# 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 5 years of legacy P measurements from maize field trials spanning over a decade in eastern North Carolina, we develop a dynamic programming model of this optimization problem, with farmer decision-making and economic optimization specified as a 'partial-observability Markov decision process' (POMDP). 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, we find that more risk-averse optimizing agents in the model apply more fertilizer, across a range of bioavailable legacy P stocks. In sensitivity analysis we find that optimizing farmers in the model do not respond to sustained increases in stochastic P fertilizer prices or to decreased monitoring costs. We discuss the implications of these findings for policy discussions seeking to address environmental externalities of P fertilizer by providing better and cheaper information to farmers about 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 Processes

*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 for ensuring sustainable food production and environmental quality protection. P is imbalanced in the global food system, and some regions lacking 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 in other advanced economies) the main social challenge in P management is the excessive application of P fertilizer on farmlands, which contributes to water quality degradation and eutrophication in surface water systems. 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).

Notably, unlike nitrogen fertilizer, P fertilizer application residuals after crop take-up can accumulate in soils. This accumulating soil stock of P – referred to 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 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). Phoshorus 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, Paudel and Crago 2020). Downing et al. (2021) estimate a substantial economic cost associated with GHG emissions from eutrophication in freshwater system globally.

Various policies have been proposed to mitigate environmental issues arising from the overuse of P fertilizers, including the Numeric Nutrient Criteria under 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 policy idea 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 paradox 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 dynamis 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 the soil P level. Goetz and Zilberman (2000) examine the intertemporal and spatial 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 legacy P, and thus cannot answer the question 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).

To our knowledge POMDP methods have yet to be applied either in a depletable resource context or in farm production economics (though Sloggy et al.'s forestry application is adjacent to such a setting), reflecting 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 these methods.

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 extends the model into what is known as a 'mixed-observability Markov decision process' or MOMDP (Kovacs et al. 2012, Sloggy et al. 2020). Inclusion of these dynamics 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. In particular, 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. The specification introduces a high-dimensional state space (i.e. a space of probability distributions) 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) that reduce the dimensionality of the belief states, while avoiding some of the restrictions and pitfalls of prior methods (e.g. use of conjugate priors or coarse discretization of the unobserved state). We also use an econometric approach

in estimating price dynamics that aids numerical tractability in the MOMDP optimization that is still informed by the empirical analysis: 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 enhanced 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 tax on fertilizer) 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, outlining 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 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. Finally, the results of the model are discussed and are integrated with Epstein-Zin preferences, highlighting

the implications of risk preferences in shaping farmers' P fertilizer application and soil sampling decisions.

## 2 Model Description and Computational Methods

A simplified schematic of our POMDP model is shown in Figure 1, with the biophysical dynamics of legacy P stocks $L_t$ at the top level of the figure. The farmer does not observe $L_t$ but receives signals $O_t$ that depend on past soil sampling $s_{t-1}$, illustrated in the middle level of the figure. Farm production decisions regarding fertilizer applications $F_t$ and realized profit $\pi_t$ are also determined at this level. The bottom-level of the figure illustrates farmer inference regarding their unobserved legacy P stocks $L_t$, with beliefs $b_t$ being updated based on the signal $O_t$. The following subsections describe the structure and equations for each of these components, as well as the economic optimization problem to be solved.

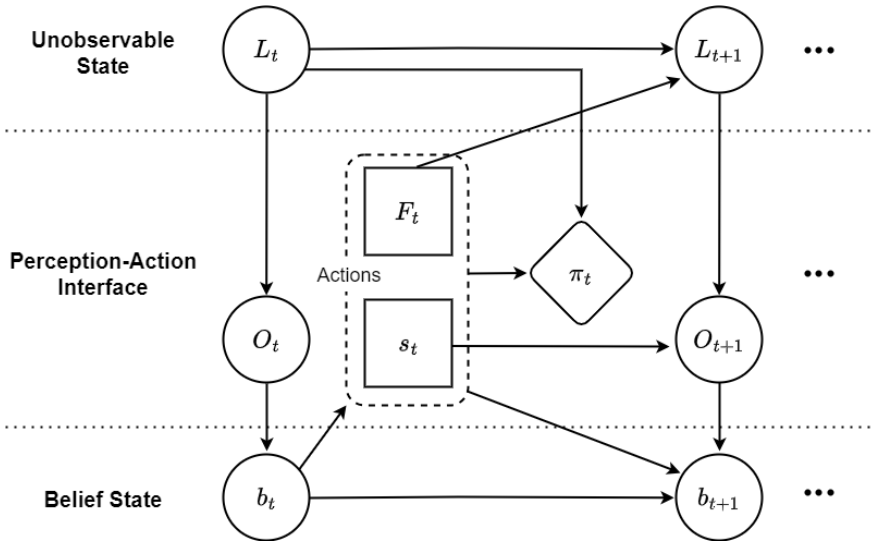
### 2.1 A Model of Stochastic Legacy Phosphorus Dynamics

We use a deterministic model of legacy P dynamics from Iho and Laukkanen (2012), to which we add stochastic behavior. The average soil-accumulated legacy P stock per hectare is given by $L_t$, with its dynamics specified in the following recursive equation:

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

where $\rho_t$ is a 'carry-over' parameter of legacy P, $F_t$ represents the amount of P fertilizer input, and $Y(L_t, F_t)$ is the crop yields at time $t$. The terms $(\gamma_3 \log(L_t) + \gamma_4)$ defines the legacy P concentration of

Figure 1: 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.

the crop yield, which increases logarithmically with $L_t$. As $L_t$ increases, the legacy P concentration also rise, initially leading to augmented yields. However, despite ongoing increases in $L_t$, the marginal yield gains attribute to each additional unit of legacy P progressively diminish. The term $(\gamma_1 + \gamma_2 L_t)$ is explain the change in legacy P brought about by a unit surplus or deficit in legacy P balance (Ekholm et al. 2005). The parameter values of $\gamma$ are summarized in Table 4.

While there are several empirically-grounded ways to introduce stochastic behavior 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, or a 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:

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

where $\mu_\rho$ is the log-mean of $\rho_t$ (so that $\mathbb{E}[\rho_t] = \exp\mu_\rho$) and $\sigma_\rho(L_t)$ is the standard deviation of log $\mu_\rho$, 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 diminished without added P fertilizer $F_t$.

Note the log-normal distribution of $\rho_t$ means that a fixed standard deviation $\sigma_\rho$ would result in the conditional variance of the annual change in legacy P stocks from 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 $\varsigma^2$, invariant with the current stock level $L_t$. This assumption implies the log standard deviation function takes the form $\sigma_\rho(L_t) = \sqrt{\ln(1+\varsigma^2\exp(-2\mu_\rho)/L_t^2)}$. 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 preliminary insights

into soil P content but serve as noisier indicators of the actual bioavailable legacy P stock across a 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).[1]

To specify the observation process, we denote the current soil sample test result as $O_t$. As a noisy measure of legacy P across the whole hectare of farmland, we assume a multiplicative test error $\lambda$ which is zero truncated and normally distributed with variance $\sigma_s^2$ determined by the type of sampling $s \in \{ss, ps\}$ (Kling et al. 2017).

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

The information gained from point v. 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 $\sigma_s^2$, the soil test may find zero 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 controls and resulting observations (Kling et al. 2017). Bayesian updating of these beliefs combines each period's prior beliefs regarding $L_t$, with projected dynamics for $L_{t+1}$, along with new information

[1]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.

$O^s_{t+1}$, 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 $O^s_t$.[2]

## 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, $\varepsilon_t$ 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 $P^Y_{t+1}$ will only be realized at the end of the season and not yet observed at the time of applying fertilizer. This means that the decision to apply

[2] 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.

fertilizer is informed by the current fertilizer price and the last harvest's crop price. Dynamics for the prices $\boldsymbol{P}_t = [P_t^Y, P_t^F]$ 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. We discuss the specific structure used for these dynamics below in econometric estimation for our application.

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 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:

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

where $\beta$ 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 conditional pdf of crop price $P_{t+1}^Y$ at the upcoming harvest, given the last observed harvest price $P_t^Y$ and 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:

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

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

$$\Pi_{EZ}(\boldsymbol{S}_t, F_t, s_t) \equiv \left( \iiint \pi(L_t, F_t, P^Y_{t+1}, P^F_t, s_t)^{1-\eta} f(P^Y_{t+1} \mid P^Y_t, P^F_t) b_t(L_t) g(\varepsilon^Y_t) d\varepsilon^Y_t \, dP^Y_{t+1} dL_t \right)^{\frac{1}{1-\eta}} \quad (9)$$

and where $\eta$ and $\psi$ indicate, respectively, the coefficient of relative risk aversion (RA) and the elasticity of intertemporal substitution (EIS): Higher $\eta$ and $\psi$ 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 the risk neutral agent with perfectly elastic

intertemporal substitution.[3]

## 2.4 Computational Methods: Density Projection and Particle Filtering

Note that one of the state variables in the above dynamic program is the continuous belief pdf $b_t(\cdot)$, which makes the model computationally intractable in its current form. Various methods have been proposed to address this common problem in POMDPs and related adaptive management applications, the simplest of which is to specify an initial prior belief pdf $b_0(\cdot)$ that is conjugate to likelihood function, so that $b_1(\cdot)$, $b_2(\cdot)$ and so for $b_t(\cdot)$ remain in the same family (e.g. normal distribution). This reduces the belief state from an infinite dimensional continuous pdf to a low-dimensional belief state corresponding to the parameters of that family (e.g. mean and variance of normal distribution). However, the use of conjugate priors is overly restrictive for most modern resource management problems, particularly in POMDP applications where the dynamics of the unobserved state variable $L_t$ need to be accounted for in belief updating, via the pdf $p(L_{t+1} \mid L_t, F_t)$ representing the stochastic transition dynamics.

To address this challenge, we follow the prevailing alternative in the resource economics literature involving density projection and particle filtering. The full algorithm used here is the same one employed by Kling et al. (2017) and Sloggy et al. (2020) in other resource management applications, and for completeness is detailed in the Appendix. In summary, the method first specifies a parametric distribution family for prior beliefs $b_t(L_t)$- here, a log-normal distribution, parameterized by a measure of central tendency and uncertainty: We parameterize the log-normal pdf here by its arithmetic mean $\mu^L$ and coefficient of variation $v^L$. The method then takes the

[3] 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 $V_{EZ}(\boldsymbol{S}) = (1-\beta)V(\boldsymbol{S})$.

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 particle filter. Density projection uses the Kullback-Liebler divergence as the distance measure between the approximating and prior pdfs. 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 only requires updating the mean and coefficient of variation.

This density projection projection procedure is integrated into computation of the dynamic programming solutions, by first discretizing the belief state parameters and actions $(\mu^L_t, \nu^L_t, F_t, s_t)$ and then calculating the discretized transition probabilities for the next-period belief parameters $(\mu^L_{t+1}, \nu^L_{t+1})$. 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. This illustrative case study represents This section describes the econometric estimation of model parameters for this context. The first subsection describes estimation of the

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.

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 re provided by (Morales et al. 2023). Table 1 provides the summary statistics from the field trial data. While the distribution of all measurements are 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:

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

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.

where $i$ denotes experiment plot, $\omega_t$ is the time fixed effect, and $\varepsilon_{i,t}^Y$ is a time-varying error component. Also, note that a $\ln(L_{i,t})^2$ 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 a panel setting, due to highly heterogeneous within-plot variation in legacy P 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 also estimate this function using OLS.

The results in Table 2 present the regression results. The positive sign for $F_{i,t}$ indicates that increasing P fertilizer leads to higher corn yields, suggesting that more fertilizer boosts productivity. However, the negative sign for the quadratic term $F_{i,t}^2$ implies diminishing marginal returns to P

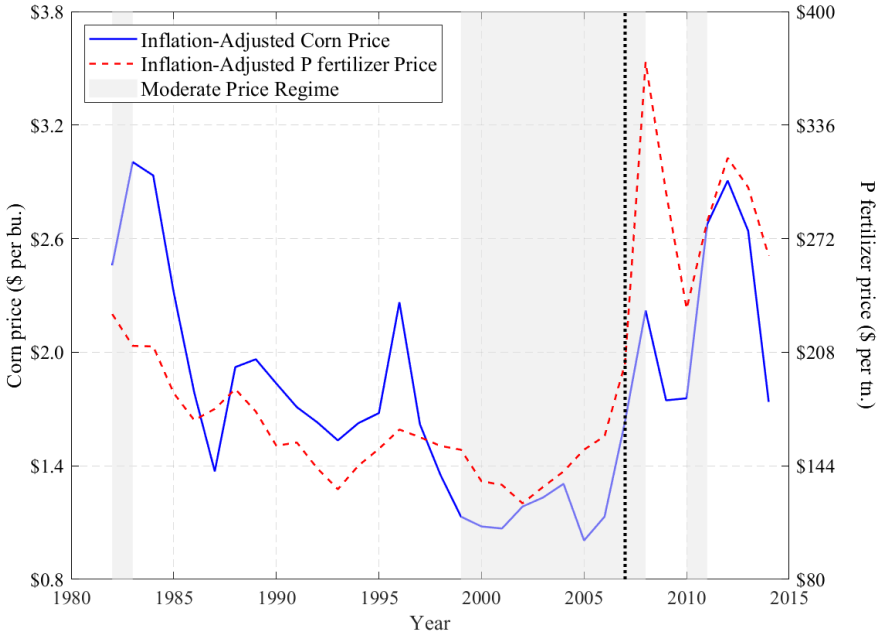
fertilizer application. As more fertilizer is applied, the incremental yield gains begin to decline, reflecting the principle of diminishing marginal productivity commonly seen in agricultural inputs.

For legacy P ($\ln(L_{i,t})$), the positive sign suggests that higher levels of legacy P could increase yields. The interaction term between $F_{i,t}$ and $L_{i,t}$ has a negative sign, indicating that when both P fertilizer and legacy P are present at high levels, they may act as substitutes, reducing each other's effectiveness. This could imply that as legacy P increases, the marginal benefit of applying additional P fertilizer decreases, which is consistent with the concept of nutrient saturation.

## 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 2). 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 MSVAR model. Markov-switching method generalizes the standard multivariate time-series vector autoregression model by allowing for probabilistic regime transitions 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 abrupt and sustained change in corn and P fertilizer price patterns after ca. 2007, as seen in Fig. 2. 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, consistent

Figure 2: Inflation-adjusted corn and phosphorus fertilizer prices




*Notes:* Inflation-adjusted prices are adjusted using the Consumer Price Index (CPI) for all urban consumer (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.

with a discrete change in market dynamics.

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

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

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 coefficient that vary depending on the regime $r_{t+1}$. $\boldsymbol{\mu}_{(r_{t+1})}$ is the regime-specific intercept, $\boldsymbol{\Sigma}$ is the covariance matrix of the error terms. In addition, the probability of regime $r_{t+1}$ can be specified as $p_{ij} = \Pr(r_{t+1} = i \mid r_t = j)$ where $p_{ij}$ represents the probability of transition from regime $j$ at time $t$ to regime $i$ at time $t+1$ (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) to estimate the transition probabilities between two regimes, moderate and high prices. Each regime is characterized by different autoregressive coefficients and covariance structures. We estimate the coefficients of the MSVAR model using a Hamiltonian Monte Carlo (HMC) method implemented via the Stan software (Osmundsen et al. 2021). Unlike traditional Gibbs sampling, HMC leverages gradient information to explore the posterior distribution efficiently, particularly in high-dimensional parameter spaces. This method fits our model, where the posterior distribution may exhibit complex geometry due to the mixture of regimes and regime transitions.

The likelihood of the model is constructed conditional on the latent state sequence, and prior distributions are placed on the model parameters, including the autoregressive coefficients, intercepts, and covariance matrices. Specifically, we use the priors on the intercept $\boldsymbol{\mu}_{(r_{t+1})} \sim \mathcal{N}(\mu_\mu, \sigma_\mu^2)$, priors on the autoregressive coefficients $\boldsymbol{\Phi}_{(r_{t+1})} \sim \mathcal{N}(\mu_\Phi, \sigma_\Phi^2)$, and priors on the covariance matrix $\boldsymbol{\Sigma} \sim \text{Wishart}(I, \nu)$, where $I$ is the identity matrix and $\nu$ is the degree of freedom.[4]

MSVAR results are presented in Tables 3. The results in Table 3 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 $\ln(P_t^Y)$ and $\ln(P_t^F)$. 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 roles of corn and P fertilizer. Corn, as a staple commodity, is more sensitive to input costs such as

[4]The mean ($\mu_\mu$, $\mu_\Phi$) and variance ($\sigma_\mu^2$, $\sigma_\Phi^2$) 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).

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, and are presented in Table 4. 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 value for the P dynamics carryover variance $\varsigma^2$, which we set at $\varsigma^2 = 9.21$ that comes from North Carolina Tidewater region data, where the variance of legacy P stock when no fertilizer was applied. This variance reflects natural fluctuations in legacy P levels due to environmental factors such as weather and soil processes, even without fertilizer input, justifying the use of the variance in the model.

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 expensive methodology to implement. Based on this information, we assumed that the observation error variance of point sampling ($\sigma_p$) 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 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, by which the variance of the mean sampled legacy P is proportional to the number of independent samples taken.

To orient readers with the biophysical dynamics from the parameterized model, Figure E2 in the Appendix displays the simulation results of legacy P accumulation (mg/dm$^3$) over 100 years without P fertilizer application, illustrating the range of stochastic paths.[5] The solid green line represents the deterministic path with 2% decay rate that assumes no uncertainty in legacy P dynamics. The shaded area represents the range of simulation sample paths from the 5% to 95% quantile, which becomes broader as the legacy P extends further into the future. Quantile lines for the 25% (blue

[5]The results depicted in Figure E2 were generated from 10,000 simulations.

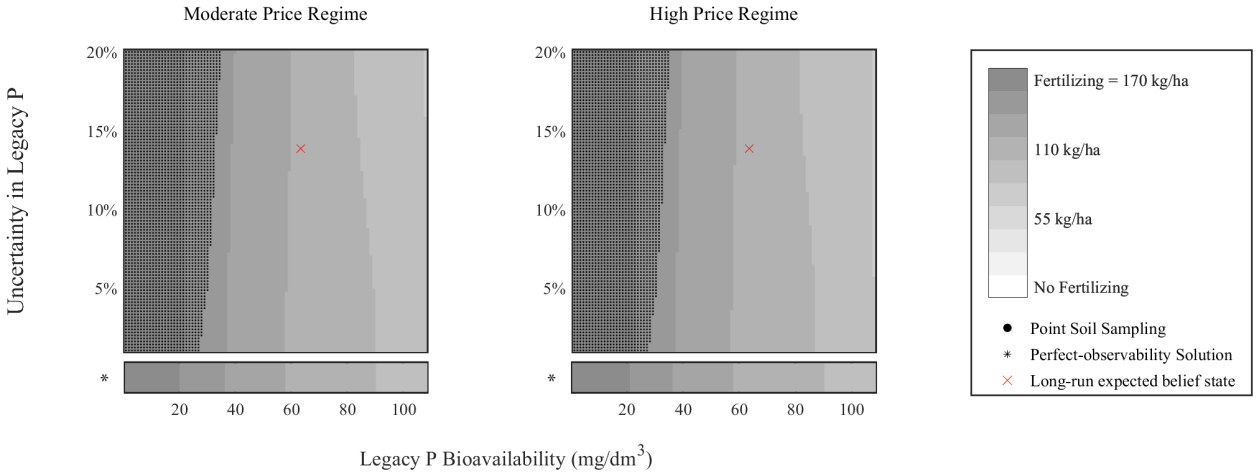
dots), 50% (red dash-dots), and 75% (green dashes) show the distribution of accumulation, with the 50% quantile also indicated as the median path. The black line represents the average of all simulation results.

The stochastic trend of legacy P dynamics follows closely to the deterministic path, suggesting that the parameters used in modeling legacy P dynamics and stochasticity do not deviate significantly from the deterministic trend. This consistency indicates that our model parameters effectively capture the essential dynamics of legacy P without substantial stochastic deviations. The light blue shaded area illustrates the variability and uncertainty in legacy P levels due to stochastic factors, showing a steady decline in legacy P, showing the gradual depletion of P reserves in the soil over time.

# 4 Model Results

In this section, we present and discuss representative solutions for the Mixed Observability Markov Decision Process (MOMDP) described above. We first present results for the risk-neutral farmer with perfectly elastic intertemporal substitution, before presenting results for the cases with risk aversion and intertemporal smoothing. The solutions presented here are meant to give the reader a sense of the qualitative features of optimal fertilizer application and soil sampling. The subsequent section then presents sensitivity analysis of the solutions with respect to key economic parameters of interest..

Figure 3: Optimal policy of P fertilizer application and soil sampling




## 4.1 Optimal Policy and Dynamics of Legacy Phosphorus

Figure 3 is composed of two graphs, each illustrating the optimal policy based on the bioavailability of legacy P, uncertainty, and the economic variables of the corn and P fertilizer prices. The horizontal axis 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.

When uncertainty in legacy P bioavailability is high and legacy P bioavailability is low, risk-neutral farmers in the model tend to apply more P fertilizer and are more likely to adopt soil sampling. The areas without dots indicate that farmers adopt standard soil sampling, while the dotted areas show where farmers opt for point sampling. When legacy P is low, farmers apply more P fertilizer to substitute for the low availability of P, ensuring sufficient nutrient supply for crop growth. Intuitively, point sampling is also more favored when estimated legacy P is low, because the precision of the soil test is more consequential in a P-limited setting: In such a setting, a test

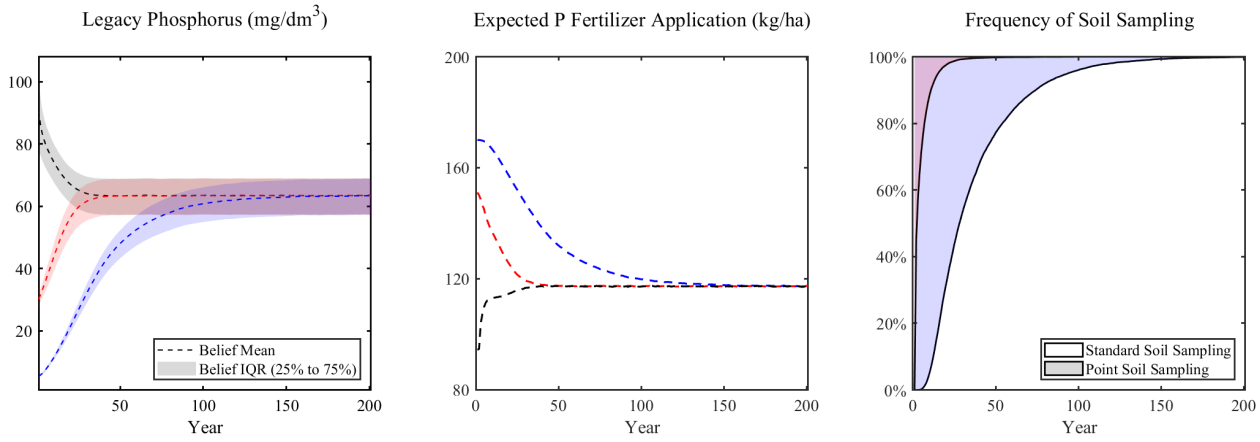
result that is used to determine optimal fertilizer application and yet is farther from the actual soil P level could have more severe consequences for yields.

The perfect observability solution at the bottom of each plot represents the scenario where farmers have perfect information about the amount of legacy P, obviating the need for soil (P) testing. The results for this scenario were derived using stochastic dynamic programming methods, which allow for optimal decision-making when the true state of legacy P is fully known. As theory dictates, as the coefficient of variation is reduced to a negligible levels, the solutions approach the perfect observability case.

We also analyzed the long-run equilibria of the model, as well as the dynamics of key variables in the presence of an optimizing farmer. In Figure 3, the "$\times$" mark represents the long-run expected belief state for legacy P. The position of this stable belief state is the 63mg/dm$^3$ level of legacy P bioavailability and 13% of uncertainty. That is, this model implies that in the long-run it is optimal for the risk-neutral farmer to apply P fertilizer so as to accumulate soil P. Figure 4, plotting moments of key variables from simulated timepaths, confirms this result, starting from three initial conditions of estimated legacy P (5mg/dm$^3$ (blue), 30mg/dm$^3$ (red), 90mg/dm$^3$ (black)). In approximately the first four decades of implementing the optimal MOMDP-derived solution, P fertilizer applications are high when the initial legacy P level is below than equilibria before transitioning to lower levels thereafter, with consequent changes to soil P stocks.

Figure 4 also shows that point sampling is favored only early in this transition period as soil P is built up, but in the long-run is completely abandoned to 100% standard sampling. This has predictable effects on the level of uncertainty in soil P, as reflected by the IQR shown in the shaded areas of the far left plot in Fig. 4: When point sampling is more favored, this IQR is small, but has a larger stable long-run range as standard sampling takes effect. These dynamics comport with the

Figure 4: Dynamics simulation of stochastic growth




*Notes:* The initial values for uncertainty is 20% for three cases. The initial conditions also include a moderate price regime. The figures were generated from simulations $i = 10,000$. The Belief IQR for a simulation $i$ at time $t$ is calculated as follows:
$\text{IQR}_{it} = \left[\exp(\mu^L_{it} + \sigma^L_{it}\Phi^{-1}(0.25)), \exp(\mu^L_{it} + \sigma^L_{it}\Phi^{-1}(0.75))\right]$ 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 $\text{IQR}_{it}$ over $i$.

optimal actions plotted in Fig. 3: As soil soil P accumulates to high levels, the relative value of more accurate information regarding soil P diminishes, since farmers can then be realtively sure of adequate soil P regardless of which test is used.

## 4.2 Risk Analysis: Epstein-Zin Preferences

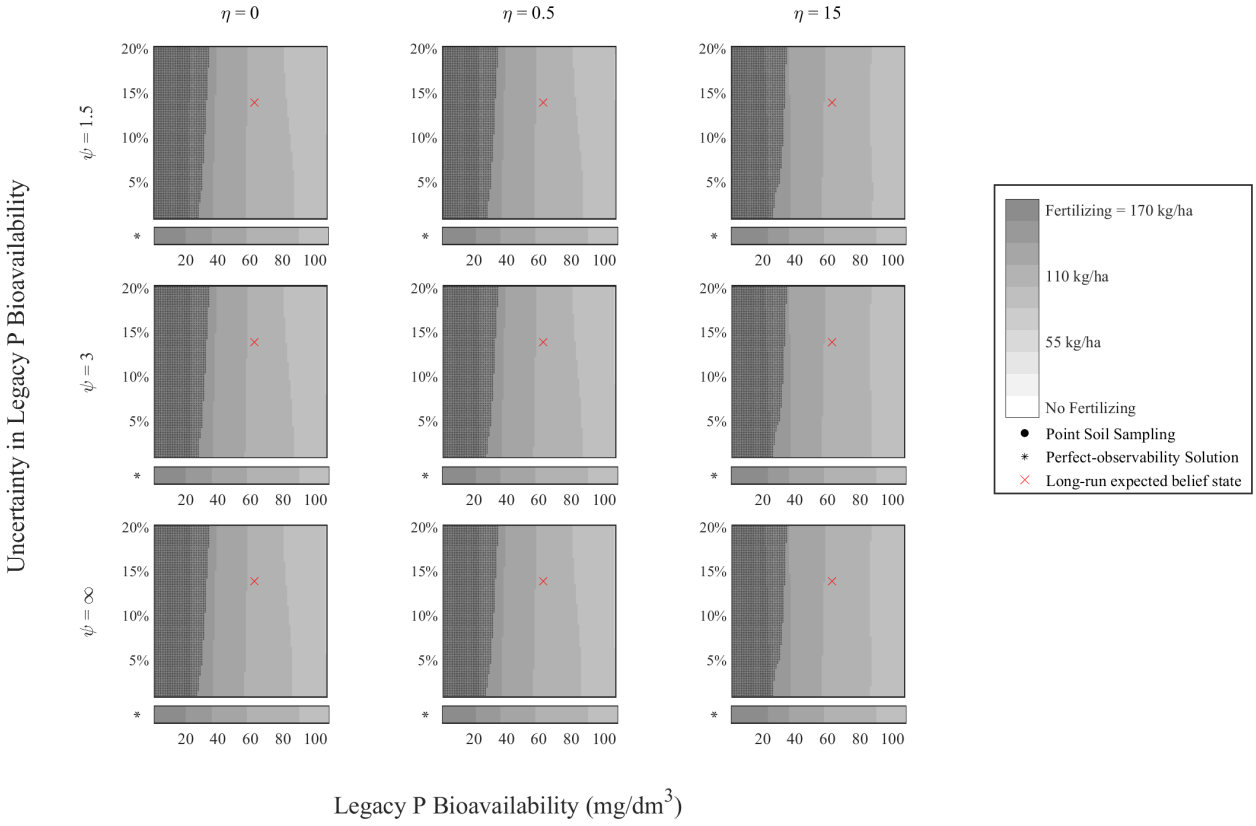
To compute solutions to the MOMDP with recursive Epstein-Zin (EZ) utility in (8) and (9), we require value ranges for the coefficient of relative risk aversion (CRRA, $\eta$) and the elasticity of intertemporal substitution (IES, $\psi$ ). Rather than solutions for one set of values for $\eta$ and $\psi$ , we are interested in how model results change across a reasonable ranges for these parameters, i.e. how does greater risk aversion or intertemporal smoothing affect optimal fertilizer applications and soil sampling? Since we have no data on farmer risk preference over time in this context that would have

permitted estimation of $\eta$ and $\psi$, we referred to the literature on environmental and agricultural studies listed in Table E1 in the Appendix. From this literature, we settled on $\eta \in (0.5, 15)$ and $\psi \in (1.5, 3)$. For easier comparison with the risk-neutral perfectly elastic IES solutions above, summary results for $\eta = 0$ and $\psi = \infty$, are included.

Figure 5 shows model results under EZ preferences. As $\eta$ increases from 0 to 15 (increasing risk aversion), the panels demonstrate that farmers apply more P fertilizer. As risk aversion increases, the shaded regions representing high fertilizer applications expand, especially legacy P bioavailability is high, as does the region where point sampling is employed.

This result aligns with the expectation that more risk-averse farmers would intensively utilize strategies to reduce uncertainty about their soil P status and ensure crops receive sufficient nutrients. The increased fertilizer application serves as a risk-mitigation strategy, reducing the likelihood of insufficient soil P affecting crop yields. Importantly, this model accounts for multiple sources of uncertainty and risk, including legacy P stocks, price fluctuations, and background yield variability, as highlighted in Equation (5). As risk aversion rises, the perceived risks associated with uncertain legacy P stocks become more significant, prompting farmers to increase fertilizer use to compensate for potential deficiencies. This leads to higher overall incentives to invest in productivity through fertilizer application. Moreover, because P fertilizer and soil P are dynamically linked substitutes, the increased fertilizer application reduces the relative importance of soil P uncertainty. Examining variations in the EIS ($\psi$ ) across the rows in Figure 5, one sees that the solutions change very little. This suggests that preferences for profit-smoothing play little role in determining the decisionmaker's optimal fertilizer and monitoring decisions.

Figure 5: Epstein-Zin preferences and optimal policy of P fertilizer application and soil sampling

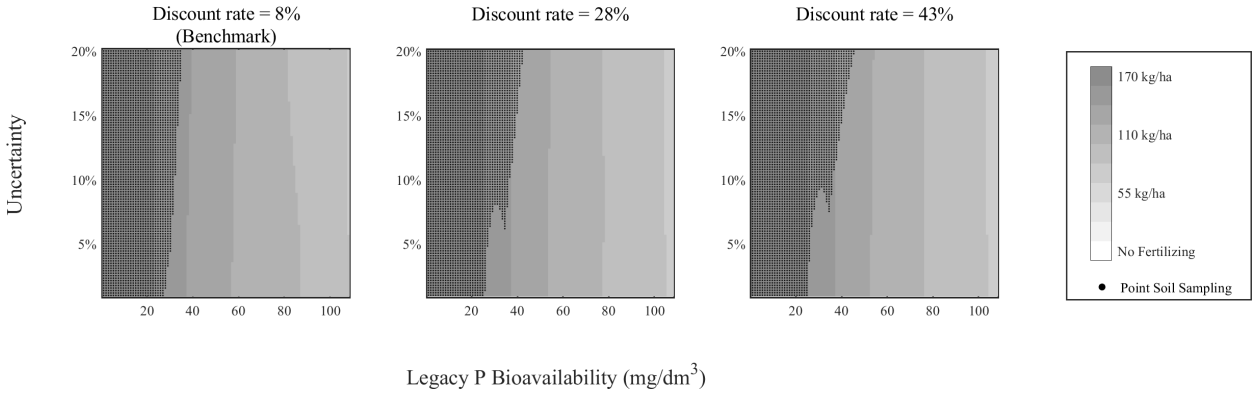



*Notes:* Results here are for the high-price regime. Results for additional low-price regime and other parameters scenarios are provided in the Appendix.

## 5 Sensitivity Analysis

Here we analyze how model solutions change with different economic parameters. We first examine the impact of the discount rate, a key parameter in dynamic economic models. We then analyze the effects of changes to P fertilizer prices (e.g. due to a persistent supply-side shock or a fertilizer tax (Sud 2020) and to the costs of enhanced soil testing (e.g. from a subsidy policy).

Figure 6: Sensitivity analysis: Discount rate




*Notes:* Initial price regime is high.

## 5.1 Discount Rate

In our sensitivity analysis, we selected three discount rates—8%, 28%, and 43%—to reflect a range of scenarios that 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 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.

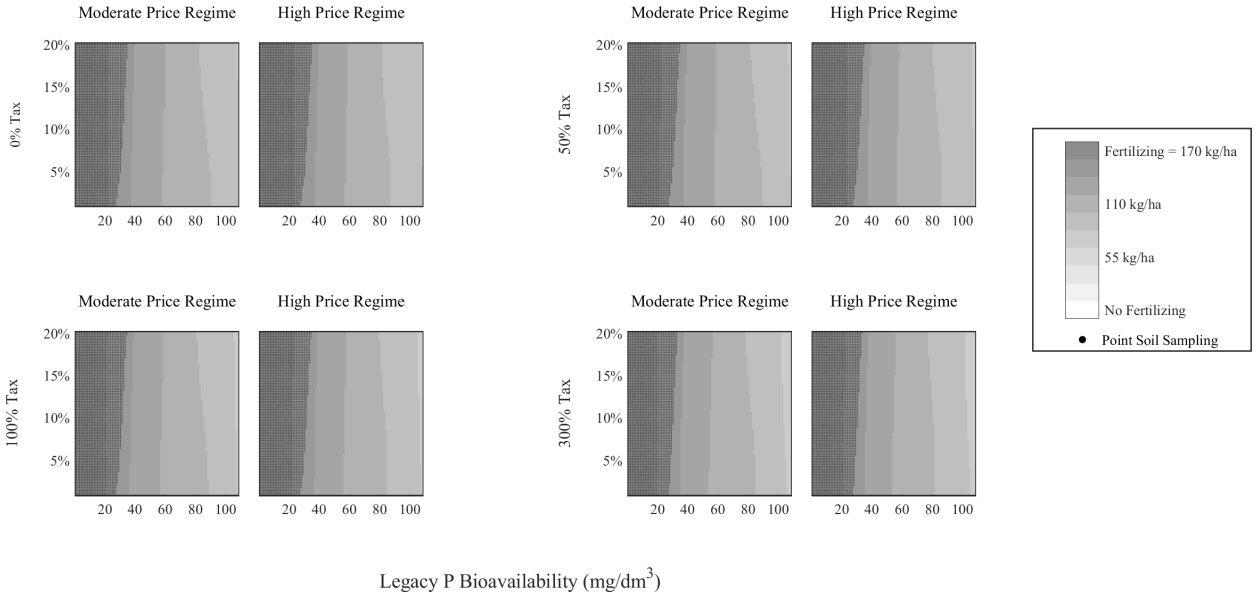
As shown in Figure 6, results are sensitive to the discount rate. Higher discount rates are associated with greater adoption of enhanced soil testing through point sampling, while fertilizer use changes very little. This suggests that farmers with high discount rates prioritize obtaining precise information about their soil P status, which enables them to apply fertilizer more accurately and efficiently rather than over-applying as a hedge against uncertainty. This behavior reflects a dynamic tradeoff where the immediate value of accurate soil information becomes critical for optimizing

fertilizer decisions under higher discount rates. By relying on point sampling, farmers reduce uncertainty about their soil P status, which allows them to fine-tune their applications without the need for excessive fertilizer use. These findings suggest that high discount rates drive farmers to adopt strategies that maximize current profitability while minimizing risks through better-informed decisions.

## 5.2 Persistent Increases in P Fertilizer Prices

Here, 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. Sun et al. 2016). 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

Figure 7: Sensitivity analysis: taxation on phosphorus fertilizer




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

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%. For exposition, we label these increases as (*ad valorem*) taxes, and modify the fertilizer price state variable as follows: $P^F_{\text{tax}} = P^F \cdot (1 + \textit{Tax Rate})$, where $P^F$ is the producer price, and $P^F_{\text{tax}}$ denotes the price of P fertilizer paid by farmers. Figure 7 illustrates how alternative price scenarios affect model results for the baseline case of a risk-neutral farmer without a preference for intertemporal smoothing. These results show that while P fertilizer application decreases with increasing soil P levels, the reduction in fertilizer use becomes less pronounced as fertilizer taxes rise. At higher prices (tax rates), farmers seem less inclined to reduce fertilizer application, likely to maintain sufficient nutrient levels for crop production. Interestingly, enhanced soil testing remains relatively unchanged across different

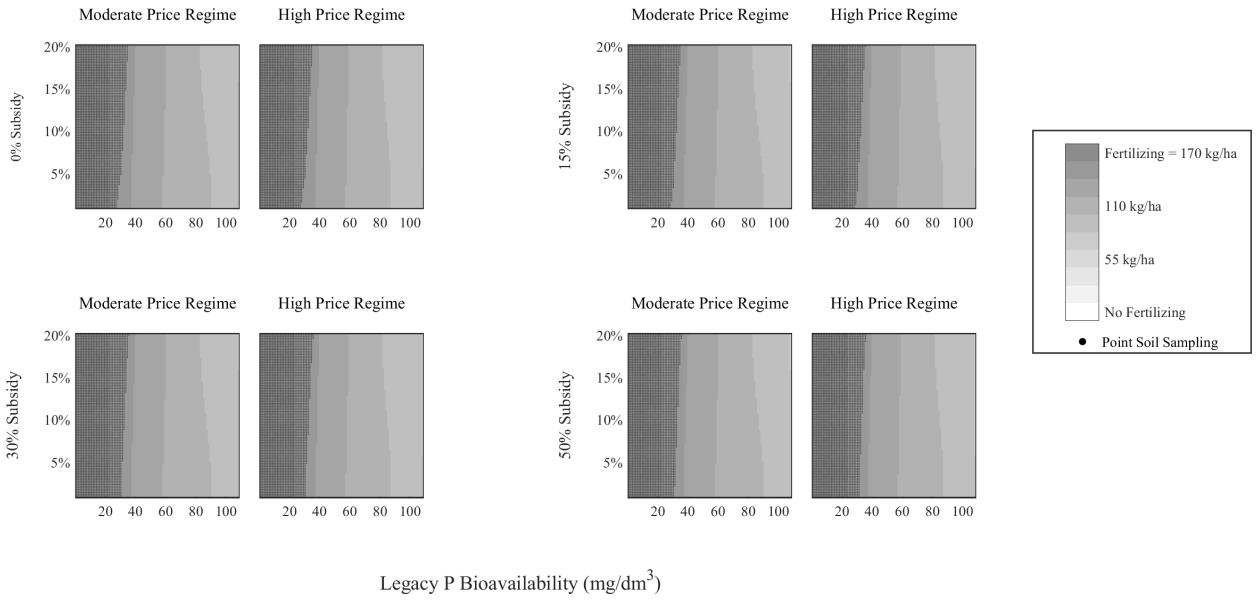
price scenarios. This suggests that while fertilizer taxes may influence the overall level of fertilizer application, they do not significantly alter the perceived value of soil testing. Farmers continue to utilize soil testing as a consistent strategy to inform their decisions, even as fertilizer costs rise. This highlights the complementary role of soil testing in maintaining productivity under varying price conditions, irrespective of tax rate changes.

## 5.3 Reduced Soil Sampling Costs

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

Figure 8 shows the effects of soil testing subsidies for the baseline case with a risk-neutral farmer with no preference for intertemporal smoothing. As expected, the figure shows that higher subsidies for enhanced testing increases adoption. At high subsidy, the decisionmaker more employs point sampling. 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. In robustness checks in the Appendix, we confirm that the presence of risk aversion does not alter these findings.

Figure 8: Sensitivity analysis: subsidy on soil sampling




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

# 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. Because of the dynamic and stochastic nature of P accumulation in soil, combined with state uncertainty about legacy P stocks, this research adopts a model-based approach to disentangling these dynamics and their effects on the fertilizer demand and soil sampling behaviors of farmers. 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.

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 more 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


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*.

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.

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.

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.

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.

KOVACS, D. L., W. LI, N. FUKUTA, AND T. WATANABE (2012): "Mixed Observability Markov decision processes for overall network performance optimization in Wireless Sensor Networks," *2012 IEEE 26th International Conference on Advanced Information Networking and Applications*.

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.

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

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.

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.

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/.

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.

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 ($\Sigma_{11}$) | 0.072 | Variance ($\Sigma_{22}$) | 0.059 | |
| | (0.0002) | | (0.0002) | |
| Covariance ($\Sigma_{12} = \Sigma_{21}$) | 0.0165 | | | |
| | (0.0001) | | | |

*Transition Probabilities*

| | Moderate ($t+1$) | 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.

Table 4: Parameters and description

| | Value | Desccription |
|---|---|---|
| Biological Parameters | | |
| $\mu_\rho$ | -0.02 | Average rate of growth (Myyrä et al. 2007) |
| $\varsigma^2$ | 9.21 | Carryover variance |
| $\gamma_1$ | 0.0032 | Response parameter of legacy P surplus (Ekholm et al. 2005) |
| $\gamma_2$ | 0.00084 | |
| $\gamma_3$ | 0.000186 | Concentration parameter on crop yield (Iho and Laukkanen 2012, Saarela et al. 1995) |
| $\gamma_4$ | 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 |
| $\beta$ | 0.9259 | Discount factor with 8% discount rate (Duquette et al. 2012) |
| $\sigma_{ss}$ | 0.638 | Observation error of standard soil sampling (Lauzon et al. 2005) |
| $\sigma_p$ | 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).

# Appendix

## Evaluating Optimal Farm Management of Phosphorus Fertilizer Inputs with Partial Observability of Legacy Soil Stocks[1]

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



[1]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).

# A Methodology and Algorithms

This appendix provides additional details and elaborates on the methodology described in the paper. It begins by outlining the foundation 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 model (MSVAR), and concludes with supplementary figures.

## A.1 Detailed formulation of the Dynamic Programming Model

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

$$\begin{aligned} V(\boldsymbol{S_t}) = &\max_{F,s} \iiint \pi(\boldsymbol{S_t}, F_t, s_t) f(P^Y_{t+1} \mid P^Y_t, P^F_t) b_t(L_t) g(\epsilon^Y_t) d\epsilon^Y_t \, dP^Y_{t+1} dL_t \\ &+ \beta \iint p(\boldsymbol{P}^{(r)}_{t+1} \mid \boldsymbol{P}^{(r)}_t) p(O^s_{t+1} \mid b_{t+1}(L_t), s_t) V(\boldsymbol{S_{t+1}}) dO^s_{t+1} d\boldsymbol{P}^{(r)}_{t+1}, \end{aligned} \tag{A1}$$

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

$$\begin{aligned} V_{EZ}(\boldsymbol{S_t}) = &\max_{F,s} \Bigg[ (1-\beta) \Bigg( \iiint \pi(\boldsymbol{S_t}, F_t, s_t)^{1-\eta} f(P^Y_{t+1} \mid P^Y_t, P^F_t) b_t(L_t) g(\epsilon^Y_t) d\epsilon^Y_t \, dP^Y_{t+1} dL_t \Bigg)^{\frac{1-\psi^{-1}}{1-\eta}} \\ &+ \beta \Bigg( \iint p(\boldsymbol{P}^{(r)}_{t+1} \mid \boldsymbol{P}^{(r)}_t) p(O^s_{t+1} \mid b_{t+1}(L_{t+1}), s_t) V_{EZ}(\boldsymbol{S_{t+1}})^{1-\eta} dO^s_{t+1} d\boldsymbol{P}^{(r)}_{t+1} \Bigg)^{\frac{1-\psi^{-1}}{1-\eta}} \Bigg]^{\frac{1}{1-\psi^{-1}}}, \end{aligned} \tag{A2}$$

under the state variables, $\boldsymbol{S_t} \equiv [b_t(L_t), P^Y_t, P^F_t]$, where $b_t(L_t)$ represents the farmer's belief about the latent legacy phosphorus (P) state at time $t$, $P^Y_t$ denotes the market price of corn at time $t$, and $P^F_t$ denotes the price of P fertilizer at time $t$. Additionally, the notation $\boldsymbol{P}^{(r)}_t$ refers to the set of corn and P fertilizer prices, $[P^Y_t, P^F_t]$, 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

Continuous state POMDP has challenges due to an infinite-dimensional belief space and because approximating belief states by discretization can lead to computational issues. Exact evaluation of the posterior distribution is difficult to address, and even structuring the belief updating process in 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 parameterized family of densities.[2] Projection mapping from the belief state $b(L)$ to exponential family of density $f(L;\theta)$, where $\theta$ 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:

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

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

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). 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 $\phi$ 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 $\theta_t$ to $\theta_{t+1}$ 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 $\theta$ 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.[3] This conversion involves discretizing the space of natural parameters $\theta$ 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 $\mu_L$ and uncertainty in legacy P bioavailability as coefficient variation $\nu_L$ ($\nu_L = \sigma_L/\mu_L$) with a parameter set $\delta = \{\mu_L, \nu_L\}$. Hence, we discretize $\theta$ by calculating the univariate log-normal parameters $\mu$ and $\sigma$ that $\theta = \{\mu, \sigma\}$ where $\sigma > 0$ from the $\delta$ (Kling et al. 2017). The calculation of $\mu$ and $\sigma$ 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, $\mu_L$ and $\sigma_L$ are discretized into a $100 \times 1$ vectors. A

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

$100 \times 100$ 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 $\delta_i$, the crop profit function is evaluated as the expected value of $\delta_i$, 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 $\delta_i$ to $\delta_j$, 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.[4]

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 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

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

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 $\rho$ 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 eq. 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 $\phi$ 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.

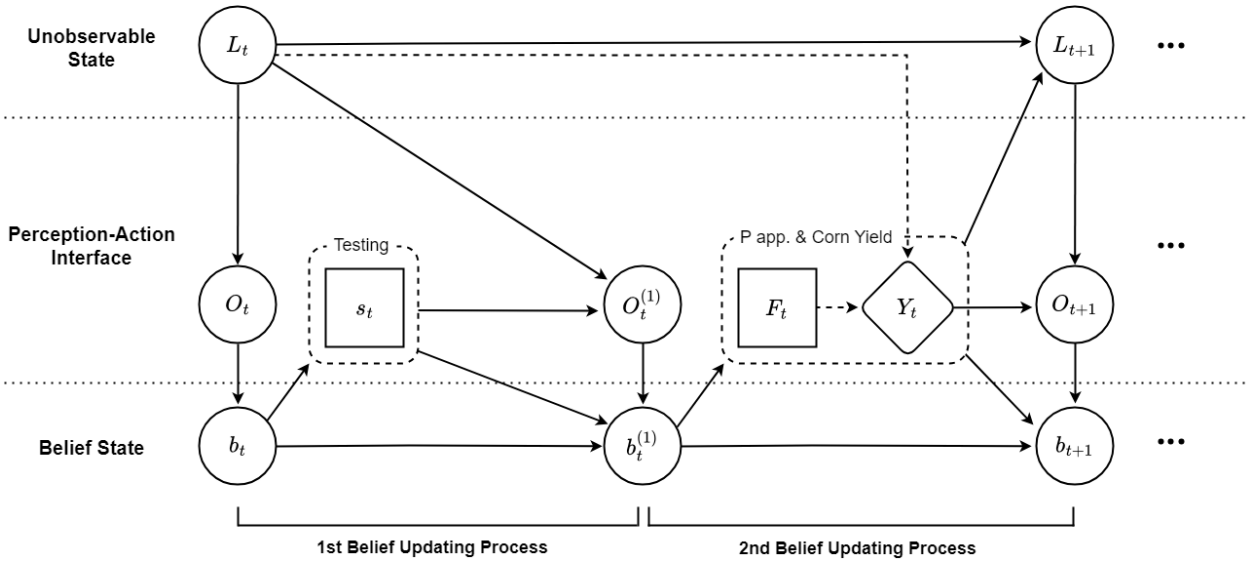
The first stage of belief updating occurs when farmers conduct soil sampling before making fertilization decisions. The second stage of belief updating takes place after the fertilization and harvest, when farmers receive additional information through the actual corn yield. This yield data, reflecting 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.

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 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 $O_t$ 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.

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 $\theta_i$ following Algorithm 1.

These Sobol points are generated from a log-normal distribution, resulting in yield samples

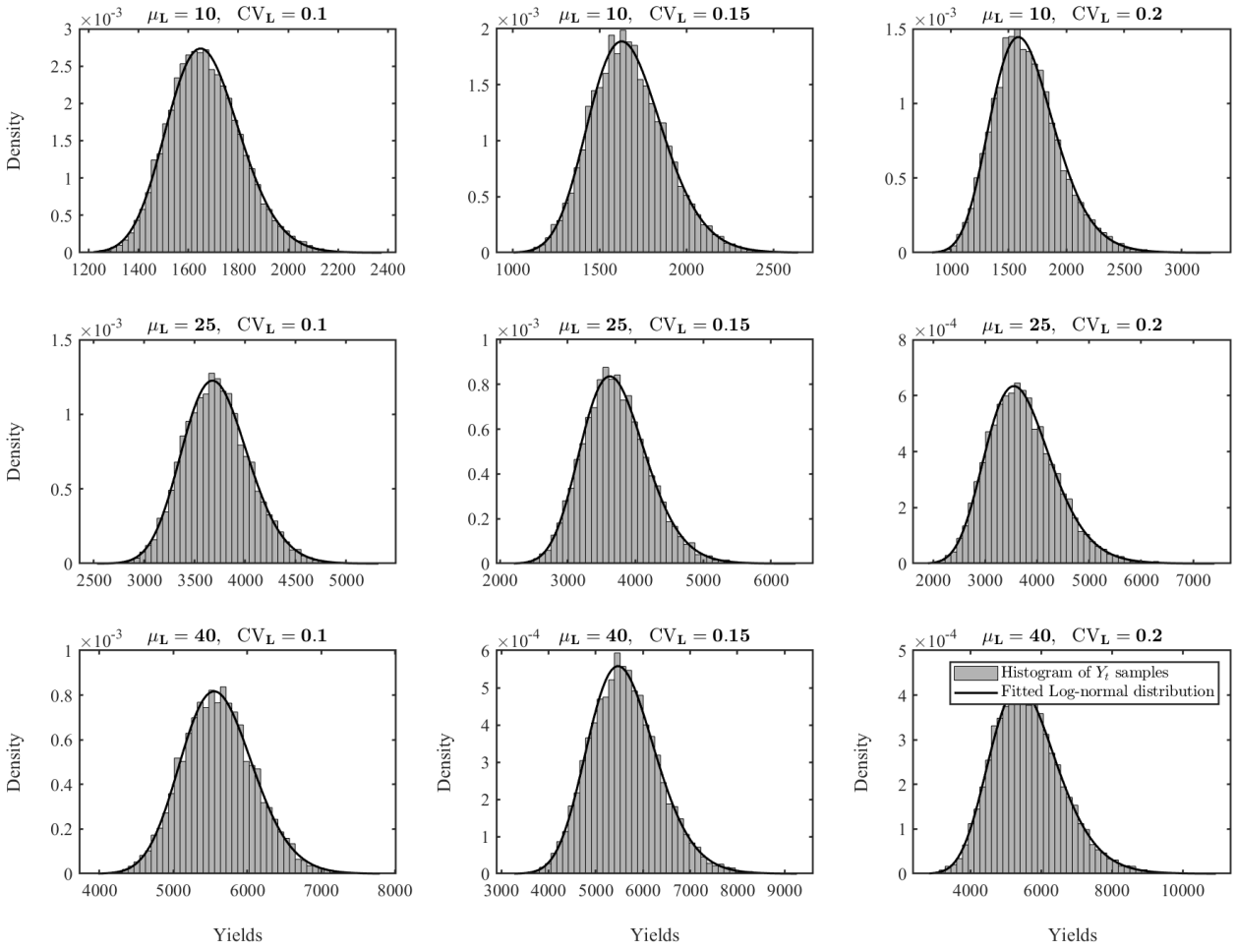
$Y_t$ that align with a log-normal distribution, as shown in Figure B2. The fitted log-normal curve effectively captures this distribution, accurately representing the central tendency, variability, and skewness inherent in the yield simulations. Defining the form of the yield distribution is essential 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, Figure 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

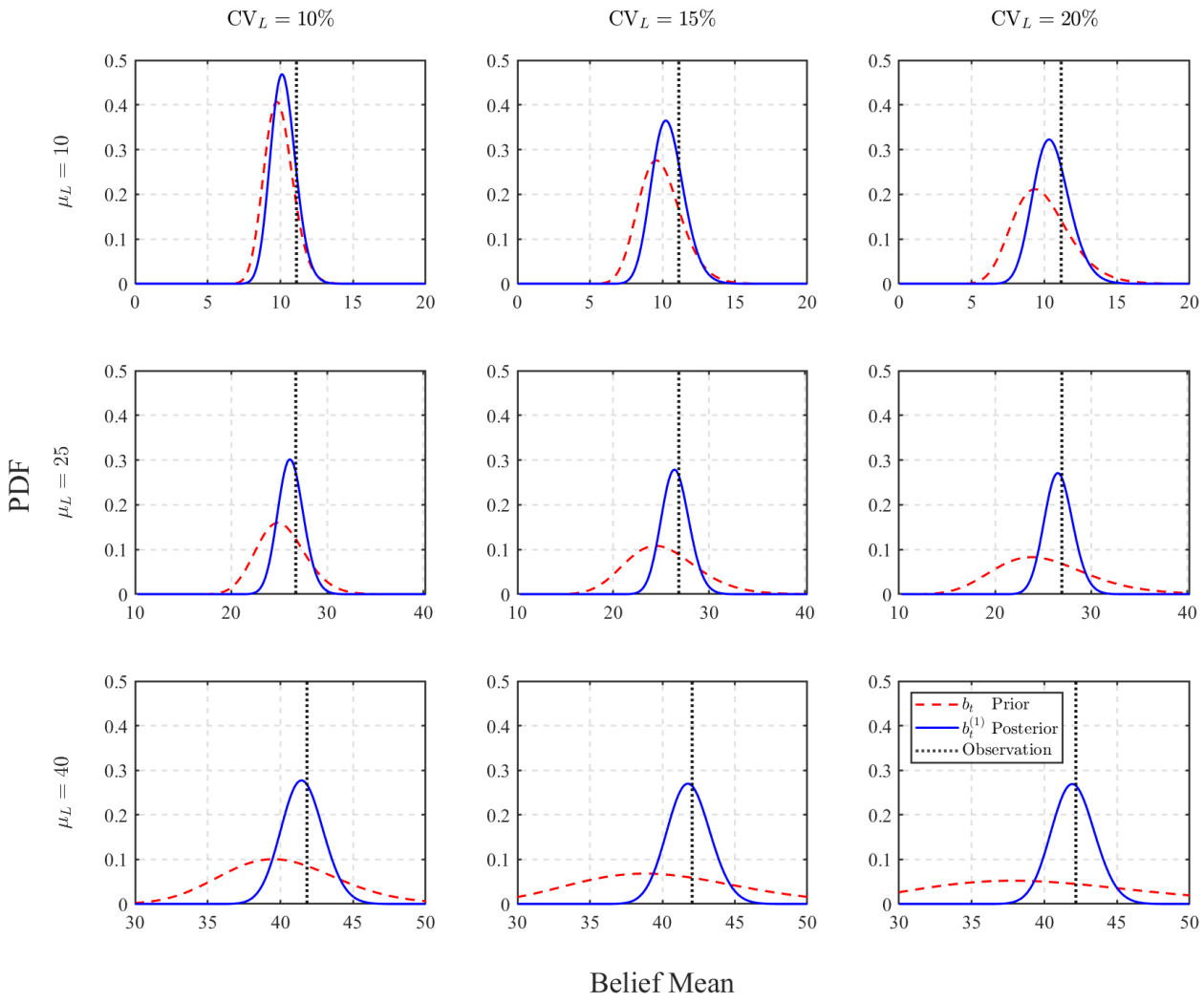
understanding of the legacy P levels compared to the soil sampling. This analysis demonstrates the critical role of soil sampling in the belief-updating process, particularly in the first stage, and in our paper, we consider soil sampling for the belief-updating process.

Figure B2: Distributions of Corn Yields



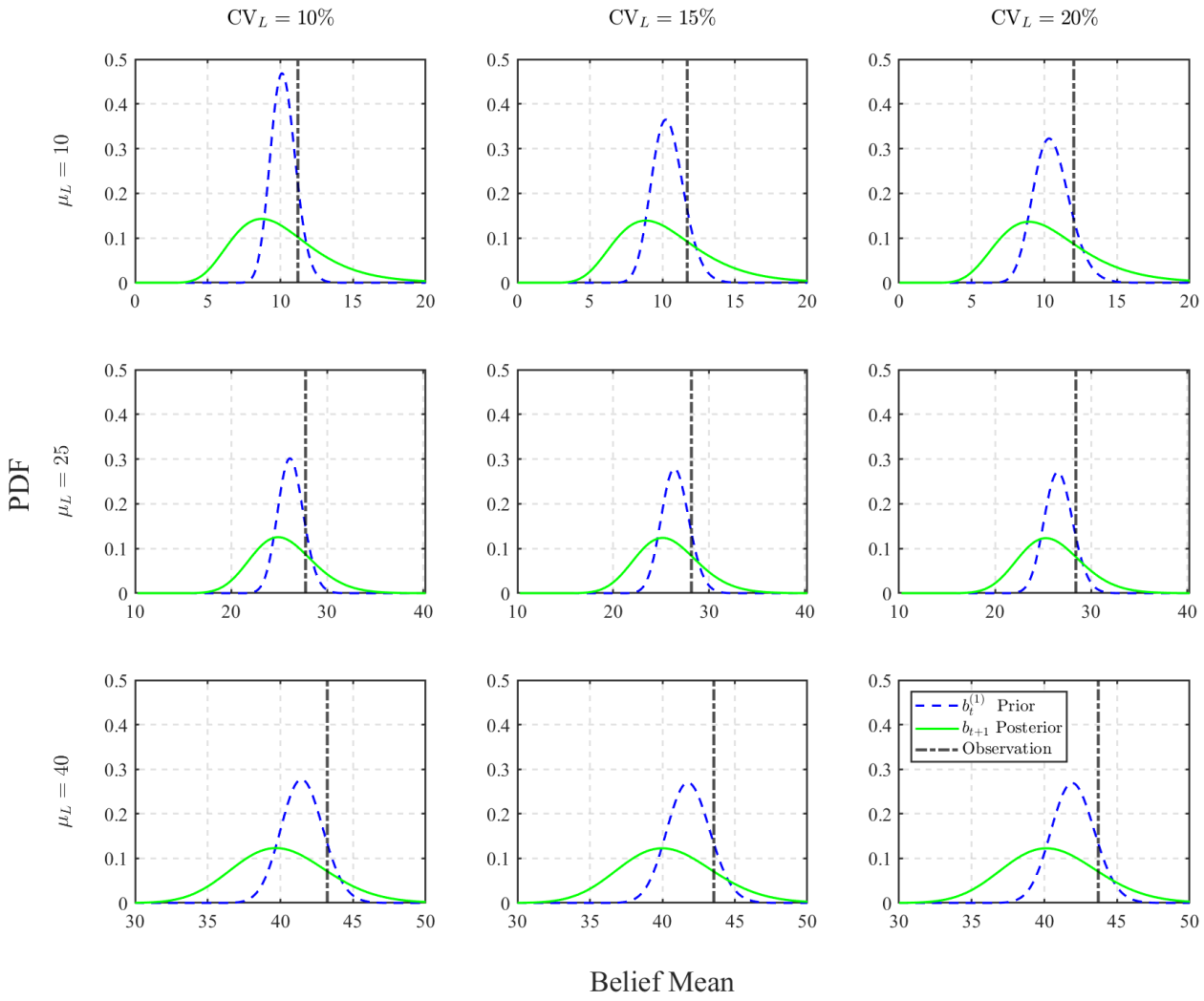

*Notes:* Figure B2 demonstrates that the distribution of corn yields closely follows a log-normal distribution across different mean legacy P levels ($\mu_L$) and coefficients of variation ($\nu_L$). The histograms of the simulated yield samples, along with the overlaid fitted log-normal curves, show 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 $\mu_L$ and uncertainty $\nu_L$. 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 $\mu_L$ and $\nu_L$, demonstrating how these parameters influence the updating process and the resulting belief distributions. As $\nu_L$ 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 $\mu_L$ and uncertainty $\nu_L$ 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 Model

In this section, We estimate a log-linear Markow-switching Dynamic Regression (MSDR) model to construct the basis parameters of prior distributions ($\mu_\mu, \sigma^2_\mu$ and $\mu_\Phi, \sigma^2_\Phi$) in MSVAR model. The specification of the MSDR of the following form:

$$\begin{aligned} \ln(P^Y_{t+1}) &= \alpha_{0,r_{t+1}} + \alpha_{1,r_{t+1}} \ln(P^Y_t) + \alpha_{2,r_{t+1}} \ln(P^F_t) + \epsilon_{t+1} \\ \ln(P^F_{t+1}) &= \beta_{0,r_{t+1}} + \beta_{1,r_{t+1}} \ln(P^F_t) + \beta_{2,r_{t+1}} \ln(P^Y_t) + \upsilon_{t+1} \end{aligned} \tag{C1}$$

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

MDSR results are presented in tables C1. To generate the basis parameters of prior distribution of intercept $(\mu_\mu, \sigma^2_\mu)$ and coefficients $(\mu_\Phi, \sigma^2_\Phi)$, we use the averaged value of estimated values $(\alpha_0, \beta_0)$ and $(\alpha_i, \beta_i)$ for each prior mean and used the averaged values of standard error to calculated the prior variances $(\sigma^2_\mu, \sigma^2_\Phi)$

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

| | Corn ($\ln(P^Y_{t+1})$) | | Phosphorus fertilizer ($\ln(P^F_{t+1})$) | |
|---|---|---|---|---|
| | Moderate | High | Moderate | High |
| $\ln(P^F_t)$ | 0.091 | 0.763*** | 0.947*** | –1.347*** |
| | (0.186) | (0.284) | (0.151) | (0.251) |
| $\ln(P^Y_t)$ | 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 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:

$$\ln(Y_{i,t}) = \beta_0 + \beta_1 \ln(F_{i,t}) + \beta_2 \ln(L_{i,t}) + \beta_3 \ln(F_{i,t})^2 + \beta_4 ln(F_{i,t}) \ln(L_{i,t}) + \omega_i + \epsilon_{i,t}, \qquad \text{(D1)}$$

where $i$ denotes experiment plot, $\omega_i$ is the experimental replication (i.e. plot) fixed effect, and $\epsilon_{i,t}$ is a timevarying error component. The estimation results are in Table D1.

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. Figure D1 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 (Figure D2), the optimal results remain similar to those presented in the original paper. 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 D1: 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 D1: Additional Analysis: Crop Yield Estimation

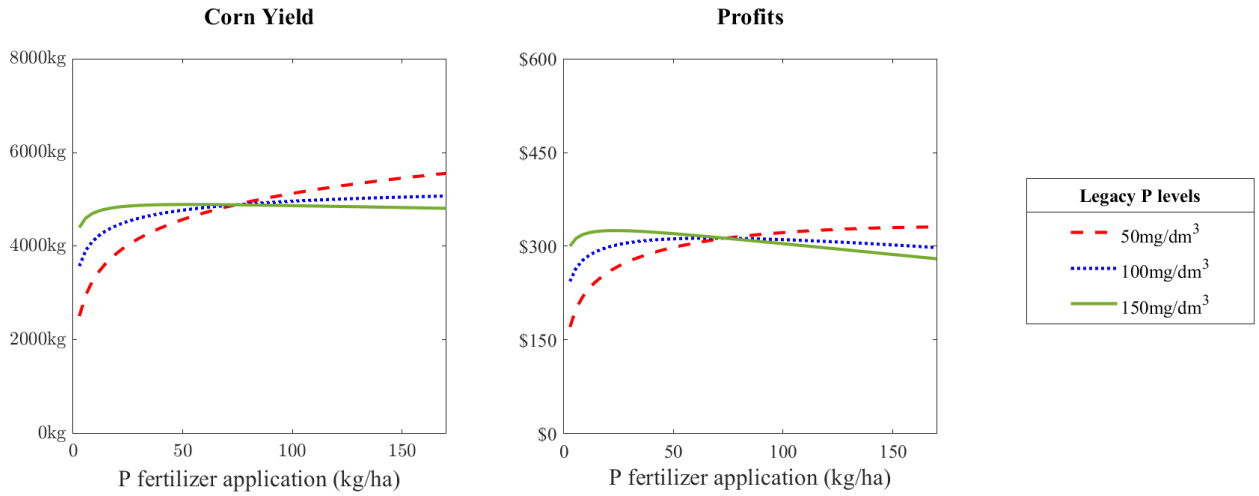
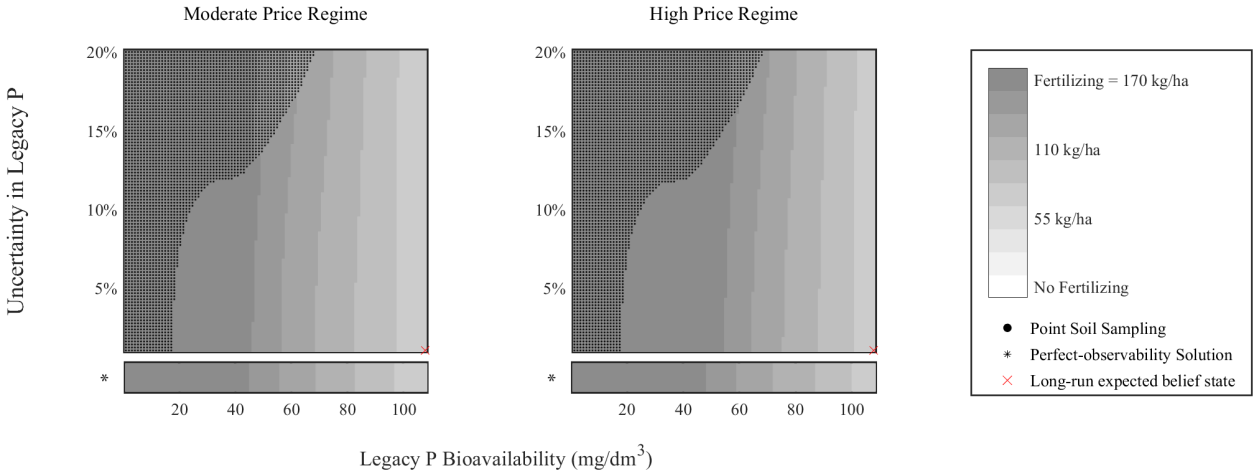
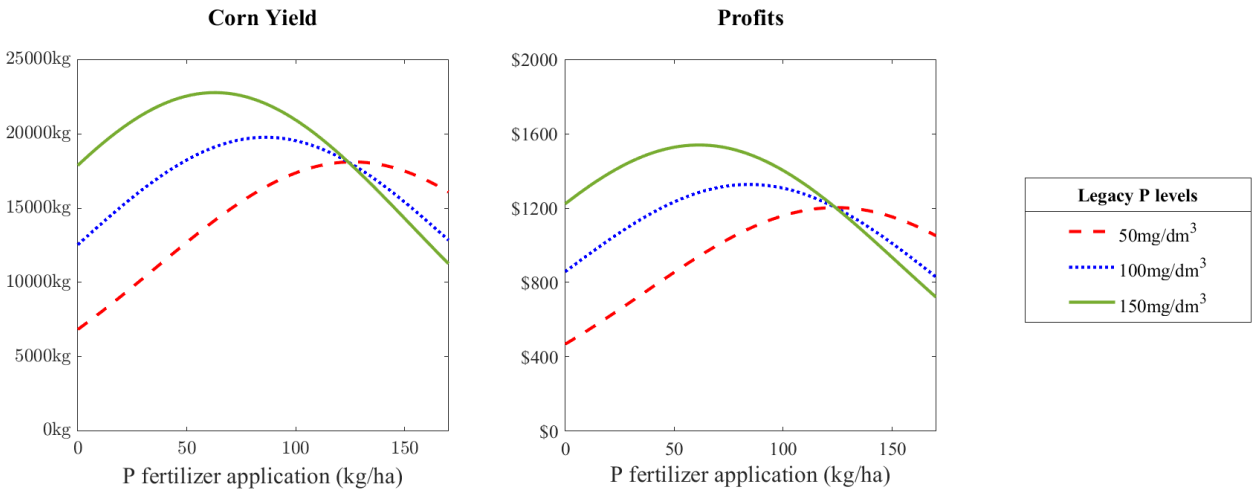


Figure D2: Additional Analysis: Crop Yield Estimation



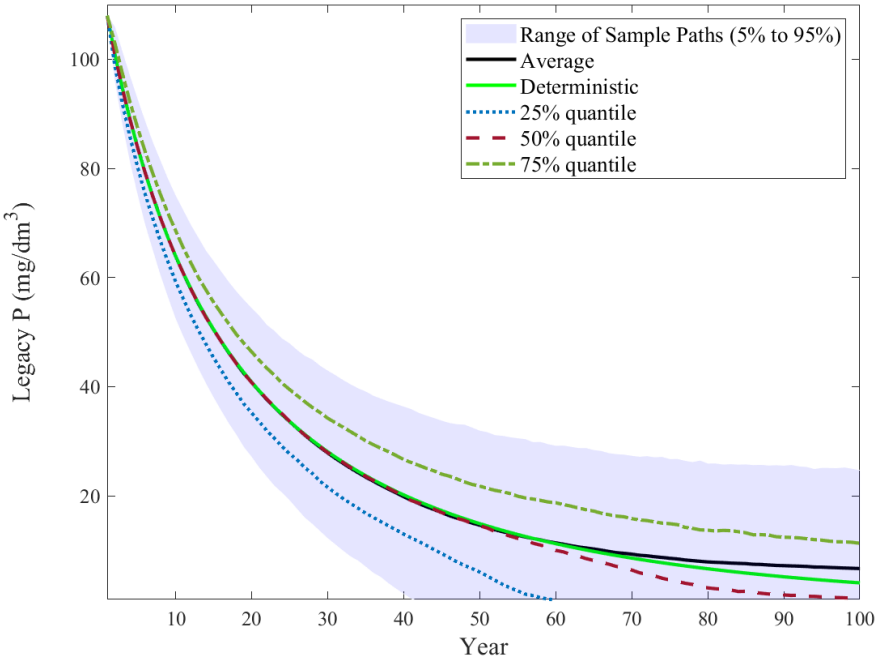
# E Supplementary Figures

Figure E1: 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 E2: Legacy phosphorus accumulation without phosphorus fertilizer application



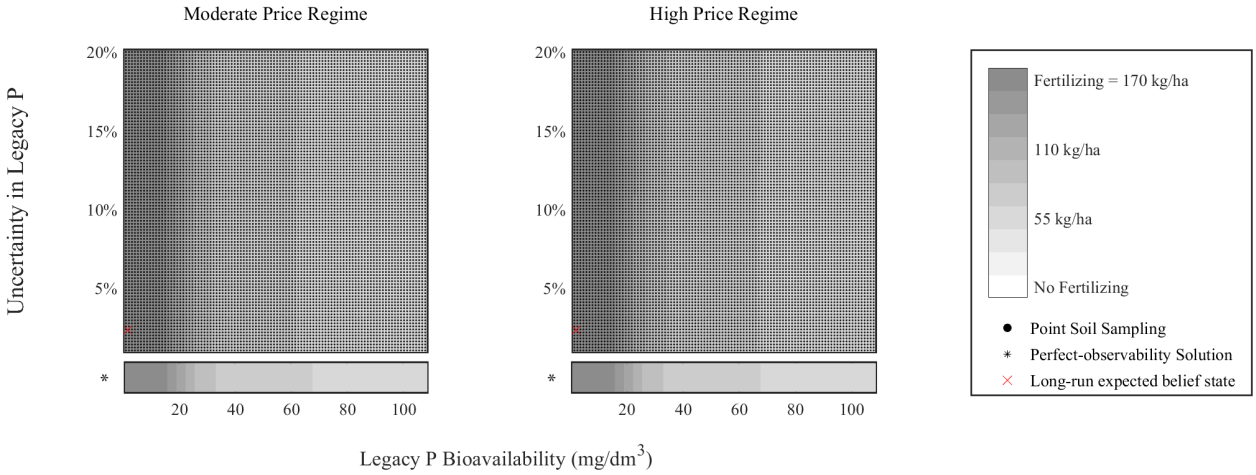

*Notes:* For the deterministic legacy P accumulation (green solid line), we employ a constant carry-over parameter $\rho_t = \rho = 0.98$ (2% decay rate) as adopted by Myyrä et al. (2007). The initial value is the 90th percentile (108 mg/dm$^3$) of legacy P in the North Carolina Tidewater data.

Table E1: Estimated value of risk aversion and elasticity of intertemporal substitution in literature

| Literature | | RA ($\eta$) | EIS ($\psi$) |
|---|---|---|---|
| Howitt et al. (2005) | California (US) | 1.4 | 0.1 |
| Lybbert and McPeak (2012) | Chalbi (Keyna) | 0.5 (OLS)<br>0.8 (IV) | 0.7(OLS)<br>0.9(IV) |
| | Dukana (Keyna) | 13.5 (OLS)<br>12.5 (IV) | 2.8(OLS)<br>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 |

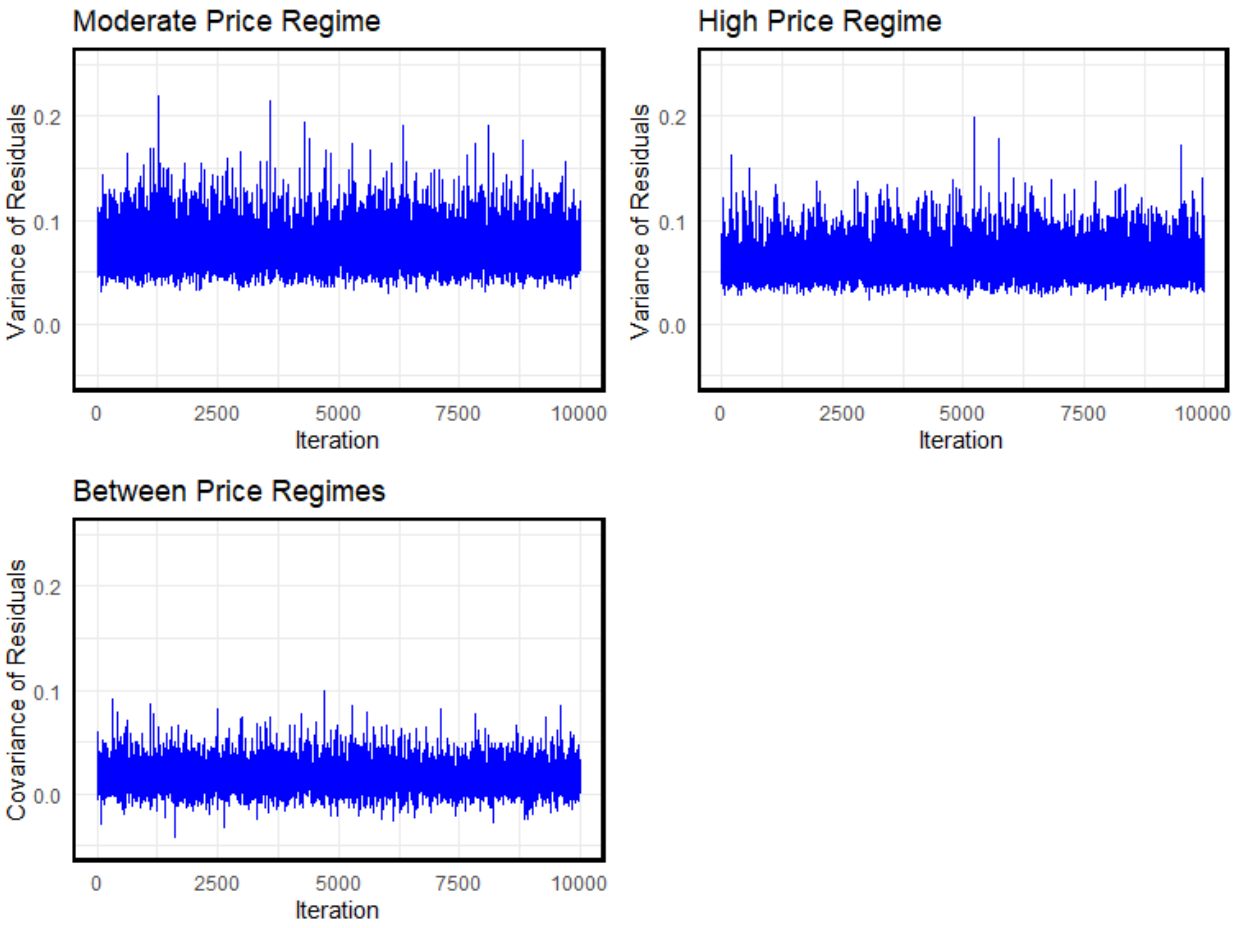
*Notes:* OLS and IV indicate Ordinary Least Squares regression and Instrumental variables estimation, respectively.

Figure E3: 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 E4: 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 E5: Risk analysis: Epstein-Zin preference (moderate price regime)

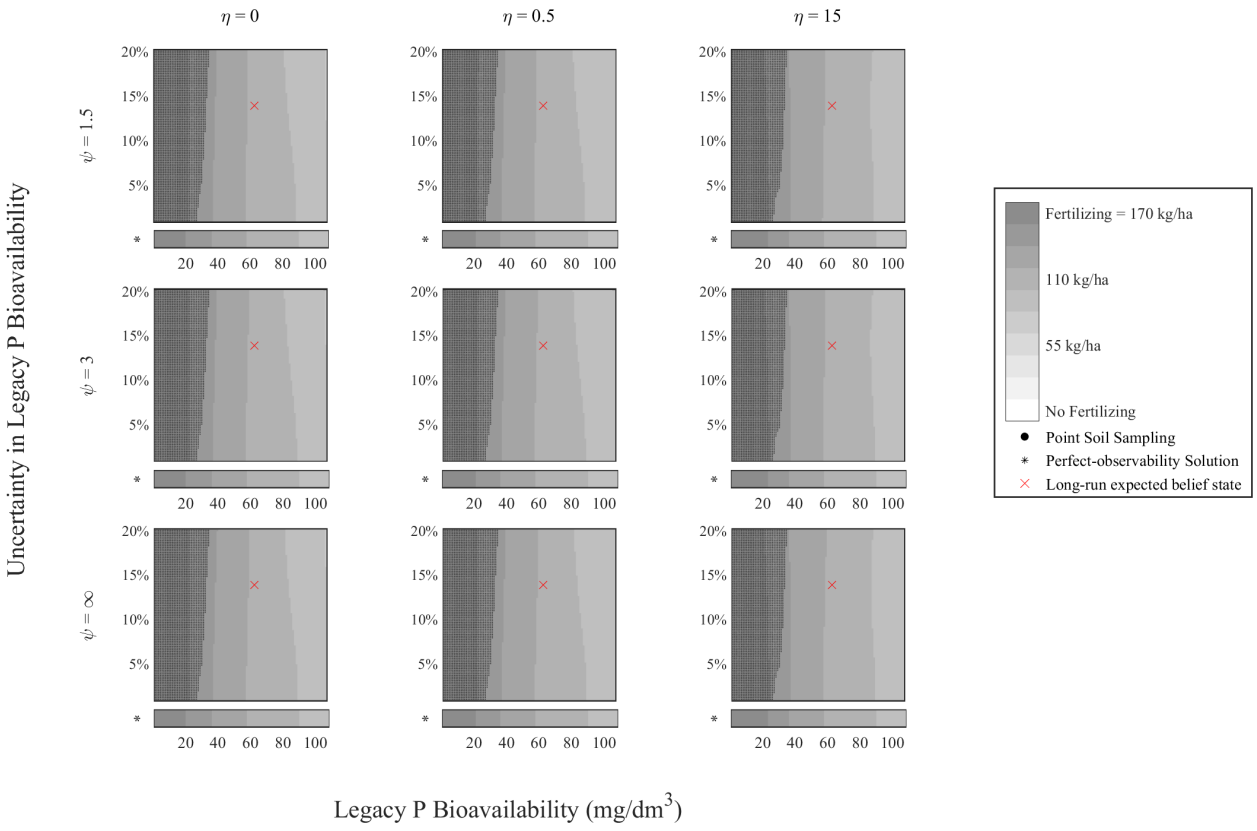


Figure E6: Risk-averse farmer responses to soil sampling subsidy (moderate price regime)

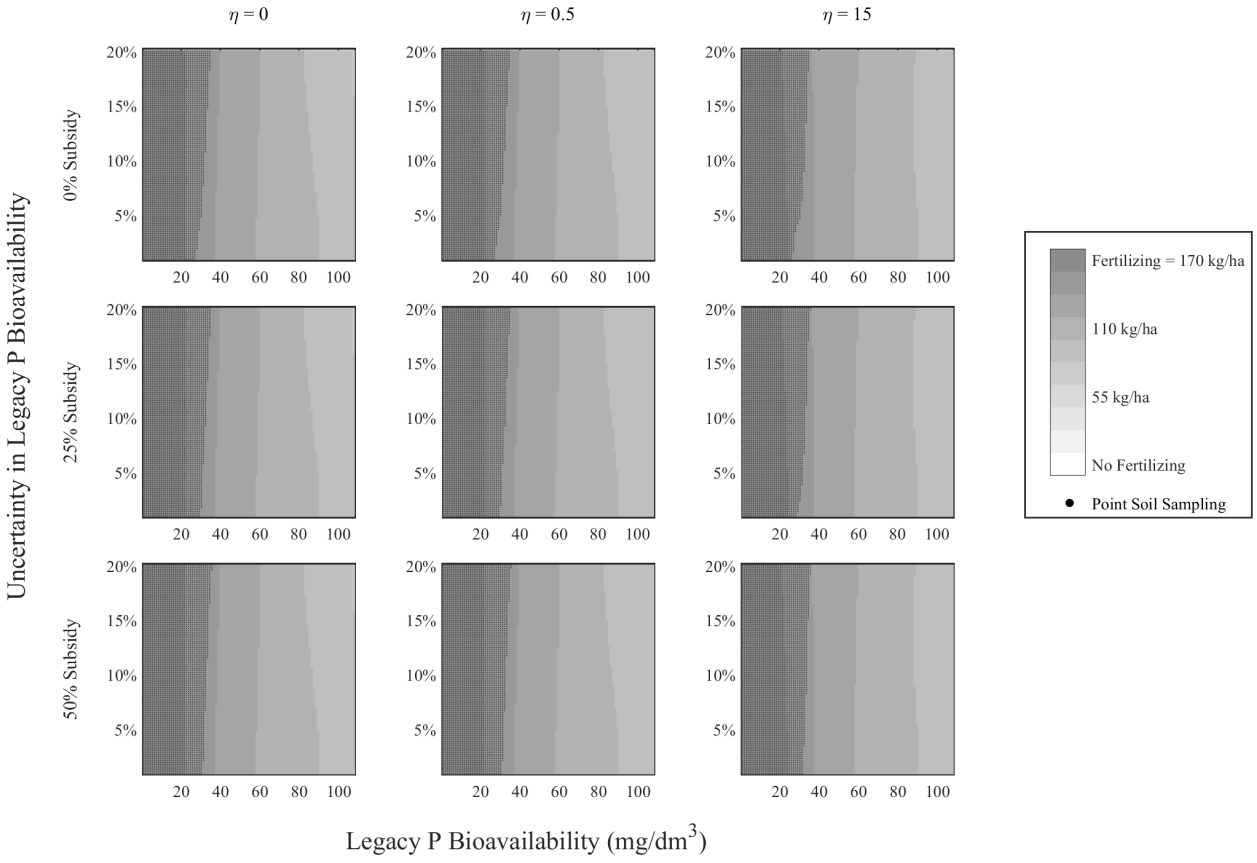
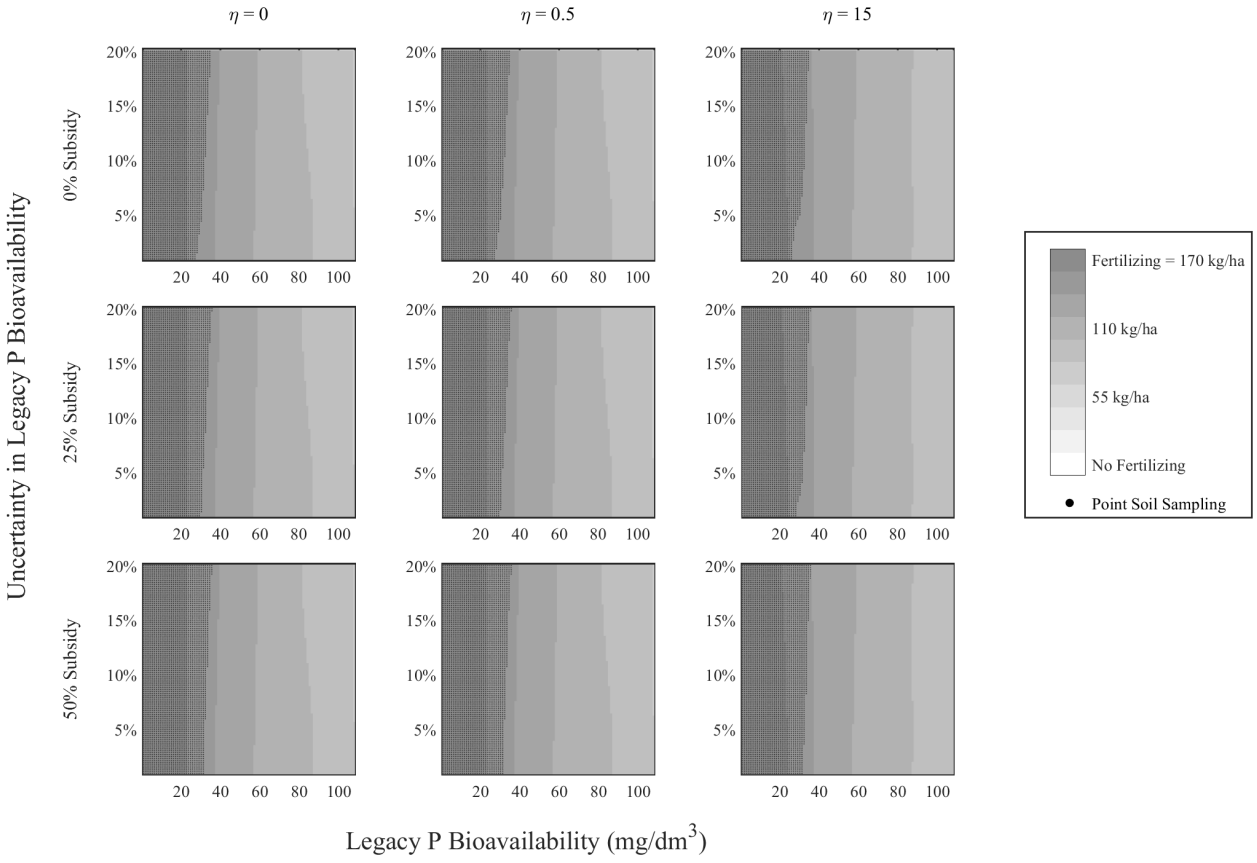


Figure E7: Risk-averse farmer responses to soil sampling subsidy (high price regime)



## 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.

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.