuations in corn prices. This explains why P fertilizer prices are only influenced by their own lagged values, with no significant impact from corn prices. Transition probabilities in Table 3 shows, for example, that the corn and fertilizer prices have a 73.5% likelihood of remaining at a moderate regime during the next period given that the process is moderate during the current period as well as a 26.5% likelihood of moving to a high regime.

3.3 Values for other model parameters

Values for the remaining model parameters not estimated above are calibrated based on the literature and expert consultation with extension colleagues. Table 4 shows the baseline point values for the first set of results we present in the next section. However, we varied these parameters in sensitivity analysis, and the output from that analysis reflects an important aspect of our findings. Here, we describe the baseline parameterization and the sources of information used.

Parameters for legacy P dynamics are primarily taken from the deterministic dynamic model of Ekholm et al. (2005). Because we introduce stochasticity into this model, we also require a value for the P dynamics carryover variance ς^2, which we set at $\varsigma^2 = 9.21$ based on calculations using North Carolina Tidewater region data – specifically using legacy P measurements from plots where no fertilizer was applied. Appendix Fig. F2 provides a visualization of the simulated legacy P dynamics with no fertilizer inputs: The mean of the simulated sample paths hews closely to the output from the deterministic model, but with significant stochasticity also visible from the plotted quartiles of the simulations and the shaded area reflecting their 90% range.

Values for soil sampling costs and precision were based on the following: Standard soil sampling typically involves collecting one soil sample per 1 hectare, costing around $4 per acre (NCAGR 2024). This assumption follows recommendations that soil samples should be taken from areas smaller than 20 acres to ensure accuracy and representativeness of the soil's nutrient levels (USDA 09.2022). Point sampling is recommended at a spacing of 209 feet, where composite samples are collected per acre, resulting in approximately 2.47 samples per hectare (Austin et al., 2020). Thus, point sampling provides more precise information on legacy P bioavailability but is a more

Table 4: Parameters and description

	Value	Desccription
Biological Parameters		
μ_ρ	-0.02	Average rate of growth (Myyrä et al. 2007)
ς^2	9.21	Carryover variance
γ_1	0.0032	Response parameter of legacy P surplus (Ekholm et al. 2005)
γ_2	0.00084	
γ_3	0.000186	Concentration parameter on crop yield (Iho and Laukkanen 2012, Saarela et al. 1995)
γ_4	0.003	
Economic Parameters		
c_{ss}	\$4	Standard soil sampling cost per hectare (NCAGR 2024)
c_p	\$9.88	Point soil sampling cost per hectare
β	0.9259	Discount factor with 8% discount rate (Duquette et al. 2012)
σ_{ss}	0.638	Observation error of standard soil sampling (Lauzon et al. 2005)
σ_p	0.104	Observation error of point soil sampling

Notes: Soil sampling cost varies depending on the institute. This paper uses the North Carolina case (NCAGR 2024, \$4 per sample).

expensive methodology to implement. Based on this information, we assumed that the observation error variance of point sampling (σ_p) was smaller than standard sampling ($\sigma_{ss} > \sigma_{ps}$), and the cost was 2.47 times higher than standard sampling ($c_{ps} = 2.47 \cdot c_{ss}$). The values for the standard deviation of observation errors of standard soil sampling, as shown in Table 4, are sourced from Lauzon et al. (2005). This study, based on data from 23 Ontario, Canada farm fields, provides the coefficient of variation for soil testing, reflecting variability in soil P measurements. For this work, we averaged these values and used it for the observation error because our observation variable is in a multiplicative form, combining legacy P bioavailability with observation error. For the point sampling standard deviation, we calculate it by dividing the standard deviation of standard soil sampling by the square of the number of samples collected through point sampling ($\sigma_p = \sigma_{ss}/2.47^2$). This assumption is motivated by a simple Central Limit Theorem application, whereby the variance of the mean sampled legacy P is proportional to the number of independent samples taken.

4 Model Results

This section presents and discusses solutions from the model using the baseline parameterization in Table 4. We first present results for the ENPV-maximizing farmer, before presenting results for

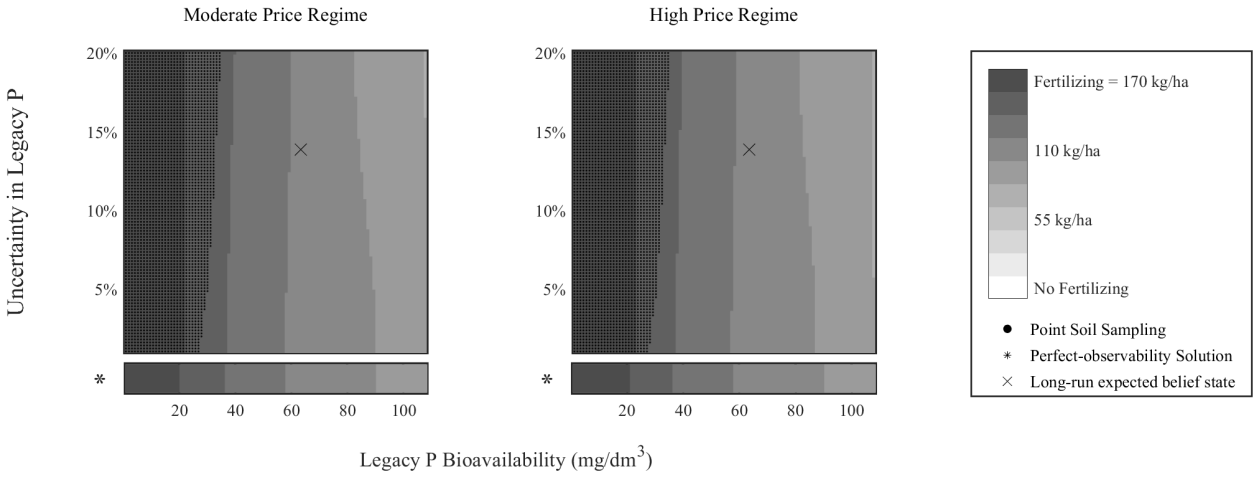
the cases with risk aversion and intertemporal smoothing.

4.1 Optimal Policy and Dynamics of Legacy Phosphorus

Figure 2 presents the agent's optimal actions across two graphs, each corresponding to a different current fertilizer price regime. The horizontal axis on each graph measures mean estimated legacy P bioavailability (mg/dm^3) within a range of 1 to 108 mg/dm^3, which captures around 90% of the legacy P measurements in the Eastern NC field trial data. The vertical axis represents uncertainty, as measured by the coefficient of variation (CV) in L beliefs, from 1% to 20% of the mean estimate. The bars at the bottom of each graph show the solution to the dynamic programming model in which legacy P is perfectly observed (obviating the need for soil testing). The $\times$ marks the long-run expected belief state, which are the average values to which the legacy P distribution mean and CV converge as $T \to \infty$, where the average is taken over a large number (10,000) of stochastic simulations.

From these graphs, we see several general features: First, optimal P fertilizer inputs monotonically decrease with higher estimated (mean) legacy P. Second, point sampling is employed only at lower levels of estimated legacy P; if estimated legacy P is high enough, the agent falls back to standard sampling, with this threshold varying with the degree of uncertainty (CV). Intuitively, the results show that more uncertainty implies more use of point sampling. Yet we also see that P fertilizer inputs are relatively insensitive to the degree of uncertainty in legacy P, and that there is a nonmonotonic relationship between legacy P uncertainty and optimal fertilizer use. In different regions of the state space, fertilizer sometimes increases or decreases with higher CV, but not by very much in either direction – at least relative to the larger, monotonic response we see with respect to the mean. We also see little variation in optimal fertilizer or sampling with respect to the fertilizer price regime the agent is currently in. These general features are consistent across all the sensitivity analyses we conducted (discussed in the next section). That optimal fertilizer use strongly and monotonically decreases with estimated legacy P is clearly intuitive, though care is required to understand the dynamic relationship here. At first glance, it may appear that P fertilizer and legacy P simply act as substitutes in the model. Yet fertilizer not only directly affects crop production but also replenishes soil P stocks, and thus acts as a form of capital investment. These substitution and capitalization effects reinforce one another: Because of diminishing marginal productivity in legacy P (estimated from the yield regression in Table 2), there are greater incentives to build up soil stocks when legacy P is low, when the marginal benefits of legacy P are highest.

Figure 2: Optimal policy of P fertilizer application and soil sampling



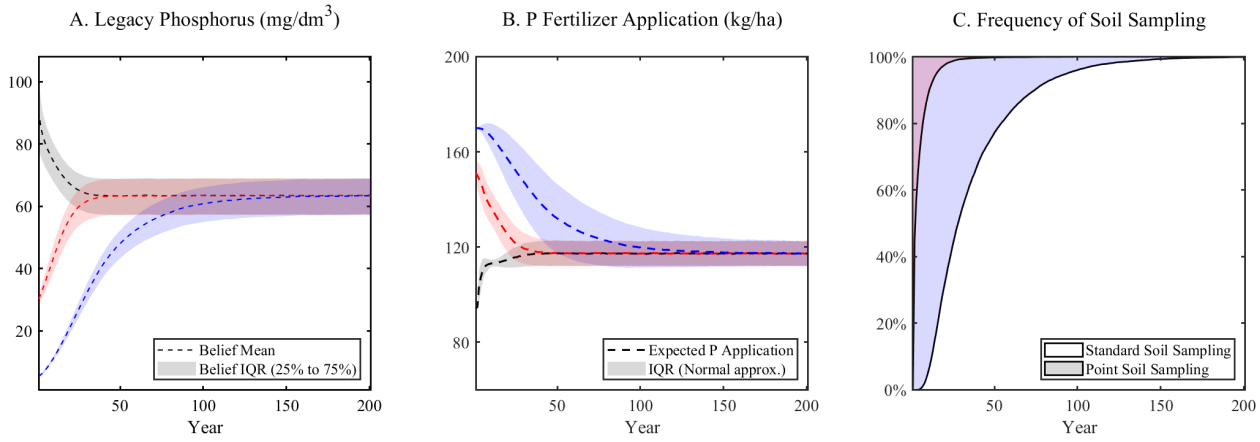
These incentives decrease as legacy P accumulates, which also decreases the marginal productivity of fertilizer on yields (given the negative interaction term in the yield regression). This substitution effect further disincentivizes fertilizer use at high legacy P levels.

Figure 3 shows clearly how these dynamic interaction effects between fertilizer use and legacy P stocks play out. This figure shows summary results from 10,000 simulations of the model under the optimal policy (Fig. 2) and three different initial conditions for legacy P, with estimated legacy P plotted over time in Panel A, fertilizer use in Panel B, and soil sampling actions in Panel C. Initial legacy low P stocks correspond to higher initial fertilizer use, with legacy P subsequently accumulating, and fertilizer use converging to a lower (but still positive) stable equilibrium.

Figures 2 and 3 clearly show that it is primarily low estimated legacy P–not the degree of uncertainty in the estimate–that drives fertilizer use. That optimal fertilizer use varies little with the degree of uncertainty, while at the same time the agent also finds it optimal to use point sampling (the only value of which is to reduce uncertainty) presents a potential paradox: Why would the agent pay for monitoring when reduced uncertainty does not appear to directly affect production decisions? This pattern is illustrated in Figure 3, Panel B, where fertilizer trajectories across uncertainty levels converge to similar paths and the variability, as indicated by the interquartile ranges, remains limited. Indeed, when we compare optimal fertilizer use to the perfect observability solution at the bottom of each graph, we see little difference.

This paradox can be resolved by recognizing that more accurate sampling indirectly affects fertilizer use, through its effect on estimated legacy P. Mean legacy P is a stochastic state variable, with its variability determined by soil test accuracy. Thus, more accurate testing reduces variability

Figure 3: Dynamics simulation of stochastic growth



Notes: The initial values for uncertainty (belief state coefficient of variation) is 20% for three cases. The initial conditions also include a moderate price regime. The figures were generated from simulations $i = 10,000$. For Panel A, the Belief IQR for a simulation i at time t is calculated as follows: $\text{IQR}_{it} = \left\{ \exp(\mu^L_{i,t} + \sigma^L_{i,t}\Phi^{-1}(p)) \mid p \in [0.25, 0.75] \right\}$ where $\mu^L_{i,t}$ and $\sigma^L_{i,t}$ are the parameters of the belief state, and $\Phi^{-1}(\cdot)$ represents the inverse cumulative distribution function of the standard normal distribution (Kling et al. 2017). In this figure, we averaged Belief IQR_{it} over i. For Panel B, the shaded areas represent a normal approximation to the IQR, calculated as $\mu^F_t \pm \sigma^F_t \Phi^{-1}(p)$ from simulation F_{it}, which corresponds to the 25th and 75th percentiles under the normal distribution.

in estimated legacy P, which then translates into reduced variability in optimal fertilizer use (which responds strongly to estimated legacy P).

In this baseline parameter scenario, we also see that in the long-run the agent always falls back to standard soil sampling (which is confirmed in simulations, as shown in Fig. 3, Panel C): This particular result is sensitive to the parameter values, as we shall see in the next section. Under other conditions, the agent's use of point sampling persists indefinitely.

4.2 Risk Analysis: Epstein-Zin Preferences

To compute solutions to the MOMDP with recursive Epstein-Zin (EZ) utility in (8) and (9), we require value ranges for the coefficient of relative risk aversion (CRRA, η) and the elasticity of intertemporal substitution (IES, ψ). Rather than analyzing the solution for one set of values for η and ψ, we are interested in how model results change across a reasonable range for these parameters, i.e., how does greater risk aversion or intertemporal smoothing affect optimal fertilizer applications and soil sampling? Because risk and intertemporal preferences are highly heterogeneous, we specified ranges for these parameters referring to the literature on environmental and agricultural studies listed in Table F1 in the Appendix. From this literature, we settled on $\eta \in \{0, 15\}$ and

Table 5: Fertilization and Sampling Strategies Corresponding to the Long-Run Stable Distribution, by Belief State and Farmer Preference Type

State		**Farmers Preference Types** (Risk Preference & Intertemporal Smoothing)			
Legacy P Bioavailability	Uncertainty in Legacy P	CRRA ($\eta = 0$) IES ($\psi = \infty$)	CRRA ($\eta = 15$) IES ($\psi = \infty$)	CRRA ($\eta = 0$) IES ($\psi = 1.5$)	CRRA ($\eta = 15$) IES ($\psi = 1.5$)
High	High	94 kg/ha [S]	113 kg/ha [S]	94 kg/ha [S]	113 kg/ha [S]
	Medium	113 kg/ha [S]	113 kg/ha [S]	113 kg/ha [S]	113 kg/ha [S]
	Low	113 kg/ha [S]	113 kg/ha [S]	113 kg/ha [S]	113 kg/ha [S]
Medium	High	151 kg/ha [P]	151 kg/ha [P]	151 kg/ha [P]	151 kg/ha [P]
	Medium	151 kg/ha [S]	151 kg/ha [S]	151 kg/ha [S]	151 kg/ha [S]
	Low	151 kg/ha [S]	151 kg/ha [S]	151 kg/ha [S]	151 kg/ha [S]
Low	High	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]
	Medium	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]
	Low	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]

Notes: Price regime is in high regime. Sampling strategies are in brackets. S and P indicate the standard sampling and point sampling, respectively. We select low, median, and high legacy P bioavailability as 1, 35, and 85 mg/dm^3, respectively, corresponding to the minimum and about the 30th and 80th percentiles of the legacy P bioavailability range. For uncertainty in legacy P bioavailability, we use CV values of 1%, 10%, and 20% to represent low, median, and high uncertainty, respectively.

$\psi \in \{1.5, \infty\}$. These ranges include the risk-neutrality and perfectly elasticicity IES cases ($\eta = 0$ and $\psi = \infty$), but also allow for very high levels of risk aversion and strong preference for intertemporal smoothing: We allow for these more extreme values for experimental purposes, to see the extent of variation in model output with respect to changes in these parameters.

Table 5 summarizes the agents' optimal actions under EZ preferences at selected points across the belief state space for high, medium, and low levels of mean legacy P and its CV. Only at high levels of estimated legacy P *and* uncertainty in this estimate do we see any variation in the actions with respect to risk aversion (η): More risk-averse optimizing agents apply more fertilizer in this case, but still do not adopt point sampling. We see no variation in optimal actions with respect to preferences for intertemporal smoothing (ψ).

While this evident lack of influence from these parameters surprised us–particularly that more

risk averse agents do not seek more intensive soil testing to reduce uncertainty–these results do offer some insight for ruling out potential explanations for P fertilizer over-application. One of the motivations for this work was to study why US farmers are not more willing to rely on their existing legacy P soil stocks, instead of applying more P fertilizer. One reason put forward has been farmer risk aversion, with more precise soil testing offered as a potential solution. However, these results suggest that risk aversion is unlikely to explain why legacy soil P is not more heavily utilized: To the extent that it does affect behavior in the model, it leads farmers to simply apply more fertilizer–likely as an easy hedge against uncertainty soil stocks–rather than opting for better soil testing to reduce this uncertainty.[8]

5 Sensitivity Analysis

Here we present several results from sensitivity analysis with respect to both the biological and economic sets of parameters in the model. We begin with results from varying the biological parameters.

5.1 Biophysical Parameters Variations

Here, we assess the impact of varying three key parameters from the biophysical model in eqs. (1) and (2) we deemed to be important in influencing agent actions in the model:[9] (i) the average decay rate of legacy P (μ_ρ), (ii) the efficiency of legacy P accumulation from surplus fertilizer (γ_1), and (iii) the productivity of legacy P in contributing to yield via P concentration in soil solution (γ_4). These parameters govern how soil P changes over time due to natural variation and crop cultivation.

Table 6 summarizes the optimal fertilizer rate (kg/ha) and sampling strategy (standard or point sampling) across representative belief states under each parameter scenario. For example, when we reduce the persistence of legacy P by setting $\mu_\rho = -0.08$ (relative to the benchmark case, $\mu_\rho = -0.02$), we observe a systematic reduction in fertilizer application, particularly in medium- and high-bioavailability regions of the state space. This outcome may seem counterintuitive if one views legacy P and fertilizer only as pure substitutes -one might expect that if legacy P decays faster, farmers would need to apply more fertilizer to compensate. However, viewed dynamically, P fertilizer is not just a yield input but also a forward-looking investment in soil P. All else equal, a

[8] Appendix Figures F5 and F6 present the optimal P fertilizer application and soil sampling outcomes across the entire state space defined by bioavailable legacy P and uncertainty, under Epstein–Zin preferences.

[9] We thank an anonymous reviewer for feedback on an earlier version of this paper, which prompted our investigation of these specific parameters.

Table 6: Fertilization and Sampling Strategy by Biophysical Parameters

State		**Biophysical Parameters**			
Legacy P Bioavailability	Unvertainty in Legacy P	Risk Neutral (Benchmark)	Decay Process ($\mu = -0.08$)	P Accumulation ($3 \times \gamma_1$)	P Concentration ($3 \times \gamma_4$)
High	High	94 kg/ha [S]	94 kg/ha [S]	94 kg/ha [S]	94 kg/ha [P]
	Medium	113 kg/ha [S]	94 kg/ha [S]	113 kg/ha [S]	75 kg/ha [P]
	Low	113 kg/ha [S]	94 kg/ha [S]	113 kg/ha [S]	75 kg/ha [S]
Medium	High	151 kg/ha [P]	132 kg/ha [P]	151 kg/ha [S]	132 kg/ha [P]
	Medium	151 kg/ha [S]	132 kg/ha [P]	151 kg/ha [S]	113 kg/ha [P]
	Low	151 kg/ha [S]	132 kg/ha [P]	151 kg/ha [S]	113 kg/ha [P]
Low	High	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]
	Medium	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]
	Low	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]

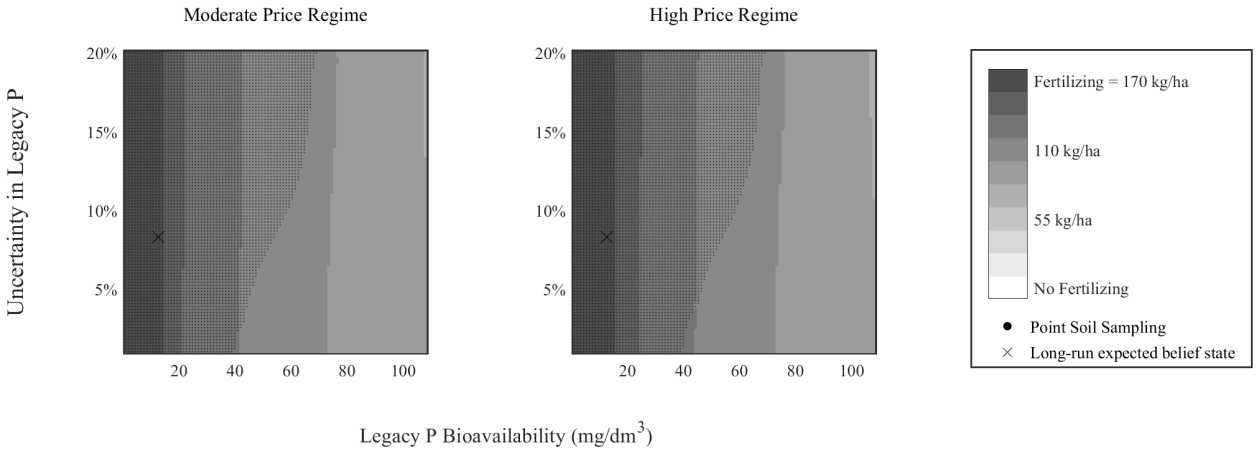
Notes: Price regime is in high regime. Sampling strategies are in brackets. S and P indicate the standard sampling and point sampling, respectively.

faster decay of soil P equates to a higher depreciation rate in this capital stock, disincentivizing investment in it.

Fig. 4 illustrates this behavior. Compared to the benchmark case in Fig. 2, the graphical results confirm low persistence decreases optimal fertilizer applications. The figure also shows that point sampling is utilized over a larger area of the state space than in the higher-persistence benchmark scenario. This pattern is also intuitive: With lower persistence and a carryover variance held fixed, soil P in each period behaves more like an unobserved random productivity shock, in which case farmers see increased value from current, accurate information to inform their fertilizer decisions.

We also explore how changes in the accumulation efficiency parameter γ_1 affect agent behavior in the model. An increase in γ_1 implies that a larger portion of P surplus is retained in the soil, increasing the legacy P stock. This enhances the value of investing in soil P and the long-term payoff of fertilization. As shown in Table 6, when soil P accumulates more readily point sampling is less utilized. This is due to greater accumulation efficiency also leading the agent to have relatively more certainty about next-period changes to soil P, based on their current P fertilizer inputs net of

Figure 4: Sensitivity analysis of legacy P decay process



crop utilization.

Finally, we examine the impact of increasing γ_4, which controls how strongly crop growth demands deplete legacy P stocks. Qualitatively, the effect of this parameter is the same as that observed from changes in a decreased carryover parameter (μ_ρ) discussed above. The intuition is clear: This parameter determines a crop-induced form of depreciation in legacy P capital stocks. For reasons already discussed above, higher depreciation leads to lower investment (less P fertilizer input) and more intensive monitoring over larger portions of the state space.[10]

5.2 Economic Sensitivity Analysis

The preceding biophysical sensitivity analysis demonstrated how changes in soil nutrient dynamics affect fertilizer use and monitoring behavior. However, farmers' decisions are also shaped by economic preferences and policy environments. In this section, we examine how economic factors -including time preferences and policy incentives- alter optimal P management strategies under uncertainty.

To summarize results in a structured and interpretable manner, Table 7 presents optimal fertilizer application rates and sampling strategies across representative belief states and policy settings. These include two levels of the discount rate (8% benchmark and 28%) (Duquette et al., 2012), two fertilizer (*ad valorem*) tax rates (50% and 500%), and two sampling subsidy levels (15% and 50%).

The table reveals consistent patterns in how farmers respond to these economic levers. First, in discount rates, we selected three discount rates -8% and 28%- to reflect a range of scenarios that

[10]We provide more detailed results from the sensitivity analyses with respect to γ_1 and γ_4 in the Appendix Fig. F7 and F8.

Table 7: Fertilization and Sampling Strategy by Economic Parameters

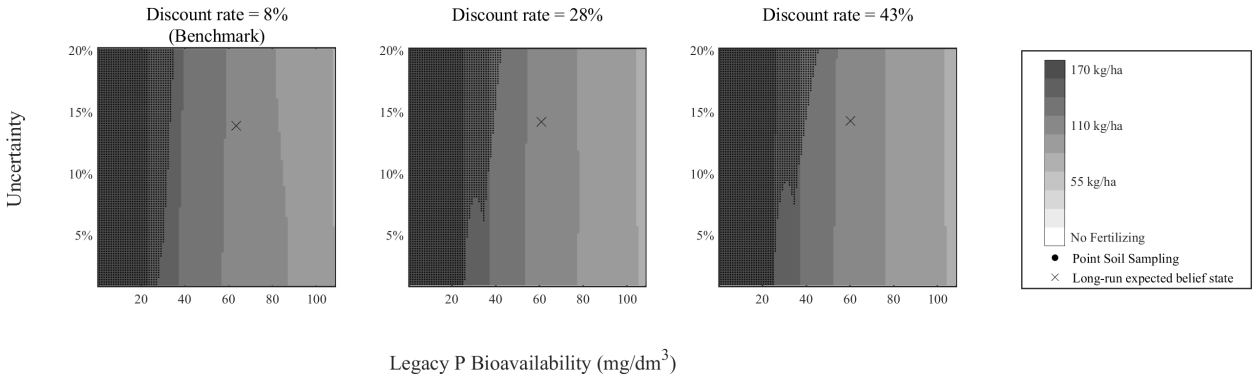
Legacy P Bioavailability	Unvertainty in Legacy P	**Discount Rate**		**Fertilizer Tax**		**Sampling Susidy**	
		Benchmark (8%)	High (28%)	Low (50%)	High (300%)	Low (15%)	High (50%)
High	High	94 kg/ha [S]	94 kg/ha [S]	94 kg/ha [S]	94 kg/ha [S]	94 kg/ha [S]	94 kg/ha [S]
	Medium	113 kg/ha [S]	94 kg/ha [S]	94 kg/ha [S]	94 kg/ha [S]	94 kg/ha [S]	94 kg/ha [S]
	Low	113 kg/ha [S]	94 kg/ha [S]	113 kg/ha [S]	94 kg/ha [S]	113 kg/ha [S]	113 kg/ha [S]
Medium	High	151 kg/ha [P]	151 kg/ha [P]	151 kg/ha [P]	151 kg/ha [S]	151 kg/ha [P]	151 kg/ha [P]
	Medium	151 kg/ha [S]	151 kg/ha [P]	151 kg/ha [S]	151 kg/ha [S]	151 kg/ha [S]	151 kg/ha [S]
	Low	151 kg/ha [S]	151 kg/ha [S]	151 kg/ha [S]	151 kg/ha [S]	151 kg/ha [S]	151 kg/ha [S]
Low	High	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]
	Medium	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]
	Low	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]	170 kg/ha [P]

Notes: Price regime is in high regime. Sampling strategies are in brackets. S and P indicate the standard sampling and point sampling, respectively.

align with both economic theory and empirical findings. While the 8% is the benchmark rate used in the model - which is within the realm of rates implied by financial markets as well as used by the US Federal government for economic analyses - we examine the effect of higher rates which are closer to those empirically measured among US farmers (Duquette et al., 2012). Farmer time preferences are heterogeneous, so it relevant to examine how model results change across this whole range.

Intuitively, simple, textbook optimal control models from capital investment theory show that discount rates affect optimal investment decisions in largely the same way that capital depreciation rates do. As shown in Table 7, this is what we observe here: Higher discount rates, as with higher depreciation rates in legacy P stocks, are associated with greater adoption of enhanced soil testing through point sampling, while fertilizer use is reduced. These shifts in behavior are visualized more clearly in Fig. 5, which shows how soil sampling and fertilization strategies vary as farmers become more myopic. While intuitive from a capital-theoretic perspective, it is worth emphasizing this result may strike other audiences as counterintuitive: Our model suggests that farmers more focused on short-v-longer term payoffs apply *less* fertilizer and engage in *more* costly and intensive monitoring.

Figure 5: Sensitivity analysis of discount rate



Notes: Price regime is in high regime.

Next, we investigate how a sustained P fertilizer price increase affects optimal P fertilizer use and soil testing in the model. Such price increases could result from tighter supply constraints (e.g. given the concentration of nonrenewable mineral P stocks in a handful of locations). Price increases may also result from policy: For example, fertilizer taxes (or, at least, removal of fertilizer subsidies) have been periodically proposed over the years as incentive-based instruments for addressing non-point source water pollution caused by fertilizer use (Sud 2020, Lungarska and Jayet 2018, Sun et al. 2016, Osteen and Kuchler 1986, Liang et al. 1998). However, the effectiveness of taxation on agricultural chemicals in reducing chemical fertilization is unclear due in part to price-inelastic fertilizer demand among farmers (e.g. Liang et al. 1998, Brunelle et al. 2015, Nasrin et al. 2022, Al Rawashdeh 2023). In a general equilibrium model, Liang et al. (1998) examined the effect of taxation on P and nitrogen on fertilizer use through two tax schemes, namely uniform and differentiated taxes. They found that a 500% tax reduced only on-farm fertilizer usage by only 8%. However, in an econometric study of fertilizer demand in Illinois and Indiana, Sun et al. (2016) found that, despite the confirmation of inelastic fertilizer demand, a fertilizer tax would be more cost-effective than output-based policies such as reducing crop subsidies at achieving nutrient reduction targets. Sud (2020) reviews a variety of policies implemented in several OECD countries for addressing fertilizer pollution from agriculture: Taxes are used infrequently, likely due to political barriers.

Here, we model increases to P fertilizer prices ranging from 50% to 300%.[11] For exposition, we label these increases as (*ad valorem*) taxes, and modify the fertilizer price state variable as follows:

[11]The entire state space of optimal P fertilizer applications and soil sampling results with other tax levels is presented in the Appendix Fig F9.

$P^F_{\text{tax}} = P^F \cdot (1 + \textit{Tax Rate})$, where P^F is the producer price, and P^F_{tax} denotes the price of P fertilizer paid by farmers. As shown in Table 7, how alternative price scenarios affect model results for the baseline case of a risk-neutral farmer without a preference for intertemporal smoothing. Only in the case of high estimated legacy P and low uncertainty in that estimate do we see any fertilizer reductions in response to the tax. Interestingly, enhanced soil testing remains relatively unchanged across different price scenarios. This suggests that while fertilizer taxes may influence the overall level of fertilizer application, they do not significantly alter the perceived value of soil testing.

Lastly, as discussed in the Introduction, part of the applied research interest in legacy P on agricultural lands lies in the potential for reducing fertilizer use. The line of thinking is that providing farmers more information about these soil stocks could reduce fertilizer use. In our model, such a hypothesis can be examined through a sensitivity analysis with respect to the costs of enhanced soil testing, e.g., from a hypothetical subsidy for enhanced soil testing. Here, we analyze how subsidies for enhanced soil testing ranging from 15% to 50% affect model output.

Contrary to the hypothesis that more information on soil P can enable lower fertilizer use, we find that subsidies for enhanced soil testing have essentially no effect on optimal fertilizer use. This nonresponsiveness of fertilizer use to soil test subsidies persists despite increased use of point sampling across portions of the belief state space. These point sampling increases, while not evident in the limited cases shown in Table 7, can be seen in a full state space visualization in Appendix Figure F10. Thus, while farmer agents in the model do see value from cheaper point sampling, the improved information they have about their soil P does not translate into appreciable changes to their fertilizer decisions. Additionally, in robustness checks in the Appendix, we confirm that the presence of risk aversion does not alter these findings.

6 Discussion

The overuse of P fertilizer in agriculture causes significant surface water pollution, necessitating policy solutions that encourage farmers to use less P fertilizer while minimizing economic losses in agricultural production. However, uncertain soil P dynamics create complex tradeoffs for farmers to consider in their decisions to continue to use P fertilizer or to try to rely more on legacy soil stocks, especially when there is heterogeneity and uncertainty in these stocks within fields. This research sheds light on these incentives using a model-based approach to study farmers' optimal, forward-looking fertilizer demand and soil sampling choices accounting for both soil P and price dynamics. We apply methods developed for resource management problems involving the partial

observability of resource stocks and advance these methods by including agent risk aversion, as well as by introducing econometric techniques for reducing the dimensionality of price dynamics.

The focus of this modeling effort is primarily on understanding how farmers' fertilizer use decisions may respond when optimally accounting for imperfectly observed legacy P stocks. In particular, we have aimed in this paper to illustrate how these potential responses vary systematically with farmers' values and preferences (over risk and time) and with the economic and policy environment (in particular, as reflected through fertilizer prices and costs for precision soil testing). Our findings provide important insights into why farmers may not fully exploit legacy P stocks, and why they may not respond to potential informational interventions (e.g. subsidies for enhanced soil testing) in the manner hoped for by policymakers and extension professionals.

Importantly, our results also suggest that both fertilizer and soil sampling decisions are relatively price inelastic. This finding holds practical significance, implying that conventional policy levers such as fertilizer taxes may be insufficient to meaningfully alter fertilizer application behavior. This result is consistent with a broader empirical literature documenting inelastic P fertilizer demand and reinforces the need for more nuanced or targeted policy approaches (Vos et al., 2025).

Our results also show that a capital-theoretic perspective can be useful for policymakers and managers to better understand and anticipate farmers' decisions in this context. For example, managers in this context may assume that high levels of fertilizer use and seeming disregard for more accurate soil test information might be the product of farmers' short-term thinking. However, results from our sensitivity analysis regarding the discount rate show that agents who weigh short-term payoffs more heavily in fact use less P fertilizer and do more costly soil testing. This finding can be readily understood once P fertilizer use is recognized as a form of investment in soil P stocks.

While we have shown that the eastern NC context we study is broadly representative of typical US corn-farming systems, we acknowledge that these results are confined to a particular case study application and may not be fully generalizable across other cropping systems and location-specific contexts. Future research can extend the POMDP methods developed herein to explore how behavioral and biophysical factors affect legacy P stock dynamics and adoption of soil sampling technologies in other contexts - particularly for production systems with lower substitutability between bioavailable legacy P and synthetic fertilizers.

In our analysis, we did not take up the economic assessment of environmental impacts of P fertilizer use in agriculture (see Kling et al. 2014, Rabotyagov et al. 2014a). Nor do we try to

directly incorporate these externalities into farmers' decision problems. However, the present modeling is relevant for addressing such externalities, because it illustrates the potential dynamic responses in farmer fertilizer demand when accumulated soil P stocks and monitoring are taken into account. This focus on farmer behavior can be extended in future research to incorporate environmental factors explicitly. For instance, expanding the model to consider the environmental and climate change implications of P management can provide a more comprehensive grasp of the overall impact of agricultural practices. Future studies can integrate spatial variability and explore interactions among farmers sharing a common watershed, as well as between farmland and adjacent areas (e.g., under urbanization). Thus, the farm-level model in this paper could be viewed as a component in future larger-scale models that address these aspects.

References

AL RAWASHDEH, R. (2023): "Estimating short-run (SR) and long-run (LR) demand elasticities of phosphate," *Mineral Economics*, 36, 239–253.

ARROW, K., P. DASGUPTA, L. GOULDER, G. DAILY, P. EHRLICH, G. HEAL, S. LEVIN, K.-G. MÄLER, S. SCHNEIDER, D. STARRETT, AND B. WALKER (2018): "Are we consuming too much?" *Journal of Economic Perspectives*, 11(1), 17.

AUGERAUD-VÉRON, E., G. FABBRI, AND K. SCHUBERT (2019): "The value of biodiversity as an insurance device," *American Journal of Agricultural Econoics*, 101(4), 1068–1081.

AUSTIN, R., L. GATIBONI, AND J. HAVLIN (2020): "Soil sampling strategies for site-specific field management," *NC State Extension AG-439-36*.

BEAULIEU, J. J., T. DELSONTRO, AND J. A. DOWNING (2019): "Eutrophication will increase methane emissions from lakes and impoundments during the 21st century," *Nature communications*, 10, 1375.

BRUNELLE, T., P. DUMAS, F. SOUTY, B. DORIN, AND F. NADAUD (2015): "Evaluating the impact of rising fertilizer prices on crop yields," *Agricultural economics*, 46, 653–666.

CAI, Y. AND T. S. LONTZEK (2019): "The social cost of carbon with economic and climate risks," *Journal of Political Economy*, 127(6), 2684–2734.

CHADES, I., J. CARWARDINE, T. MARTIN, S. NICOL, R. SABBADIN, AND O. BUFFET (2012): "MOMDPs: a solution for modelling adaptive management problems," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 26, 267–273.

CHADèS, I., T. G. MARTIN, S. NICOL, M. A. BURGMAN, H. P. POSSINGHAM, AND Y. M. BUCKLEY (2011): "General rules for managing and surveying networks of pests, diseases, and endangered species," *Proceedings of the National Academy of Sciences*, 108(20), 8323–8328.

CHEN, L. (2022): "Essays on the Economics of Soil Health Practices," *North Carolina State University*.

CLARK, C. W. (2010): "General rules for managing and surveying networks of pests, diseases, and endangered species," *Mathematical Bioeconomics: The Mathematics of Conservation. Wiley*.

CONLEY, D. J., H. W. PAERL, R. W. HOWARTH, D. F. BOESCH, S. P. SEITZINGER, K. E. HAVENS, C. LANCELOT, AND G. E. LIKENS (2009): "Controlling eutrophication: Nitrogen and phosphorus," *Science*, 323(5917), 1014–1015.

DENBALY, M. AND H. VROOMEN (1993): "Dynamic fertilizer nutrient demands for corn: A cointegrated and error-correcting system. American Journal of Agricultural Economics," *American Journal of Agricultural Economics*, 75(1), 203–209.

DIXIT, A. K. AND R. S. PINDYCK (1994): *Investment under uncertainty*, Princeton university press.

DOWNING, J. A., S. POLASKY, S. M. OLMSTEAD, AND S. C. NEWBOLD (2021): "Protecting local water quality has global benefits," *Nature Communications*, 12(1), 2709.

DUQUETTE, E., N. HIGGINS, AND J. HOROWITZ (2012): "Farmer discount rates: Experimental evidence," *American Journal of Agricultural Economics*, 94(2), 451–456.

EKHOLM, P., E. TURTOLA, J. GRÖNROOS, P. SEURI, AND K. YLIVAINIO (2005): "Phosphorus loss from different farming systems estimated from soil surface phosphorus balance," *Agriculture, Ecosystems and*

Environment, 110(3-4), 266–278.

EPSTEIN, L. G. AND ZIN (1991): "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis," *Journal of Political Economy*, 99(2), 261–286.

EPSTEIN, L. G. AND S. E. ZIN (1989): "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework," *Econometrica*, 57(4), 937.

FACKLER, P. AND K. PACIFICI (2014): "Addressing structural and observational uncertainty in resource management," *Journal of Environmental Management*, 133, 27–36.

FACKLER, P. L. (2014): "Structural and observational uncertainty in environmental and natural resource management," *International Review of Environmental and Resource Economics*, 7(2), 109–139.

FAN, X., M. I. GèMEZ, S. S. ATALLAH, AND J. M. CONRAD (2020): "A Bayesian state-space approach for invasive species management: the case of spotted wing drosophila," *American Journal of Agricultural Economics*, 102(4), 1227–1244.

FEDERAL RESERVE BANK OF MINNEAPOLIS (2024.04): "Consumer Price Index, 1913-," *https://www.minneapolisfed.org/about-us/monetary-policy/inflation-calculator/consumer-price-index-1913-* .

GATIBONI, L., D. OSMOND, AND D. HARDY (2022): "Changes in the Phosphorus Fertilizer Recommendations for Corn, Soybean, and Small Grains in North Carolina," *NC State Extension AG-439-30*.

GILBERT, R. J. (1979): "Optimal depletion of an uncertain stock," *The Review of Economic Studies*, 46(1), 47.

GOETZ, R. U. AND D. ZILBERMAN (2000): "The dynamics of spatial pollution: The case of phosphorus runoff from agricultural land," *Journal of Economic Dynamics and Control*, 24(1), 143–163.

HAIGHT, R. G. AND S. POLASKY (2010): "Optimal control of an invasive species with imperfect information about the level of infestation," *Resource and Energy Economics*, 32(4), 519–533.

HAMILTON, J. D. (1989): "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica*, 57(2), 357.

IHO, A. AND M. LAUKKANEN (2012): "Precision phosphorus management and agricultural phosphorus loading," *Ecological Economics*, 77, 91–102.

INNES, R. (2000): "The economics of Livestock Waste and its regulation," *American Journal of Agricultural Economics*, 82(1), 97–117.

KILE, L. K., L. GATIBONI, D. L. OSMOND, A.-M. MARSHALL, A. JOHNSON, AND O. W. DUCKWORTH (2025): "Why Does Overapplication of Phosphorus Fertilizers Occur: Insights from North Carolina Farmers," *Agriculture*, 15, 606.

KLING, C. L., Y. PANAGOPOULOS, S. S. RABOTYAGOV, A. M. VALCU, P. W. GASSMAN, T. CAMPBELL, AND M. J. W. ET AL. (2014): "LUMINATE: linking agricultural land use, local water quality and Gulf of Mexico hypoxia," *European Review of Agricultural Economics*, 41(3), 431–459.

KLING, D. M., J. N. SANCHIRICO, AND P. L. FACKLER (2017): "Optimal Monitoring and control under state uncertainty: Application to lionfish management," *Journal of Environmental Economics and Management*, 84, 223–245.

LAKE ERIE LAMP (2011): "Lake Erie Binational Nutrient Management Strategy: Protecting Lake Erie by

Managing Phosphorus," *Prepared by the Lake Erie LaMP Work Group Nutrient Management Task Group*.

LAUZON, J. D., I. P. O'HALLORAN, D. J. FALLOW, A. P. VON BERTOLDI, AND D. ASPINALL (2005): "Spatial variability of soil test phosphorus, potassium, and pH of Ontario soils," *Agronomy Journal*, 97(2), 524–532.

LIANG, C. L. K., S. B. LOVEJOY, AND J. G. LEE (1998): ""Green Taxes": Impacts On National Income, Social Welfare, And Environmental Quality," *Working Paper*.

LIVINGSTON, M., M. J. ROBERTS, AND Y. ZHANG (2015): "Optimal sequential plantings of corn and soybeans under price uncertainty," *American Journal of Agricultural Economics*, 97(3), 855–878.

LÖTJÖNEN, S., E. TEMMES, AND M. OLLIKAINEN (2020): "Dairy farm management when nutrient runoff and climate emissions count," *American Journal of Agricultural Economics*, 102(3), 960–981.

LOURY, G. C. (1978): "The optimal exploitation of an unknown reserve," *The Review of Economic Studies*, 45(3), 621–638.

LUNGARSKA, A. AND P. A. JAYET (2018): "Impact of spatial differentiation of nitrogen taxes on French farms' compliance costs," *Environmental and Resource Economics*, 69, 1–21.

LYBBERT, T. J. AND J. MCPEAK (2012): "Risk and intertemporal substitution: Livestock portfolios and off-take among Kenyan pastoralists," *Journal of Development Economics*, 97(2), 415–426.

MACLACHLAN, M. J., M. R. SPRINGBORN, AND P. L. FACKLER (2017): "Learning about a moving target in resource management: optimal Bayesian disease control," *Journal of Environmental Economics and Management*, 99(1), 140–161.

MCDONALD, R. AND D. SIEGEL (1986): "The value of waiting to invest," *The quarterly journal of economics*, 101, 707–727.

MELBOURNE, B. A. AND A. HASTINGS (2008): "Extinction risk depends strongly on factors contributing to stochasticity," *Nature*, 454(7200), 100–103.

MEMARZADEH, M., G. L. BRITTEN, B. WORM, AND C. BOETTIGER (2019): "Rebuilding global fisheries under uncertainty," *Proceedings of the National Academy of Sciences*, 116, 15985–15990.

MORALES, N. A., L. GATIBONI, D. OSMOND, R. VANN, S. KULESZA, C. CROZIER, AND D. HARDY (2023): "Critical soil test values of phosphorus and potassium for soybean and corn in three long-term trials in North Carolina," *Soil Science Society of America Journal*, 87(2), 278–290.

MYYRÄ, S., K. PIETOLA, AND M. YLI-HALLA (2007): "Exploring long-term land improvements under land tenure insecurity," *Agricultural Systems*, 92(1-3), 63–75.

NASRIN, M., P. VORTIA, S. SALAM, AND M. S. PALASH (2022): "Is fertilizer demand elastic to its own price? Assessing the consequences of fertilizer subsidy policy in Bangladesh," *SN Business & Economics*, 2, 110.

NCAGR (2024): "Agronomic Services - Soil Testing: Tests and Fees- Homeowners." *North Carolina Department of Agricultura & Consumer Service, Retrieved from https://www.ncagr.gov/divisions/agronomic-services/soil-testing/homeowners/tests-and-fees*, accessed October 11, 2024.

OSMUNDSEN, K. K., T. S. KLEPPE, AND A. OGLEND (2021): "MCMC for Markov-switching models—Gibbs sampling vs. marginalized likelihood," *Communications in Statistics-Simulation and Computation*, 50(3), 669–690.

OSTEEN, C. AND F. KUCHLER (1986): "Potential Bans of Corn and Soybean Pesticides: Economic Implications

for Farmers and Consumers. AER-546," *Economics Research Service*, U.S.Department of Agriculture.

PAUDEL, J. AND C. L. CRAGO (2020): "Environmental externalities from agriculture: Evidence from water quality in the United States," *American Journal of Agricultural Economics*, 103(1), 185–210.

PINDYCK, R. S. (1993): "Investments of uncertain cost," *Journal of financial Economics*, 34, 53–76.

PORTA, J. M., N. VLASSIS, M. T. SPAAN, AND P. POUPART (2006): "Point-based value iteration for continuous POMDPs," .

RABOTYAGOV, S. S., T. D. CAMPBELL, M. WHITE, J. G. ARNOLD, M. L. N. JAY ATWOOD, AND C. L. K. ET AL (2014a): "Cost-effective targeting of conservation investments to reduce the northern Gulf of Mexico hypoxic zone," *Proceedings of the National Academy of Sciences*, 111(52), 18530–18535.

RABOTYAGOV, S. S., C. L. KLING, P. W. GASSMAN, N. N. RABALAIS, , AND R. E. TURNER (2014b): "The economics of dead zones: Causes, impacts, policy challenges, and a model of the Gulf of Mexico hypoxic zone," *Review of Environmental Economics and Policy*, 8, 58–79.

RINGEVAL, B., J. DEMAY, D. S. GOLL, X. HE, Y. P. WANG, E. HOU, ..., AND S. PELLERIN (2018): "A global dataset on phosphorus in agricultural soils. Scientific Data," *Scientific Data*, 11(1), 17.

ROUT, T. M., J. L. MOORE, AND M. A. MCCARTHY (2014): "Prevent, search or destroy? A partially observable model for invasive species management," *Journal of Applied Ecology*, 51(3), 804–813.

SAARELA, I., A. JäRVI, H. HAKKOLA, AND K. RINNE (1995): "Field Trials on Fertilization 1977-1994." *Agrifood Research Finland Bulletin*, 16.

SATTARI, S. Z., A. F. BOUWMAN, K. E. GILLER, AND M. K. VAN ITTERSUM (2012): "Residual soil phosphorus as the missing piece in the global phosphorus crisis puzzle," *Proceedings of the National Academy of Sciences*, 109(16), 6348–6353.

SCHNITKEY, G. D. AND M. J. MIRANDA (1993): "The Impact of Pollution Controls on Livestock-Crop Producers," *Journal of Agricultural and Resource Economics*, 18(1), 25–36.

SIMS, C., D. FINNOFF, A. HASTINGS, AND J. HOCHARD (2017): "Listing and delisting thresholds under the Endangered Species Act," *American Journal of Agricultural Economics*, 99(3), 549–570.

SLOGGY, M. R., D. M. KLING, AND A. J. PLANTINGA (2020): "Measure twice, cut once: Optimal Inventory and harvest under volume uncertainty and Stochastic Price Dynamics," *Journal of Environmental Economics and Management*, 103, 102357.

SPRINGBORN, M. AND J. N. SANCHIRICO (2013): "A density projection approach for non-trivial information dynamics: Adaptive management of stochastic natural resources," *Journal of Environmental Economics and Management*, 66(3), 609–624.

SUD, M. (2020): "Managing the biodiversity impacts of fertiliser and pesticide use: Overview and insights from trends and policies across selected OECD countries," *OECD Publishing, Paris*, Working Papers, No. 155.

SUN, S., M. S. DELGADO, AND J. P. SESMERO (2016): "Dynamic adjustment in agricultural practices to economic incentives aiming to decrease fertilizer application. Journal of Environmental Management," *Journal of Environmental Management*, 177, 192–201.

TOMBERLIN, D. AND T. ISH (2007): "When is logging road erosion worth monitoring. In Proceedings of the International Mountain Logging and 13th Pacific Northwest Skyline Symposium (pp. 239-244)," *Oregon*

State University, College of Forestry Corvallis Oregon.

US EPA (1995): "In National Nutrient Assessment Workshop: Proceedings, December 4-6," *United States Environmental Protection Agency, Washington, D.C.*

USDA (09.2022): "Small Scale Solutions for Your Farm - Soil Testing," *U.S. Department of Agriculture-Natural Resource and Convservation Service*.

USDA (2020): "Legacy Phosphorus Project," *United States Department of Agriculture*.

USDA (2024a): "Fertilizer Use and Price," *United States Department of Agriculture*, https://www.ers.usda.gov/data-products/fertilizer-use-and-price/.

——— (2024b): "U.S. Bioenergy Statistics," *United States Department of Agriculture*, https://www.ers.usda.gov/data-products/u-s-bioenergy-statistics/.

VOS, R., J. GLAUBER, C. HEBEBRAND, AND B. RICE (2025): "Global shocks to fertilizer markets: Impacts on prices, demand and farm profitability," *Food Policy*, 133, 102790.

WHITE, B. (2005): "An economic analysis of Ecological Monitoring," *Ecological Modelling*, 189(3-4), 241–250.

WIRONEN, M. B., E. M. BENNETT, AND J. D. ERICKSON (2018): "Phosphorus flows and legacy accumulation in an animal-dominated agricultural region from 1925 to 2012," *Global Environmental Change*, 50, 88–99.

ZHOU, E., M. C. FU, AND S. I. MARCUS (2010): "Solving continuous-state POMDPs via density projection," *IEEE Transactions on Automatic Control*, 55(5), 1101–1116.

ZOU, T., X. ZHANG, AND E. A. DAVIDSON (2022): "Global trends of cropland phosphorus use and sustainability challenges," *Nature*, 611(7934), 81–87.

Appendix

Evaluating Optimal Farm Management of Phosphorus Fertilizer Inputs with Partial Observability of Legacy Soil Stocks

Table of Contents

This appendix provides additional details and elaborates on the methodology described in the paper. It begins by outlining the foundations of partially observable Markov decision processes (POMDP), introduces a two-stage belief-updating process for soil sampling and crop yield, describes the basis of the Markov-switching Vector Autoregressive (MSVAR) model, discusses the representativeness of North Carolina in phosphorus (P) fertilizer use, presents an alternative crop yield function with results, and concludes with supplementary figures.

A Methodology and Algorithms

A.1 Detailed formulation of the Dynamic Model

The Bellman equation for the recursive expected utility function (6) can be detailed as:

$$\begin{aligned} V(\boldsymbol{S_t}) = \max_{F,s} & \iiint \pi(\boldsymbol{S_t}, F_t, s_t) f(P_{t+1}^Y \mid P_t^Y, P_t^F) b_t(L_t) g(\epsilon_t^Y) d\epsilon_t^Y \, dP_{t+1}^Y dL_t \\ & + \beta \iint p(\boldsymbol{P}_{t+1}^{(r)} \mid \boldsymbol{P}_t^{(r)}) p(O_{t+1}^s \mid b_{t+1}(L_t), s_t) V(\boldsymbol{S_{t+1}}) dO_{t+1}^s d\boldsymbol{P}_{t+1}^{(r)}, \end{aligned} \tag{A1}$$

and given Epstein-Zin preferences, (8) can be further detailed as:

$$\begin{aligned} V_{EZ}(\boldsymbol{S_t}) = \max_{F,s} \Bigg[& (1-\beta) \Bigg(\iiint \pi(\boldsymbol{S_t}, F_t, s_t)^{1-\eta} f(P_{t+1}^Y \mid P_t^Y, P_t^F) b_t(L_t) g(\epsilon_t^Y) d\epsilon_t^Y \, dP_{t+1}^Y dL_t \Bigg)^{\frac{1-\psi^{-1}}{1-\eta}} \\ & + \beta \Bigg(\iint p(\boldsymbol{P}_{t+1}^{(r)} \mid \boldsymbol{P}_t^{(r)}) p(O_{t+1}^s \mid b_{t+1}(L_{t+1}), s_t) V_{EZ}(\boldsymbol{S_{t+1}})^{1-\eta} dO_{t+1}^s d\boldsymbol{P}_{t+1}^{(r)} \Bigg)^{\frac{1-\psi^{-1}}{1-\eta}} \Bigg]^{\frac{1}{1-\psi^{-1}}}, \end{aligned} \tag{A2}$$

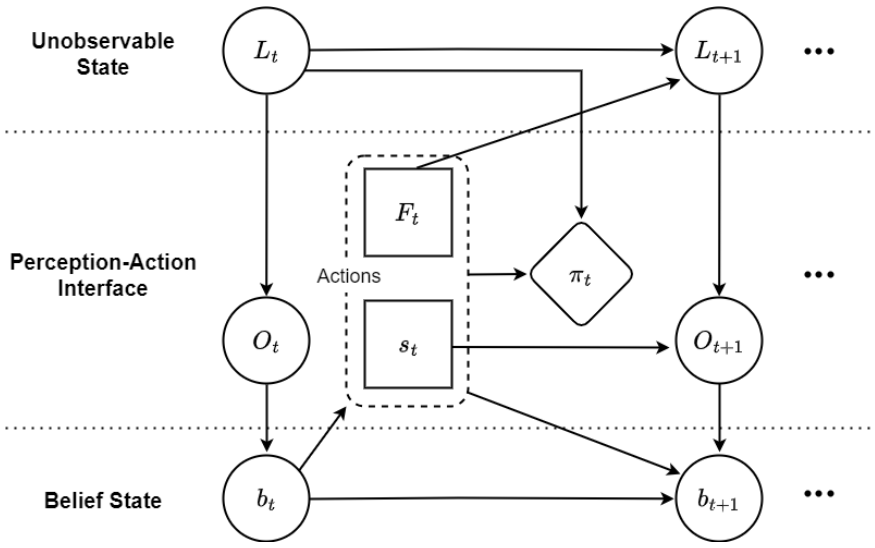
under the state variables, $\boldsymbol{S_t} \equiv [b_t(L_t), P_t^Y, P_t^F]$, where $b_t(L_t)$ represents the farmer's belief about the latent legacy P state at time t, P_t^Y denotes the market price of corn at time t, and P_t^F denotes the price of P fertilizer at time t. Additionally, the notation $\boldsymbol{P}_t^{(r)}$ refers to the set of corn and P fertilizer prices, $[P_t^Y, P_t^F]$, that are associated with the price regime r_t at time t. Here, the price regime r_t captures the underlying economic or policy conditions influencing the joint price dynamics of corn and P fertilizer.

A.2 Solution Methods of Projected Belief

A simplified schematic of our POMDP model is shown in Figure A1, with the biophysical dynamics of legacy P stocks L_t at the top level of the figure. A continuous-state POMDP is challenging because it involves an infinite-dimensional belief space, and discretizing this space to approximate beliefs can lead to significant computational difficulties. Exact evaluation of the posterior distribution is difficult to address, and even structuring the belief updating process in a discretized space is often infeasible. To address this challenge, a density projection technique suggested by Zhou et al. (2010) and employed by Kling et al. (2017) in economics is utilized.

Density projection projects the infinite-dimensional belief space onto a low-dimensional param-

Figure A1: Schematic of Partially Observable Markov Decision Process



Notes: The farmer infers their unobserved legacy P stock L_t through observations O_t and updates their belief state b_t. Phosphorus fertilizer application F_t and soil sampling s_t influence the state transition L_{t+1} and future observations O_{t+1}, respectively.

eterized family of densities.[1] Projection mapping from the belief state $b(L)$ to exponential family of density $f(L;\theta)$, where θ is a natural parameter, is achieved by minimizing the Kullback-Leibler divergence between $b(L)$ and $f(L;\theta)$ as:

$$\begin{aligned} b^P(L) &\triangleq \arg\min_f \; D_{KL}(b \parallel f) \\ &\text{where } D_{KL}(b \parallel f) \triangleq \int b(L) \log \frac{b(L)}{f(L;\theta)} dL \\ &\forall L, b(L) > 0 \leftrightarrow f(L;\theta) > 0 \end{aligned} \tag{A3}$$

and thus belief $b(L)$ and its projection $f(L;\theta)$ satisfies:

$$\mathbb{E}_b[T_j(L)] = \mathbb{E}_\theta[T_j(L)] \quad \text{for } j = 1, 2, ..., J \tag{A4}$$

where $T(L)$ is the sufficient statistics of the probability density (Zhou et al. 2010).

Bayesian updating of projected belief state is implemented adopting a particle filtering, which uses a Monte Carolo simulation approach to estimate the belief state with a limited set of particles (samples) and simulates the transition of the belief state (De Freitas 2001, Arulampalam et al. 2002).

[1]Technical interpretation of density projection and particle filtering hereafter closely follows Zhou et al. (2010).

In the particle filtering, particles L_t^i for $i = 1, 2, ..., Z$ are drawn from $b_t(L_t)$ and L_{t+1}^i from the propagation $p(L_{t+1}|L_t, F_t, s_t)$. This allows for the approximation of $b_{t+1}(L_{t+1})$ by the probability mass function (Zhou et al. 2010):

$$b_{t+1}(L_{t+1}) \approx \sum_{i=1}^{Z} \tau_{t+1}^i \phi(L_{t+1} - L_{t+1}^i) \tag{A5}$$

where $\tau_{t+1}^i \propto p(O_{t+1}^l|L_{t+1}^i, F_t, s_t)$, denoting the associated weight and ϕ represent the Kronecker delta function. Substituting equation (A5) into (A4), the approximation becomes:

$$\begin{aligned}
\mathbb{E}_{b_{t+1}}[T_j(L_{t+1})] &= \int T_j(L_{t+1}) b_{t+1}(L_{t+1}) dL_{t+1} \\
&\approx \int T_j(L_{t+1}) \left[\sum_{i=1}^{Z} \tau_{t+1}^i \phi(L_{t+1} - L_{t+1}^i) \right] dL_{t+1} \\
&= \sum_{i=1}^{Z} \tau_{t+1}^i T_j(L_{t+1}^i) \\
&= \mathbb{E}_{\theta_{t+1}}[T_j(L_{t+1})]
\end{aligned} \tag{A6}$$

simplified by the properties of the Kronecker delta function. Thus, if the particles L_t^i are drawn from the projected belief state $b_t^P = f(\cdot; \theta_t)$ and their propagation L_{t+1}^i satisfy the $\sum_{i=1}^{Z} \tau_{t+1}^i T_j(L_{t+1}^i) = \mathbb{E}_{\theta_{t+1}}[T_j(L_{t+1})]$, the transition probability of θ_t to θ_{t+1} can be calculated.

Density projection effectively reduces infinite-dimensional density to low-dimensional, parameter-defined density, transforming the belief Markov decision process (MDP) into a more manageable and solvable form referred to as 'projected belief MDP'. In this paper, the legacy P states are defined as the natural parameters of log-normal distribution and transform to the θ in the 'projected belief MDP' calculation (Kling et al. 2017). The utilization of the log-normal distribution in parameterized density is particularly advantageous, primarily due to its tractability to positive-valued state variables and its parametric simplicity characterized by two parameters: mean and coefficient variation (Sloggy et al. 2020).

While there are numerous ways to solve the projected belief MDP, we follow Kling et al. (2017) and discretize the projected belief MDP space into a discrete-state space. Because the value function in eq. (6) and (8) is a function both of the belief and price states, we then compute the value function on a grid of all discretized possible belief and price state combinations.

The projected belief MDP is a low-dimensional, continuous state MDP (Zhou et al. 2010).

To facilitate the value iteration, we first convert the projected belief MDP into a discrete state MDP.[2] This conversion involves discretizing the space of natural parameters θ in the exponential distribution $f(\cdot|\theta)$ (Zhou et al. 2010). In this paper, we employ the log-normal distribution to define legacy P bioavailability μ_L and uncertainty in legacy P bioavailability as coefficient variation ν_L ($\nu_L = \sigma_L/\mu_L$) with a parameter set $\delta = \{\mu_L, \nu_L\}$. Hence, we discretize θ by calculating the univariate log-normal parameters μ and σ that $\theta = \{\mu, \sigma\}$ where $\sigma > 0$ from the δ (Kling et al. 2017). The calculation of μ and σ is follows:

$$\mu = \ln\left(\frac{\mu_L^2}{\sqrt{\mu_L^2 + \sigma_L^2}}\right), \qquad \sigma^2 = \ln\left(1 + \frac{\sigma_L^2}{\mu_L^2}\right). \tag{A7}$$

For the estimation in discretized space, μ_L and σ_L are discretized into a 100×1 vectors. A 100×100 mesh grid $\{\delta_i\}_{i=1}^N = G$ is then calculated, incorporating all grid points $\delta_i = \{\mu_{L,i}, \nu_{L,i}\}$ where $\nu_{L,i} = \sigma_{L,i}/\mu_{L,i}$. Within this discretized state space δ_i, the crop profit function is evaluated as the expected value of δ_i, in associated with controls F, s and prices P^Y, P^F. By defining the transition probability as $\tilde{p}(\delta_i, F, s)(\delta_j)$, representing the probability to transitioning from δ_i to δ_j, the discretized belief MDP for eq (A1) is formulated as:

$$\begin{aligned}\tilde{V}(\delta_i, P^Y, P^F) = &\max_{F,s} \tilde{\pi}(\delta_i, F, P^{Y'}, P^F, s) \\ &+ \beta \sum^{P^{(r)'}} \sum_{j=1}^{N} p(\boldsymbol{P}^{(r)'} \mid \boldsymbol{P}^{(r)}) \tilde{p}(\delta_i, F, s)(\delta_j) \tilde{V}(\delta_j, P^{Y'}, P^{F'}),\end{aligned} \tag{A8}$$

where $p(\boldsymbol{P}^{(r)'} \mid \boldsymbol{P}^{(r)})$ denote the discretized transition probability of corn price P^Y and P fertilizer price P^F, estimated from the MSVAR model.[3]

The profit function $\tilde{\pi}(\delta_i, F, P^Y, P^F, s)$ and transition probability $\tilde{p}(\delta_i, F, s)(\delta_j)$ associated with controls F and s can be estimated by using Monte-Carlo simulation, as denoted in Algorithm 1 and Algorithm 2 (Zhou et al. 2010).

$\boldsymbol{\omega}^k$ is the set of Sobol points $\boldsymbol{\omega}^k = \{\boldsymbol{\omega}_1^k, \boldsymbol{\omega}_2^k, ..., \boldsymbol{\omega}_Z^k\}$ that derived from Sobol sequence. For the estimation of crop profit function and transition probability, we use the three-dimensional ($k = 3$) Sobol points $\boldsymbol{\omega}^k$ that includes Z =10,000 points. In the draw process, the Sobol draw omits an initial 1,000 points, then select every 101st point thereafter (MathWorks. 2024). We also apply a random

[2]The discretization and estimation methods are adopted from Zhou et al. 2010 and Kling et al. 2017.

[3]In price dynamics, the $t+1$ state is represented by $'$ notation.

linear scramble along with a random digit shift. In the estimation of log-likelihood function, Sobol draw is efficient methods. To achieve the same precision level of 1,000 Sobol draws in the estimation of log-likelihood function value, the estimation requires the 1,661 Haltom draws, 4,155 Modified Latin Hyper Cube Sampling draws or 9,987 pseudo-random draws (Czajkowski and Budziński 2019). With a five-dimensional Sobol draw, the desired precision level requires at least 2,100 points (Czajkowski and Budziński 2019), and we choose the number of points to 10,000 to increase the precision level.

Estimation of transition probability $\tilde{p}(\delta_i, F, s)(\delta_j)$ is in Algorithm 2. Based on the output from Algorithm 2. and the estimated transition probabilities of corn and P fertilizer price, we proceed to calculate the comprehensive of transition probabilities $p(\boldsymbol{P}^{(r)'} \mid \boldsymbol{P}^{(r)})\tilde{p}(\delta_i, F, s)(\delta_j)$. The combination of these probabilities is achieved through the Kronecker product of probability matrices for corn and P fertilizer prices, as well as the transition probabilities $\mathbf{P}_r \otimes \tilde{\mathbf{P}}$ (Sloggy et al. 2020), where $\mathbf{P}_r$ is the probability matrix for corn and P fertilizer prices over moderate and high regimes and $\tilde{\mathbf{P}}$ is the estimated probability matrix of $\forall i, j,\ \tilde{p}(\delta_i, F, s)(\delta_j)$.

Algorithm 1. Estimation of Crop Profit Function

Input: $\delta_i, P^Y, P^F, F, s, \boldsymbol{\omega}^2$

Output: $\tilde{\pi}(\delta_i, F, P^Y, P^F, s)$

Step 1. Sampling:

$$\boldsymbol{L} = f^{-1}(\boldsymbol{\omega}^2|\theta_i) \qquad \text{where } \boldsymbol{L} = \{L_1, L_2, ..., L_Z\}$$

Step 2. Estimation:

$$\tilde{\pi}(\delta_i, F, P^{Y'}, P^F, s) = \frac{1}{Z}\sum_{j=1}^{Z}\sum^{P^{Y'}}\pi(L_j, F, P^{Y'}, P^F, s)f(P^{Y'} \mid P^Y, P^F)$$

Source: Zhou et al. (2010)

Algorithm 2. Estimation of transition probability

Input: $\delta_i, P^Y, P^F, F, s, \boldsymbol{\omega}^1, \boldsymbol{\omega}^2, \boldsymbol{\omega}^3$

Output: $\tilde{p}(\delta_i, F, s)(\delta_j)$

Step 1. Sampling:

$$\boldsymbol{L} = f^{-1}(\boldsymbol{\omega}^2|\theta_i) \quad \text{where } \boldsymbol{L} = \{L_1, L_2, ..., L_Z\}$$

Step 2. Compute $\tilde{\boldsymbol{L}}$ by propagation of $\boldsymbol{L}$ according to the dynamics of legacy P (eq. 1) using controls F and s, and carry-over parameter ρ that is generated using $\boldsymbol{\omega}^1$.

Step 3. Compute $O_1^s, O_2^s, O_3^s, ..., O_Z^s$ from $\tilde{\boldsymbol{L}} = \{\tilde{L}_1, \tilde{L}_2, \tilde{L}_3, ..., \tilde{L}_Z\}$ using (3) and observation error $\{\lambda_i^s\}_{i=1}^Z$ that is generated by $\boldsymbol{\omega}^3$, where s is determined by the controls ss and ps.

Step 4. For each O_k^s, $k = 1, 2, ..., Z$, compute the updated belief state

$$\tilde{b}_k = \sum_{i=1}^{Z} \tau_i^k \phi(L - \tilde{L}_i),$$

where ϕ is the Kronecker delta product function and

$$\tau_i^k = \frac{p(O_k^s \mid \tilde{L}_i, F, s)}{\sum_{i=1}^{Z} p(O_k^s \mid \tilde{L}_i, F, s)}$$

Step 5. For $k = 1, 2, ..., Z$ project each $\tilde{b}_k$ onto the lognormal density to find $\tilde{\theta}_k$, and compute $\hat{\delta}_k$ from $\tilde{\theta}_k$.

Step 6. For each $k = 1, 2, ..., Z$, calculate the bilinear interpolation weight for $\tilde{\delta}_k$ on G. For each $\tilde{\delta}_k$, sum the bilinear interpolation weight.

$$\tilde{p}(\delta_i, F, s)(\delta_j) = \frac{\text{sum of bilinear interpolation weights assigned to } \delta_j}{Z}$$

Source: Zhou et al. (2010), Kling et al. (2017)

B Soil Sampling and Yield-Based Information Update

In this section, we discuss the process behind the two-stage belief updating mechanism. The two-stage approach considers the dynamic nature of decision-making in agricultural practices, where information is acquired at different points in time.

The first stage of belief updating occurs when farmers conduct soil sampling before making fertilization decisions. The second stage of belief updating takes place after the fertilization and harvest, when farmers receive additional information through the actual corn yield. This yield data, reflecting the results of their fertilization decisions, provides a further opportunity to update their beliefs about the legacy P state.

B.1 Two-Stage Belief Updating Process

Figure B1 illustrates the two-stage belief updating process within a POMDP framework. In this model, the unobservable state of legacy P L_t represents the true but hidden condition of the bioavailability at time t, which evolves to a new state L_{t+1} by the next time period. Farmers, unable to directly observe this state, rely on a sequence of observations and actions to update their beliefs about the legacy P condition.

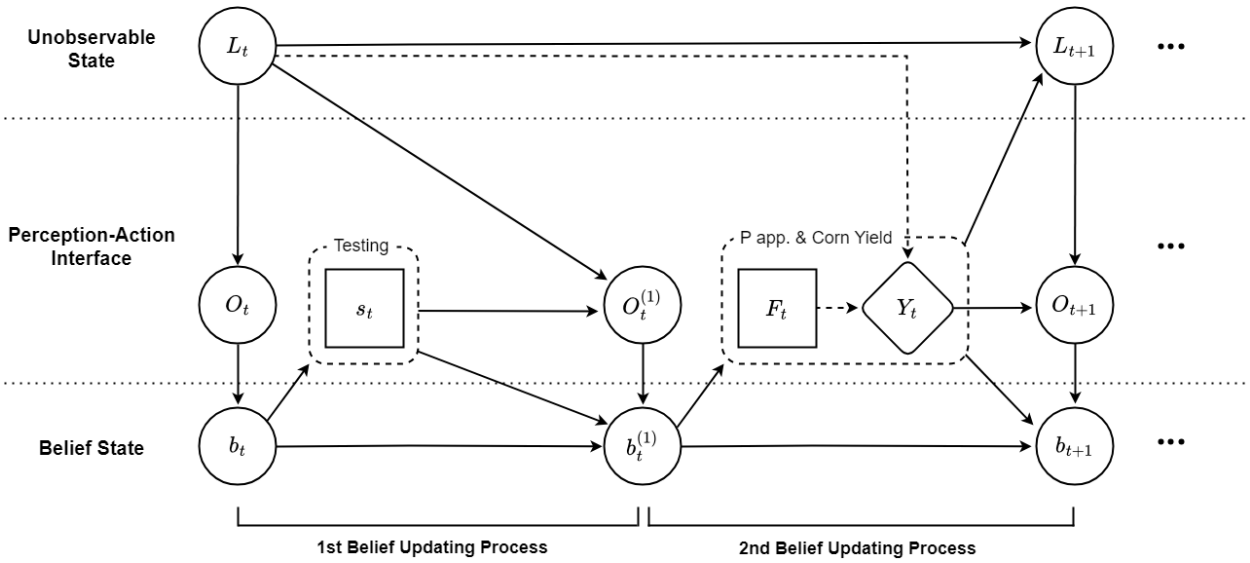
The process begins with soil sampling s_t, where farmers obtain an initial observation $O_t^{(1)}$ that provides partial information about the current state L_t. This observation is used to update their belief from b_t to $b_t^{(1)}$ within a period, forming the first stage of belief updating. Following this, farmers apply P fertilizer F_t, and the resulting corn yield Y_t offers additional information. This yield data leads to a second update of their belief, from $b_t^{(1)}$ to b_{t+1}, as they refine their understanding of the legacy P state.

The arrows in the figure indicate the flow of information between these components, showing how observations from soil sampling and yield outcomes interact with the unobservable state to update the belief state over time. In the first belief updating process, because farmers adopt soil sampling before making a P fertilizer decision, the belief updating process begins with the following equation:

$$b_t^{(1)}(L_t) \propto p(O_t^s \mid L_t, s_t) b_t(L_t), \tag{B1}$$

where O_t^s represents the observation obtained from soil sampling s_t. The belief $b_t(L_t)$ is updated to $b_t^{(1)}(L_t)$ based on the new information provided by the soil sampling. This updated belief reflects the farmer's revised understanding of the legacy P state L_t after considering the soil test results.

Figure B1: Schematic of Two-Stage Belief Updating Process



Notes: Figure B1 illustrates the two-stage belief updating process for farmers' decision-making in a POMDP framework. The first stage involves updating the belief state b_t based on soil sampling s_t, and the second stage further updates the belief using corn yield Y_t after P fertilizer application F_t.

The next step in the belief updating process occurs after the corn yield Y_t is realized. The corn yield is calculated based on the current legacy P state, and it is conditional on the P fertilizer application F_t. In our model, we assume that farmers directly obtain information from the corn yield Y_t. Consequently, we assume that the observation O_{t+1} at time $t+1$ is equivalent to the yield Y_t. This assumption is based on the fact that the yield is a direct and observable outcome that strongly influences the farmer's beliefs about the soil's legacy P levels. For instance, if the yield Y_t is high, farmers are likely to believe that the soil has a high level of legacy P, suggesting that their prior application of fertilizer was effective or that the soil had sufficient nutrient reserves. Conversely, a low yield might lead farmers to adjust their beliefs toward the soil having lower legacy P levels. By equating O_{t+1} with Y_t, we simplify the belief updating process while still capturing the essential feedback mechanism that guides farmers' future management decisions. The belief updating process at this stage is represented by the equation:

$$b_{t+1}(L_{t+1}) \propto \int p(Y_t \mid L_t, F_t) p(L_{t+1} \mid L_t, F_t) b_t^{(1)}(L_t)\, dL_t. \tag{B2}$$

Here, $b_{t+1}(L_{t+1})$ is the updated belief at time $t+1$, taking into account the information provided by the corn yield Y_t. The term $P(Y_t \mid L_t, F_t)$ represents the likelihood of observing the yield given

the previous legacy P state and the fertilizer application, while $p(L_{t+1} \mid L_t, F_t)$ represents the propagation of the legacy P state from time t to $t+1$ given the fertilizer application F_t.

B.2 Defining the Corn Yield Distribution and Bayesian Updating

In this section, we first define the corn yield distribution for likelihood $p(Y_t \mid L_t, F_t)$ in our Bayesian updating process. Corn yield distributions are derived from Sobol points $\boldsymbol{L}$ for each natural parameter θ_i following Algorithm 1.

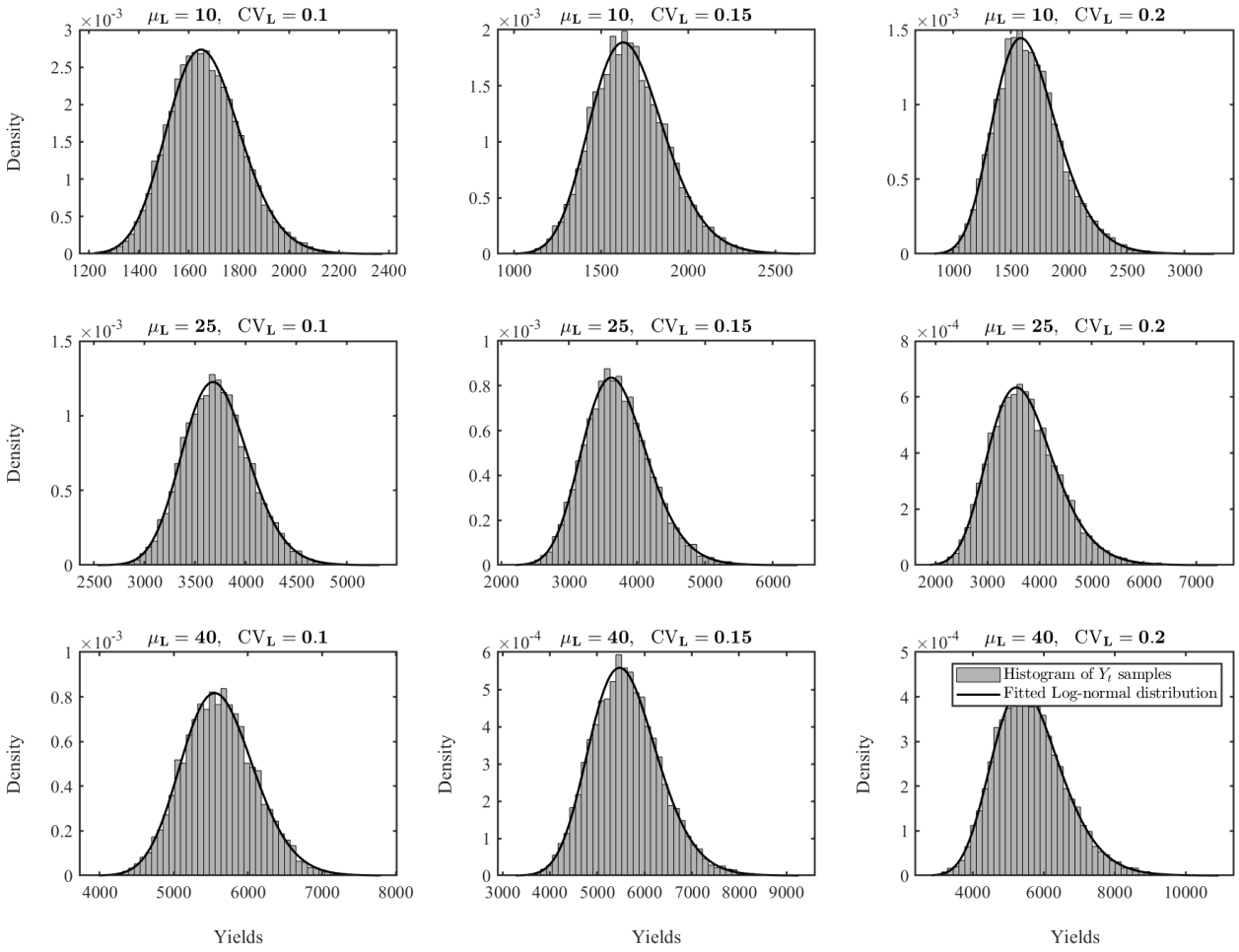
These Sobol points are generated from a log-normal distribution, resulting in yield samples Y_t that align with a log-normal distribution, as shown in Fig. B2. The fitted log-normal curve effectively captures this distribution, accurately representing the central tendency, variability, and skewness inherent in the yield simulations. Defining the form of the yield distribution is essential for our Bayesian updating because the yield serves as an observation in the likelihood function. By understanding the distribution of yields under varying P fertilizer applications, we can better inform the belief updating process, ensuring that our model realistically reflects the probabilistic nature of yield outcomes.

Figure B3 depicts the belief update based on information obtained from soil sampling. The prior belief distribution b_t, (shown by the red dashed line) is updated to the posterior distribution $b_t^{(1)}$, (shown by the blue solid line) after incorporating the soil sampling observation (indicated by the black dotted line). As seen in the figure, the observation provides significant information, leading to a notable shift in the belief from the prior to the posterior distribution. This substantial update indicates that the soil sampling results are effective in refining the farmer's understanding of the legacy P state, which is crucial for making informed fertilization decisions.

In contrast, Fig. B4 represents the second stage of belief updating, where the belief is further adjusted based on information from the crop yield. The posterior distribution from the first stage $b_t^{(1)}$ now serves as the prior distribution in this stage, and the crop yield observation (again indicated by the black dotted line) informs the update to the final posterior distribution b_{t+1} (shown by the green solid line). However, in this stage, the crop yield provides less additional information, resulting in a less pronounced shift from the prior to the posterior distribution. The reason for this is that the crop yield, while reflective of the legacy P state, is also influenced by other factors, leading to greater uncertainty and less precise updating of the belief. Consequently, the posterior distribution has a relatively wider variation, indicating that the yield data does not significantly help the farmer's understanding of the legacy P levels compared to the soil sampling. This analysis

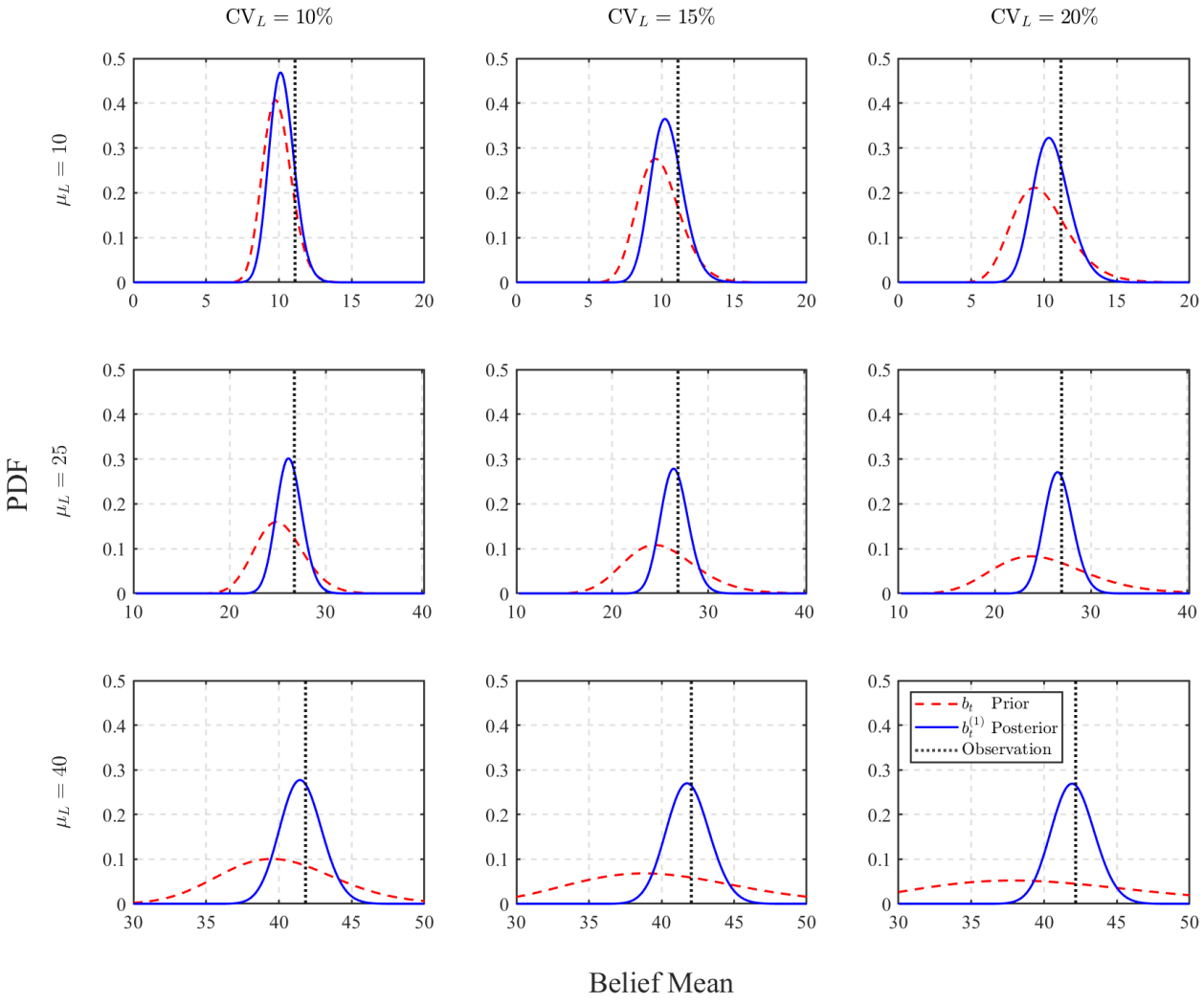
demonstrates the critical role of soil sampling in the belief-updating process, particularly in the first stage, and in our paper, we consider soil sampling for the belief-updating process.

Figure B2: Distributions of Corn Yields



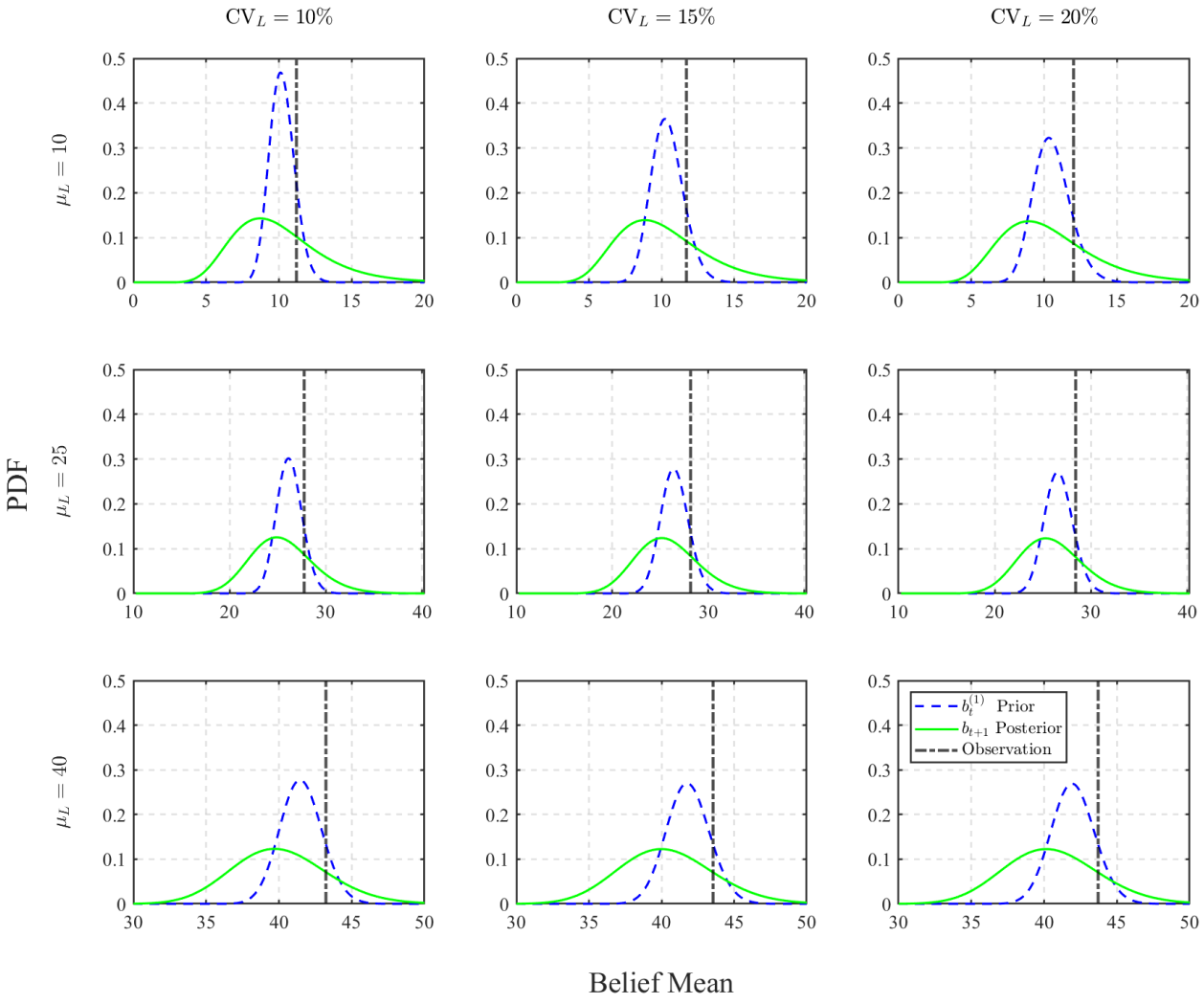
Notes: Figure B2 demonstrates that the distribution of corn yields closely follows a log-normal distribution across different mean legacy P levels (μ_L) and coefficients of variation (ν_L). The histograms of the simulated yield samples, along with the overlaid fitted log-normal curves, show a strong alignment between the empirical simulations and the log-normal distribution. This consistency across varying scenarios justifies the assumption that corn yield distributions can be appropriately modeled using a log-normal distribution in subsequent analyses.

Figure B3: First Stage Belief Updating Process



Notes: Figure B3 illustrates the prior b_t and posterior $b_t^{(1)}$ distributons across different combinations of legacy P bioavailability μ_L and uncertainty ν_L. The observation (black dotted line) is derived from soil sampling, which provides partial information about the legacy P levels. The prior belief (red dashed line) is updated to the posterior belief (blue solid line) after incorporating the soil sampling observation. Each subplot corresponds to a different combination of μ_L and ν_L, demonstrating how these parameters influence the updating process and the resulting belief distributions. As ν_L increases, the posterior distribution becomes wider, indicating greater uncertainty in the belief about the legacy P state.

Figure B4: Second Stage Belief Updating Process



Notes: Figure B4 shows the progression from prior $b_t^{(1)}$ to posterior b_{t+1} belief distributions after incorporating additional information from the crop yield, across different combinations of legacy P bioavailability μ_L and uncertainty ν_L for the legacy P state. The observation from the crop yield is indicated by the black dashed line, which further informs the belief updating process. The prior distribution (blue dashed line) represents the belief after the first stage of updating, while the posterior distribution (green solid line) reflects the updated belief after considering the crop yield observation.

C Estimation of Price Transition: Markov-switching Dynamic Regression

In this section, We estimate a log-linear Markow-switching Dynamic Regression (MSDR) model to construct the basis parameters of prior distributions (μ_μ, σ_μ^2 and μ_Φ, σ_Φ^2) in MSVAR model. The specification of the MSDR of the following form:

$$\begin{aligned} \ln(P_{t+1}^Y) &= \alpha_{0,r_{t+1}} + \alpha_{1,r_{t+1}} \ln(P_t^Y) + \alpha_{2,r_{t+1}} \ln(P_t^F) + \epsilon_{t+1} \\ \ln(P_{t+1}^F) &= \beta_{0,r_{t+1}} + \beta_{1,r_{t+1}} \ln(P_t^F) + \beta_{2,r_{t+1}} \ln(P_t^Y) + v_{t+1} \end{aligned} \tag{C1}$$

where $\alpha_{0,r_{t+1}}, \beta_{0,r_{t+1}}$ are the intercepts for price regime r_{t+1} and ϵ_{t+1}, v_{t+1} are the identical distribution (i.i.d.) normal errors with mean zero and regime-dependent variance $\sigma^2_{\epsilon,r_{t+1}}, \sigma^2_{v,r_{t+1}}$, respectively. We allow for two price regimes in the model, $r_t \in \{\text{moderate, high}\}$.

MDSR results are presented in table C1. To generate the basis parameters of prior distribution of intercept (μ_μ, σ_μ^2) and coefficients $(\mu_\Phi, \sigma_\Phi^2)$, we use the averaged value of estimated values (α_0, β_0) and (α_i, β_i) for each prior mean and used the averaged values of standard error to calculated the prior variances $(\sigma_\mu^2, \sigma_\Phi^2)$

Table C1: Markov switching dynamics regression for corn and phosphorus fertilizer prices

	Corn ($\ln(P_{t+1}^Y)$)		Phosphorus fertilizer ($\ln(P_{t+1}^F)$)	
	Moderate	High	Moderate	High
$\ln(P_t^F)$	0.091	0.763***	0.947***	–1.347***
	(0.186)	(0.284)	(0.151)	(0.251)
$\ln(P_t^Y)$	0.633***	0.280	–0.034	2.234***
	(0.199)	(0.205)	(0.097)	(0.470)
Const.	–0.410	–3.408**	0.275	11.866***
$(\alpha_{0,r_t}, \beta_{0,r_t})$	(0.923)	(1.448)	(0.723)	(1.862)
Std Dev.	0.109		0.075	
$(\sigma_{\epsilon,r_t}, \sigma_{v,r_t})$	(0.014)		(0.009)	
Log-likelihood	12.309		31.502	
AIC	–0.207		–1.406	

Notes: Robust standard errors are in parentheses. In the regression, constant standard deviation $\sigma^2 = \sigma_i^2 = \sigma_j^2$ is assumed for $r_t \in \{i, j\}$, $i \neq j$. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

D Representativeness of North Carolina in P Fertilizer Use

To support the generalizability of our eastern North Carolina case study, we present a comparison of P fertilizer application across U.S. corn-producing states using United States Department of Agriculture (USDA) data. Specifically, we use long-term averages (1946–2018) provided by the USDA for 27 major corn-producing states (USDA, 2024).

Figure D1 shows the percentage of corn acreage receiving P fertilizer. North Carolina (highlighted in blue) aligns closely with the U.S. average (gray dashed line) and median (red dashed line), falling well within the interquartile range of most other states.

Figure D2 presents the average P fertilizer application rate per fertilized acre. Once again, North Carolina is near the U.S. average and median, and is not an outlier in terms of intensity.

These two indicators—fertilizer coverage and application intensity—suggest that North Carolina is broadly representative of typical U.S. corn farming practices with respect to P management. This supports our use of the North Carolina context as a meaningful case study for evaluating behavior under legacy P uncertainty.

Figure D1: Percentage of corn acreage receiving phosphorus fertilizer

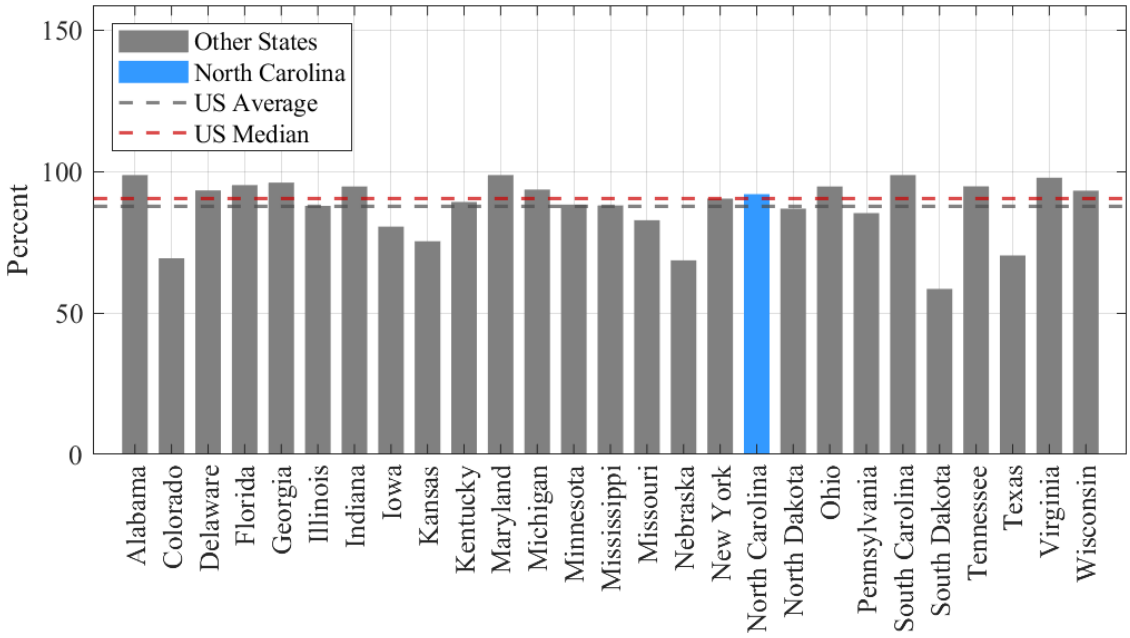
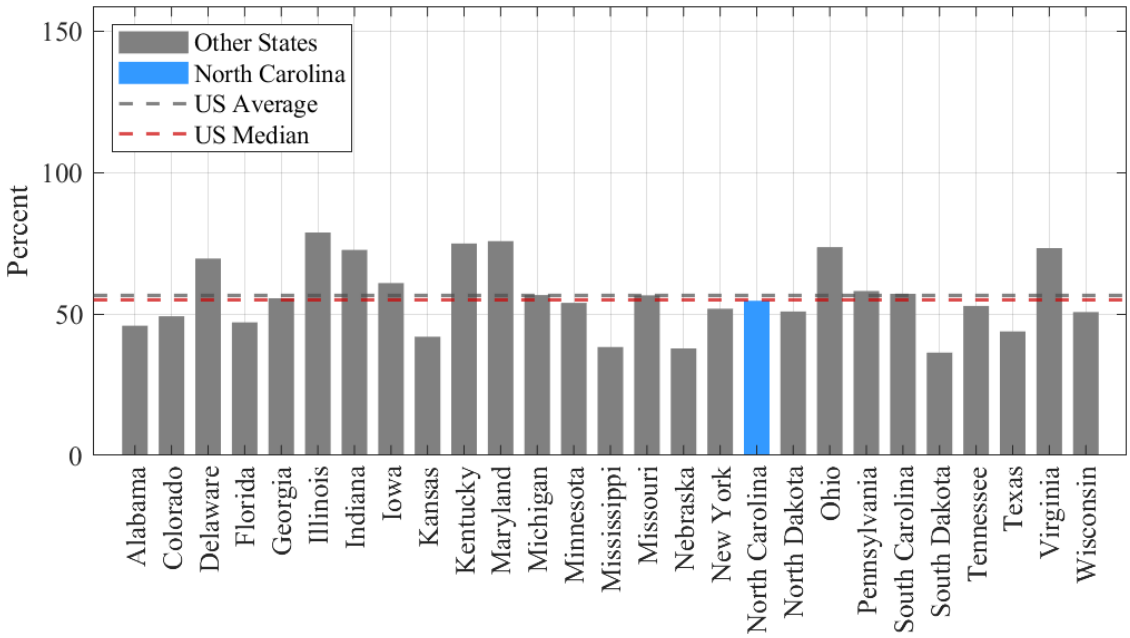


Figure D2: Phosphate used on corn, rate per fertilized acre receiving phosphorus fertilizer



E Additional Analysis: Effects of Modified Crop Yield Estimation

In this section, we evaluate how an alternative crop yield estimation influences fertilizer application and soil testing decisions. To permit flexible estimation of the substitutability between legacy P and fertilizer applications, we use a slight modification of a translog production function, as follows:

$$\ln(Y_{i,t}) = \beta_0 + \beta_1 \ln(F_{i,t}) + \beta_2 \ln(L_{i,t}) + \beta_3 \ln(F_{i,t})^2 + \beta_4 \ln(F_{i,t}) \ln(L_{i,t}) + \omega_i + \epsilon_{i,t}, \tag{E1}$$

where i denotes experiment plot, ω_i is the experimental replication (i.e. plot) fixed effect, and $\epsilon_{i,t}$ is a timevarying error component. The estimation results are in Table E1.

Unlike the concave yield response used in paper, the additional yield function converges to a stable point as fertilizer applications increase, reflecting diminishing returns to additional inputs. Fig. E1 shows the simulated estimation results. Under this yield structure, results show that farmers apply fertilizer more conservatively, avoiding excessive use that would have limited yield or profit benefits. Specifically, yields rise initially with increasing fertilizer application but saturate at higher levels, while profits exhibit a similar plateau effect.

Interestingly, when analyzing the risk-neutral farmer case based on this crop yield estimation (Fig. E2), the optimal results remain similar to those presented in the main manuscript. Despite the adoption of a modified crop yield estimation, which converges rather than follows a concave response, the key patterns in fertilizer application and soil testing decisions persist. Point soil sampling continues to provide value, particularly in regions with high uncertainty about legacy P availability and low legacy P level, and fertilizer application strategies remain aligned with the original findings. This consistency reinforces the robustness of the original analysis, demonstrating that the overall conclusions regarding the role of risk preferences and enhanced monitoring are not sensitive to the specific yield response function used.

Table E1: Corn yield estimation

Log Corn Yield (kg/ha)	
$\ln(F_{i,t})$	0.891**
	(0.415)
$\ln(L_{i,t})$	0.680
	(0.696)
$\ln(F_{i,t})^2$	-0.0125
	(0.0349)
$\ln(F_{i,t}) \times \ln(L_{i,t})$	-0.158
	(0.154)
Constant	4.891***
	(2.401)
Experiment Fixed Effect	Yes
Observations	139
Adjusted R-squared	0.419

Notes: Experiment plot clustered standard errors in parenthese. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Figure E1: Additional Analysis: Crop Yield Estimation

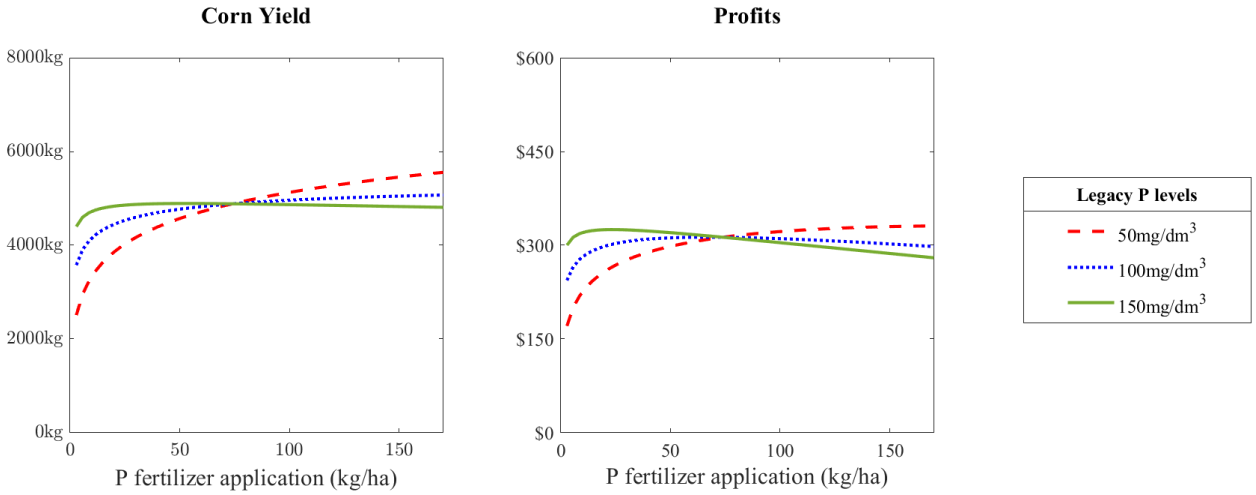
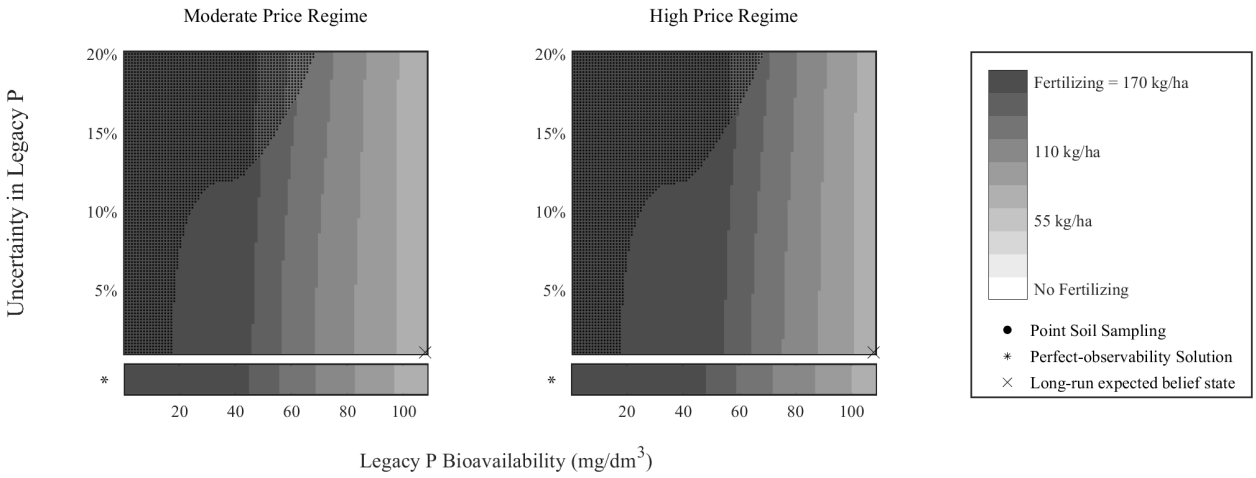
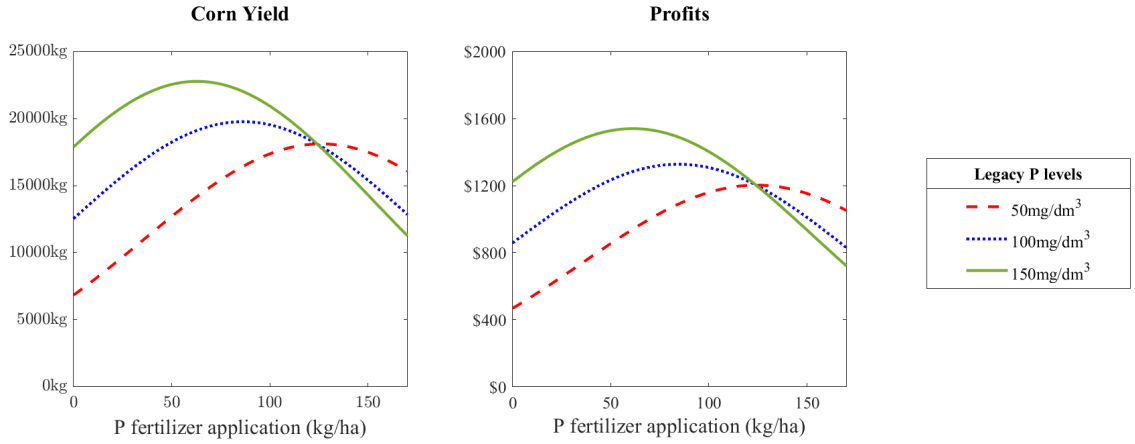


Figure E2: Additional Analysis: Crop Yield Estimation



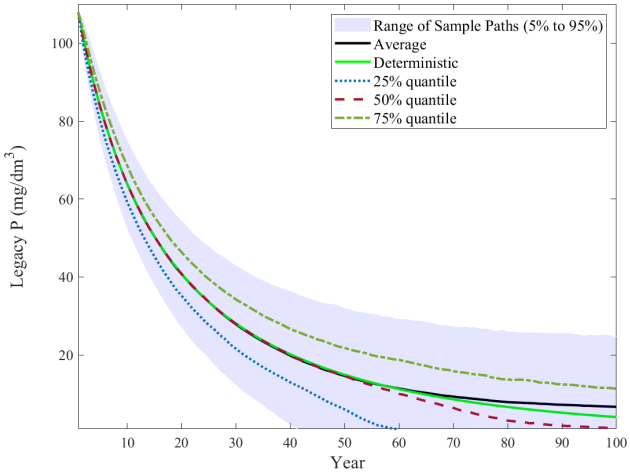
F Additional Figures and Tables

Figure F1: Example of corn yield estimation



Notes: This figure illustrates the relationship between P fertilizer application rates (kg/ha) and two key outcomes: corn yield (left panel) and profits (right panel), across three levels of legacy P concentrations: 50 mg/dm^3 (red dashed line), 100 mg/dm^3 (blue dotted line), and 150 mg/dm^3 (green solid line).

Figure F2: Legacy phosphorus accumulation without phosphorus fertilizer application



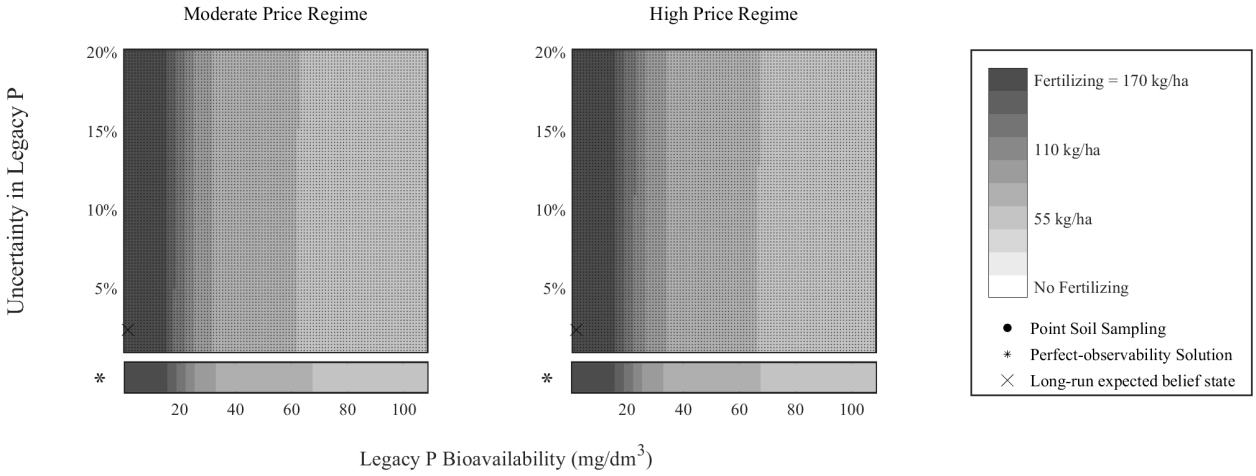
Notes: For the deterministic legacy P accumulation (green solid line), we employ a constant carry- over parameter $\rho_t = \rho = 0.98$ (2% decay rate) as adopted by Myyrä et al. (2007). The initial value is the 90th percentile (108 mg/dm^3) of legacy P in the North Carolina Tidewater data.

Table F1: Estimated value of risk aversion and elasticity of intertemporal substitution in literature

Literature		RA (η)	EIS (ψ)
Howitt et al. (2005)	California (US)	1.4	0.1
Lybbert and McPeak (2012)	Chalbi (Kenya)	0.5 (OLS) 0.8 (IV)	0.7(OLS) 0.9(IV)
	Dukana (Kenya)	13.5 (OLS) 12.5 (IV)	2.8(OLS) 3.3(IV)
Augeraud-Véron et al. (2019)		0.5-11	0.1-2
Cai and Lontzek (2019)		10	0.5, 1.5
Daniel et al. (2019)		1.1-15	0.6-1.2

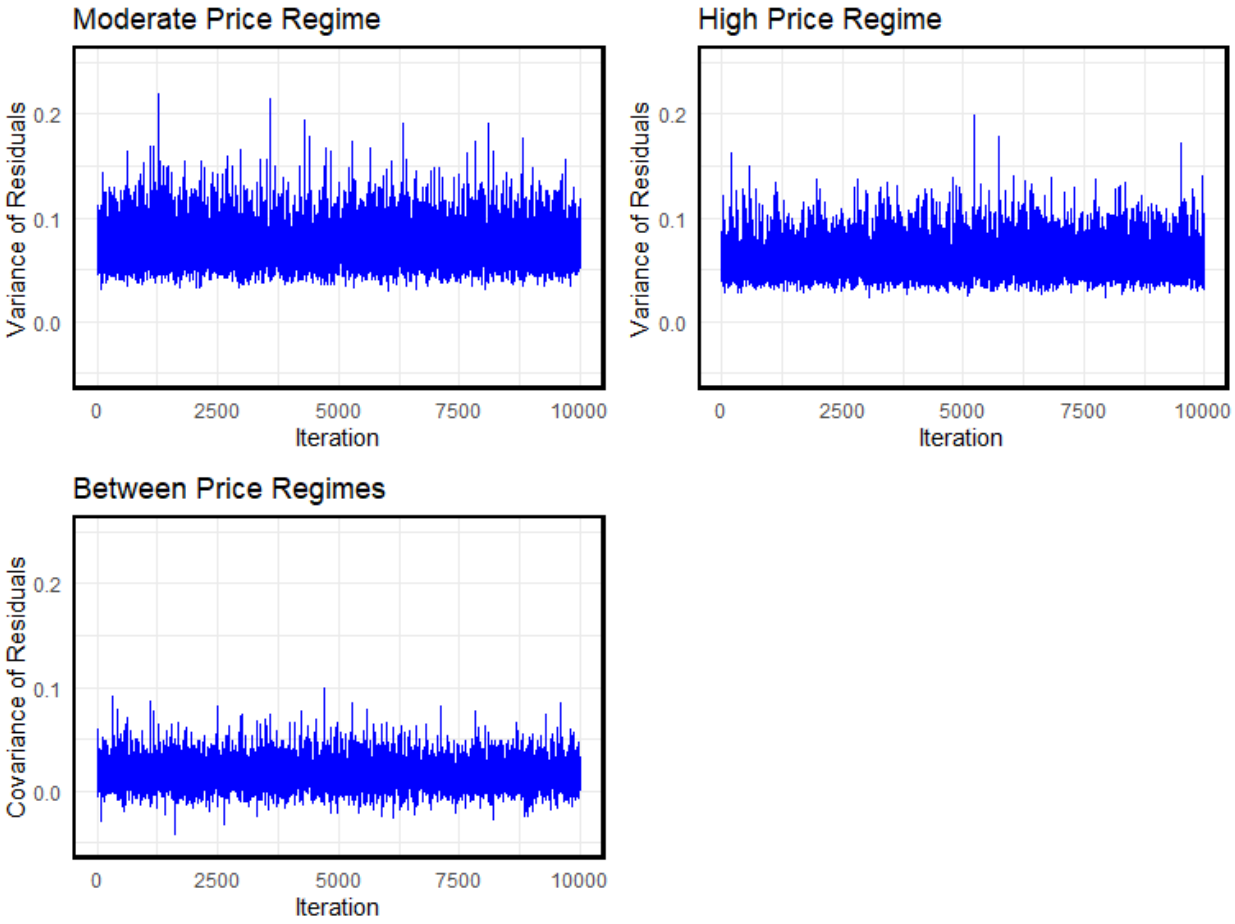
Notes: OLS and IV indicate Ordinary Least Squares regression and Instrumental variables estimation, respectively.

Figure F3: Optimal policy with fixed stochasticity in growth rate



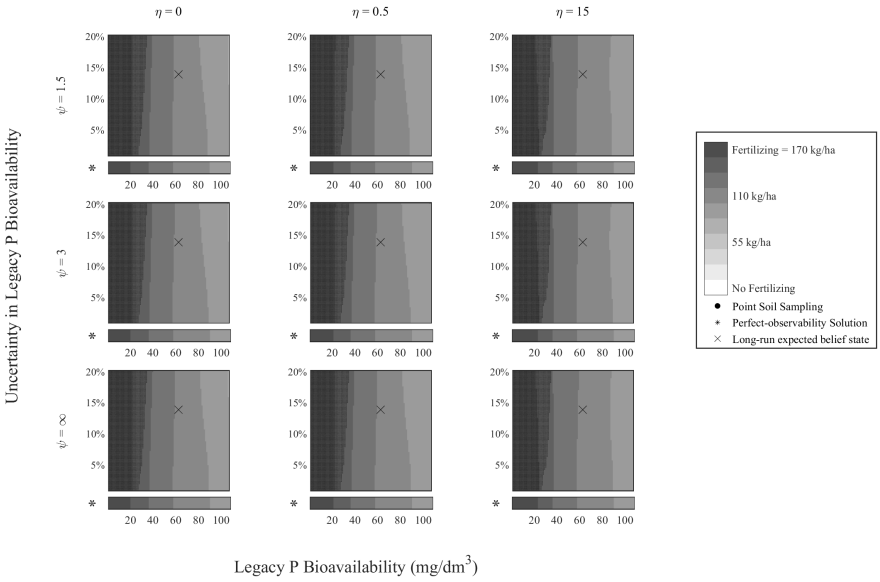
Notes: : In the original model, the standard deviation $\sigma_\rho(L)$ of the log percentage growth rate is inversely related to the legacy P level (equation 2), reflects an assumption in the model that more abundant legacy P stocks are assumed to be relatively more predictable in terms of their carry-over to the next period. Because we have no quantitative data with which to estimate the form of $\sigma_\rho(L)$, we investigate the effects of the alternative assumption that $\sigma_\rho(L) = \varsigma$ is fixed at uncertainty coefficient. This figure shows the model output derived under the alternative assumption. Given the uncertainty regardless in the dynamics scaling with legacy P levels, an optimal approach is to employ substantially more point sampling across state spaces.

Figure F4: Variance and Covariance of Residuals Across Price Regimes in MSVAR Model



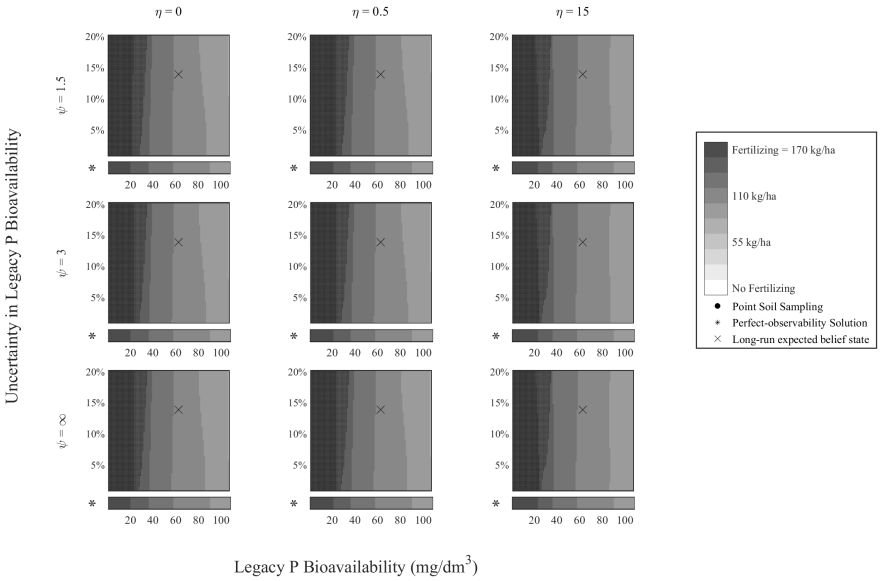
Notes: : The figure illustrates the variance and covariance of residuals for corn and phosphorus fertilizer price regimes over 10,000 iterations, based on a Markov-switching vector autoregression (MSVAR) model. The top left panel displays the variance of residuals for the moderate price regime, showing stable fluctuations. Similarly, the top right panel represents the variance of residuals for the high price regime, which follows a similar pattern of stabilization. Finally, the bottom panel shows the covariance of residuals between the moderate and high price regimes. The covariance exhibits more consistent fluctuations throughout the iterations. This figure highlights the dynamics of the model, with variances and covariances are consistent, indicating the model's convergence.

Figure F5: Risk analysis: Epstein-Zin preference (moderate price regime)



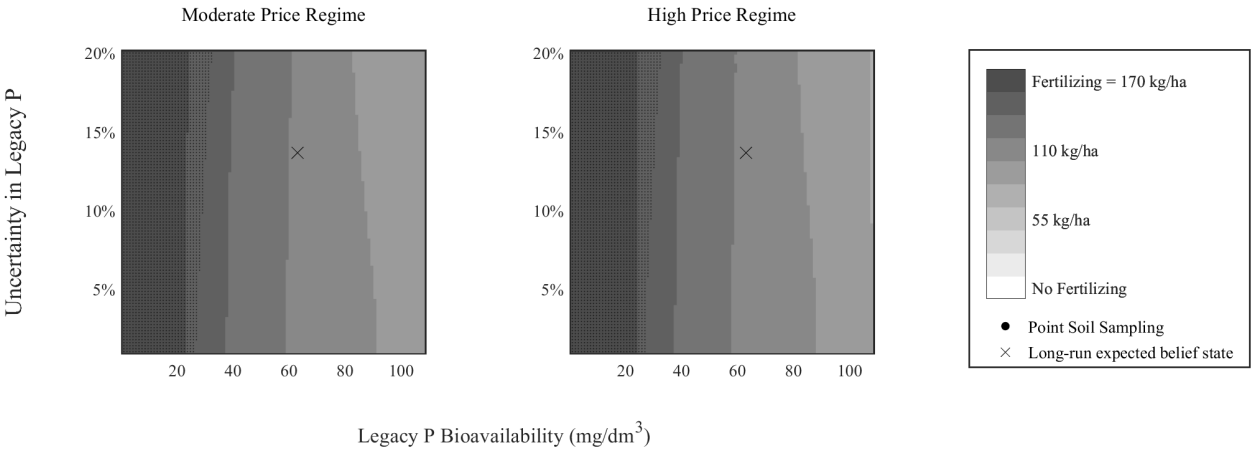
Notes: x-axis and y-axis indicate legacy P bioavailability and uncertainty in legacy P bioavailability, respectively. CRRA and IES parameters: $\eta = \psi^{-1} = 0$.

Figure F6: Risk analysis: Epstein-Zin preference (high price regime)



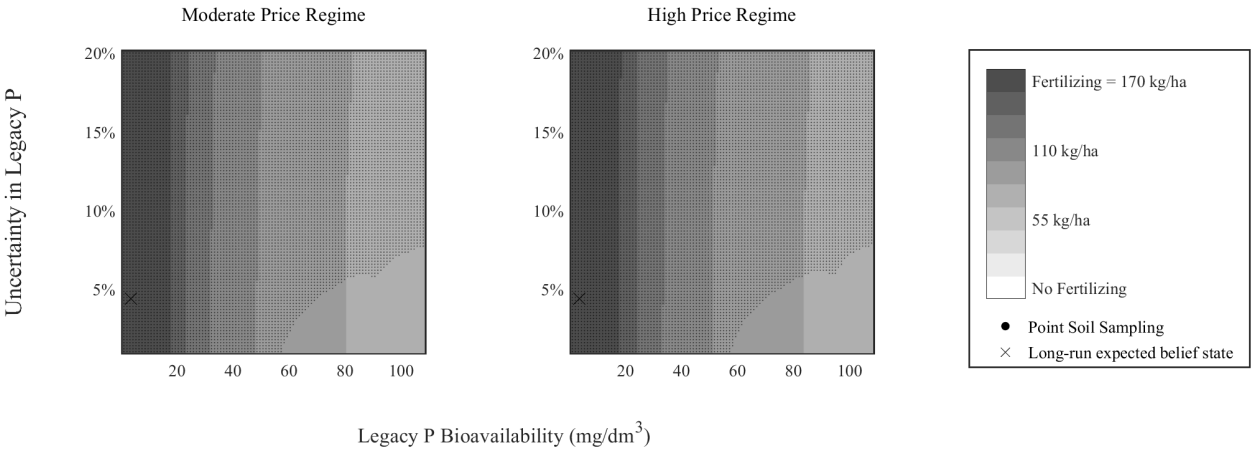
Notes: x-axis and y-axis indicate legacy P bioavailability and uncertainty in legacy P bioavailability, respectively. CRRA and IES parameters: $\eta = \psi^{-1} = 0$.

Figure F7: Sensitivity Analysis of P accumulation



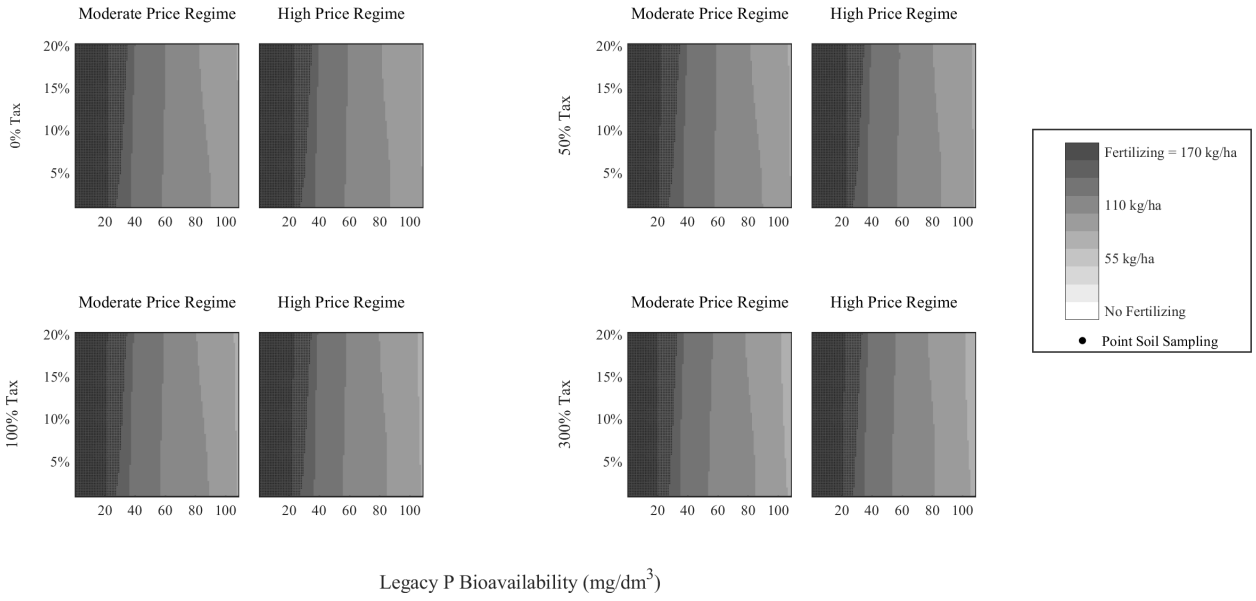
Notes: This figure presents the optimal fertilizer application and soil sampling strategy under a sensitivity scenario in which the P accumulation efficiency parameter (γ_1) is tripled. Compared to the baseline, this scenario results in more extensive fertilizer application, especially when the legacy P level is low, reflecting the increased long-term value of P fertilizer as a capital investment.

Figure F8: Sensitivity Analysis of P concentration



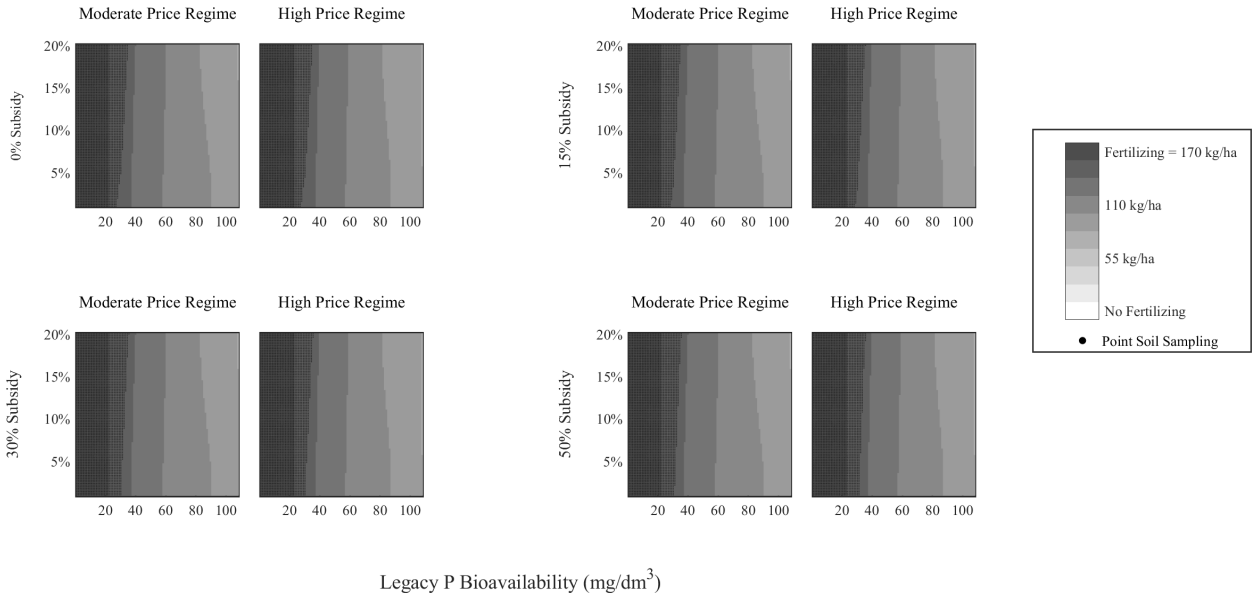
Notes: This figure presents the optimal fertilizer application and soil sampling strategy under a sensitivity scenario in which the phosphorus concentration parameter in crop yield (γ_3) is tripled. Increasing γ_3 raises the amount of P removed through harvest, making the system more depletable. As a result, farmers reduce fertilizer application, especially in low legacy P areas, and rely more on precise monitoring (point sampling) to avoid over- or under-application under greater volatility.

Figure F9: Sensitivity analysis: taxation on phosphorus fertilizer



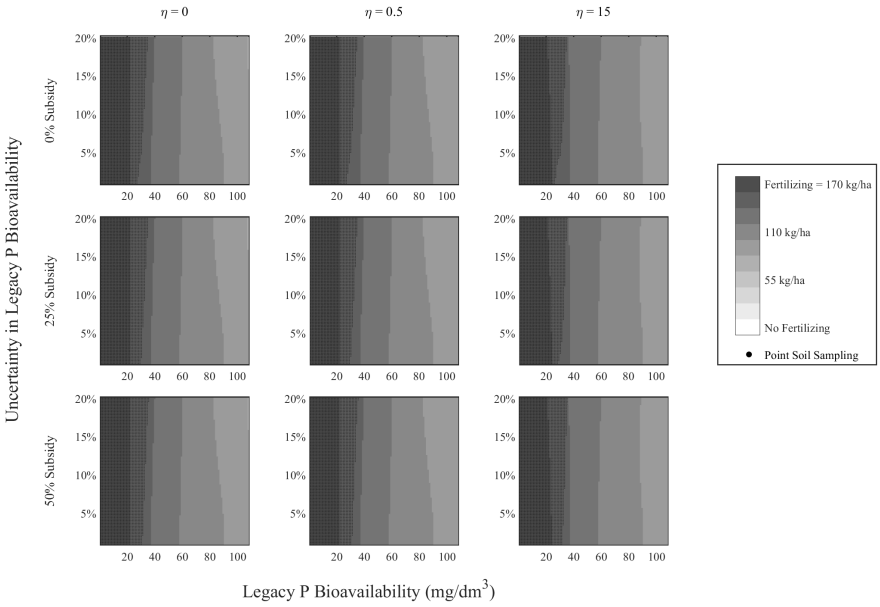
Notes: x-axis and y-axis indicate legacy P bioavailability and uncertainty in legacy P bioavailability, respectively.

Figure F10: Sensitivity analysis: subsidy on soil sampling



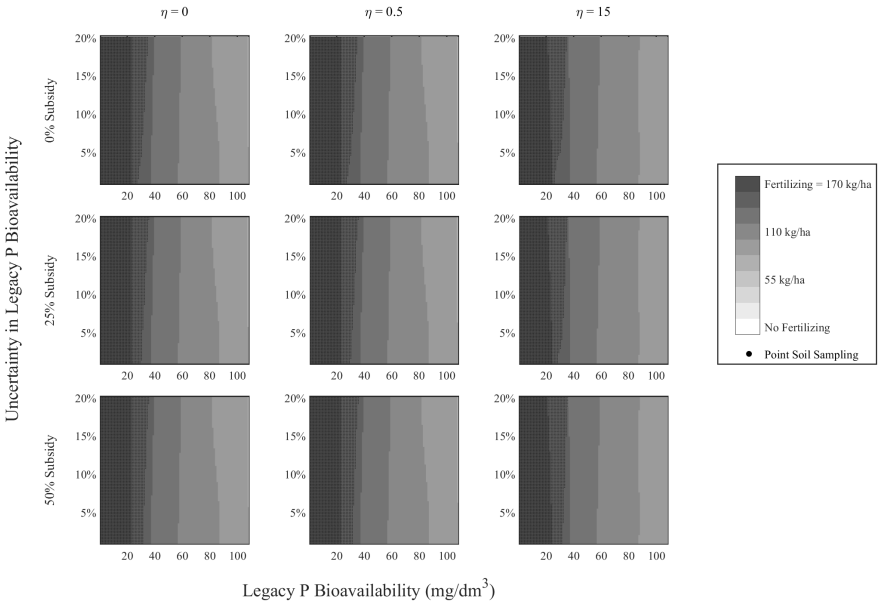
Notes: x-axis and y-axis indicate legacy P bioavailability and uncertainty in legacy P bioavailability, respectively.

Figure F11: Risk-averse farmer responses to soil sampling subsidy (moderate price regime)



Notes: x-axis and y-axis indicate legacy P bioavailability and uncertainty in legacy P bioavailability, respectively.

Figure F12: Risk-averse farmer responses to soil sampling subsidy (high price regime)



Notes: x-axis and y-axis indicate legacy P bioavailability and uncertainty in legacy P bioavailability, respectively.

References

ARULAMPALAM, M. S., S. MASKELL, N. GORDON, AND T. CLAPP (2002): "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on signal processing*, 50(2), 174–188.

AUGERAUD-VÉRON, E., G. FABBRI, AND K. SCHUBERT (2019): "The value of biodiversity as an insurance device," *American Journal of Agricultural Econoics*, 101(4), 1068–1081.

CAI, Y. AND T. S. LONTZEK (2019): "The social cost of carbon with economic and climate risks," *Journal of Political Economy*, 127(6), 2684–2734.

CZAJKOWSKI, M. AND W. BUDZIŃSKI (2019): "Simulation error in maximum likelihood estimation of discrete choice models," *Journal of Choice Modelling*, 31, 73–85.

DANIEL, K. D., R. B. LITTERMAN, AND G. WAGNER (2019): "Declining CO2 price paths," *Proceedings of the National Academy of Sciences*, 116(42), 20886–20891.

DE FREITAS, N. (2001): "Sequential Monte Carlo methods in practice (Vol. 1, No. 2, p. 2)," *A. Doucet, & N. J. Gordon (Eds.). New York: springer*.

HOWITT, R. E., S. MSANGI, A. REYNAUD, AND K. C. KNAPP (2005): "Estimating intertemporal preferences for Natural Resource Allocation," *American Journal of Agricultural Economics*, 87(4), 969–983.

KLING, D. M., J. N. SANCHIRICO, AND P. L. FACKLER (2017): "Optimal Monitoring and control under state uncertainty: Application to lionfish management," *Journal of Environmental Economics and Management*, 84, 223–245.

LYBBERT, T. J. AND J. MCPEAK (2012): "Risk and intertemporal substitution: Livestock portfolios and off-take among Kenyan pastoralists," *Journal of Development Economics*, 97(2), 415–426.

MATHWORKS. (2024): "Sobol Sequence, MATLAB Documentation." *Retrieved from https://www.mathworks.com/help/stats/sobolset.html*, accessed September 23, 2024.

MYYRä, S., K. PIETOLA, AND M. YLI-HALLA (2007): "Exploring long-term land improvements under land tenure insecurity," *Agricultural Systems*, 92(1-3), 63–75.

SLOGGY, M. R., D. M. KLING, AND A. J. PLANTINGA (2020): "Measure twice, cut once: Optimal Inventory and harvest under volume uncertainty and Stochastic Price Dynamics," *Journal of Environmental Economics and Management*, 103, 102357.

USDA (2024): "Fertilizer Use and Price," *United States Department of Agriculture*, https://www.ers.usda.gov/data-products/fertilizer-use-and-price/.

ZHOU, E., M. C. FU, AND S. I. MARCUS (2010): "Solving continuous-state POMDPs via density projection," *IEEE Transactions on Automatic Control*, 55(5), 1101–1116.