Estimating Crop Rotations as Dynamic Cycles using Field Data

Duncan MacEwan (macewan@primal.ucdavis.edu)
Richard Howitt (howitt@primal.ucdavis.edu)

University of California, Davis


Copyright 2011 by Richard Howitt and Duncan MacEwan. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided this copyright notice appears on all such copies.
Abstract

Crop rotation systems are an important part of agricultural production for managing pests, diseases, and soil fertility. Recent interest in sustainable agriculture focuses on low input-use practices which require knowledge of the underlying dynamics of production and rotation systems. Policies to limit chemical application depending on proximity to waterways and flood management require field-level data and analysis. Additionally, supply elasticity estimates based on crop production as independent activities omit the dynamic effects of a cyclical rotation. We estimate a dynamic programming model of crop rotation which incorporates yield and cost inter-temporal effects in addition to field-specific factors including salinity and soil quality. Using an Optimal Matching algorithm from the Bioinformatics literature we determine empirically observed rotations using a geo-referenced panel dataset of 14,000 fields over 13 years. We estimate the production parameters which satisfy the Euler Equations of the field-level rotation problem and solve an empirically observed four-crop rotation.

Introduction

Crop rotations are an integral part of agricultural production for environmental, agronomic, and financial objectives. Rotations reduce production costs and increase crop yields by managing pests and soil organic matter. For example, alfalfa may be rotated with one year cotton to control root-knot nematodes and a minimum two year rotation of cotton or grains to control dodder weeds. Similarly, tomatoes are rotated annually to increase soil productivity and are rotated up to three years with various small grains to control diseases (UCD IPM). Rotation systems are based on a tradeoff between the immediate profits from a crop planted this season with the future costs (or benefits) realized through changes in yield and production costs for future crops. Field-specific characteristics including soil classification, micro-climate, farmer management skills, and water quality also factor into the rotation decision. In other words, agricultural production decisions are very often a dynamic process but rarely modeled as the solution to a dynamic economic problem.

The field is the minimum decision unit of the farmer and, in many cases, the relevant level of disaggregation for agricultural policy analysis. Farm-wide considerations such as risk and input
smoothing are certainly important for aggregate planting decisions but, fundamentally, we hypothesize farmers understand variation in land characteristics and manage rotations on individual fields. Within a region, we observe that agricultural production exhibits significant spatial specialization, and it follows that where may be as important as what for many agricultural-environmental policies. For example, environmental effects of nitrogen runoff depend on the spatial location (e.g. proximity to waterways) of the field(s) producing a specific set of crops. As another example, regulations in California are being considered to limit chemical application to agriculture within a specific distance of surface water and irrigation.\footnote{More information at: http://www.cdpr.ca.gov/docs/legbills/regsdeve.htm}

Additionally, supply response to exogenous shocks, or policies, may depend on dynamic rotation adjustment costs. The effect of breaking a rotation cycle is not captured in acreage response elasticity estimates based on aggregate data. Treating agricultural production as part of a dynamic cycle at the field level offers valuable insights into these topics.

Agricultural production is increasingly constrained by environmental concerns. Agricultural-environmental policies need to take into account spatial specialization in production. Examples include nitrogen runoff from agriculture into groundwater or local streams, groundwater salinity, and a tradeoff between water for environmental, urban, and agricultural use. Sustainable agriculture suggests a shift to lower input (chemical and pesticide) use and is likely to require farmers to employ tighter rotation management or shift rotation systems. For example, Southern root knot nematode, in cotton, can be controlled through rotation with nematode resistant alfalfa varieties instead of application of chemicals like Temik or Vapam. As a contemporary policy example, the California Department of Pesticide Regulation has proposed banning ground application of pesticides within 25 feet of waterways and within 100 feet for aerial applications. The proposed regulation would affect specific fields within farms and, like the other examples cited, is best modeled at field-level disaggregation.

Farmers manage field rotation systems to capture inter-temporal spill-over effects between crops. Spill-over effects potentially include cost savings and yield increases. As such, when faced with changing prices, or other exogenous shocks, farmers respond to both relative profitability and the long-term cost of breaking a rotation system. Researchers have established, in various contexts, that there is a difference between dynamic and static supply elasticities (Nerlove 1979, Orazem...
and Miranowski 1994, Tegene, Huffman and Miranowski 1988). An interesting extension, which our framework encompasses, is how production cycles, or rotation systems, respond to price shocks. Two different sets of fields may be planted to cotton in a given season but one may be part of a one year cotton-vegetable rotation, say to control seedling diseases, whereas the other may be in a multi-year cotton-alfalfa-vegetable rotation, perhaps to control nematodes. Both changes in relative prices and the dynamic cost of shifting out of breaking the rotation are important. In other words, the fields (corresponding sets of fields) are in two different dynamic cycles and are likely to respond differently to exogenous shocks.

Historically, decisions at the field-level of detail have been difficult to observe consistently across time in anything other than experimental plots. We employ a unique geo-referenced panel dataset of field-level production covering over 14,000 fields (over 1 million acres) and 13 years. Using these data we estimate empirically observed rotations using an Optimal Matching algorithm originally developed for determining common genetic sequences (DNA base pairs). We set up the farmer field-level rotation problem and solve the model using dynamic programming. We estimate the parameters of rotation problem, including yield and cost carry-over effects as well as soil and salinity effects, for a four-crop, seven year alfalfa-cotton-grain-fallow rotation. We estimate parameters using Generalized Maximum Entropy based on observed farmer decisions and show that the model results in a dynamic cycle which responds as expected to price shocks and other perturbations. Finally, we conclude by simulating a grain price shock and show how supply response may differ from aggregate models depending on where fields are in the rotational cycle.

Existing Literature on Crop Rotations

Heady (1948) first formalized the crop rotation problem with a static analysis of the hay-grain rotations observed in the U.S. Corn Belt. Heady (1948) followed work by Johnson (1933), who should be credited as the first to consider the rotation problem. Burt and Allison (1963) formalized the dynamics of rotations in the context of a wheat-fallow rotation. They considered a dynamic programming approach to crop planting decisions, wheat or fallow, in every year dependent on the underlying state of the field (soil moisture). Contemporary research on the economics of crop rotations stems from these seminal works and falls into four main areas: (i) linear programming models of production with fixed proportion rotation constraints, (ii) models
that lend themselves to econometric analysis and control for lagged crop choice, (iii) dynamic analysis which models crop rotation as a control variable consisting of the proportion of total land use, and (iv) multiple-phase optimal control dynamic models which estimate the switch point between two successive crops.

Linear programming models of production with fixed rotation constraints commonly appear in the literature. Hildreth and Reiter (1951) analyzed a corn-oats-hay rotation in the Corn Belt of the United States and treated specific rotations as individual production processes. Initial linear programming models imposed rotation constraints, in essence fixed proportions, on production activities (Swanson 1956, Peterson 1955, Beneke and Winterboer 1973). El-Nazer and McCarl (1986) built on the previous methodology and specified a set of rotational constraints that allowed the program to choose the optimal rotation. They specified a static model that they hypothesized would satisfy the steady-state conditions and, consequently, represent a dynamic solution. The tendency for overspecialization limits linear programming methods, making it necessary to impose significant constraints on the model in order to reproduce a rotation, limiting the effectiveness of the model for policy simulations.

Hennessey (2006) formalized the theory behind models that lend themselves to econometric specification. He considered rotation effects through changes in yield or changes in input use in a framework that allows for positive effects without excluding the possibility of negative effects. Current plantings can affect future plantings through increased yields (soil quality effects) or reduced inputs (capital savings), depending on the length of memory (how many previous years matter) of the rotation. Recent studies that employed a reduced form econometric specification include Wu et. al. (2004) and Wu and Babcock (1998). Both authors specified a reduced form multinomial logit model to analyze land use decisions. The multinomial logit approach estimates the share of land producing each crop and a lagged crop choice term captures crop rotation. Tanaka and Wu (2004) evaluated the Conservation Reserve Program in the United States in a similar framework to investigate the effect of taxes on fertilizer, payments for land retirement, and payments on rotations. They found that taxes on fertilizer have the largest direct impact on environmental effects, whereas a general increase in payments for land retirement leads to retirement of less fertilizer intensive lands. Other research that analyzes the effects of rotations and land use decisions in the context of environmental concerns include studies on wildlife.
abundance (Langpap and Wu 2008), watershed ecosystem protection (Langpap, Hascic and Wu 2008), and ecosystem services (Antle and Stoorvogel 2006, Antle and Valdivia 2006).

Economists often explicitly model the dynamics of crop rotations in a framework where aggregate land use proportions represent crop rotations. Modeling land in proportions makes the control variable continuous, as well as the underlying state equations, so traditional control theory applies. Jaenicke (2000) analyzed the importance of soil quality in corn-soybean rotations within this framework. Orazem and Miranowski (1994) applied a similar framework to a four crop rotation model. The majority of this vein of literature focuses on the optimal fertilizer application rate and dynamic carryover effect (Kennedy et al. 1973, Kennedy 1981, Kennedy 1986, Taylor 1983). In these models, farmers choose the dynamically optimal fertilizer application based on expectations about future returns, current returns, and the different fertilizer carryover rates between sequential crops. In the most compelling paper in this field, Thomas (2002) developed a dynamic model of crop rotation for optimal nitrogen management. Crops utilize the nitrogen stock available in the soil over time, with different crops using or replacing nitrogen at different rates. As such, the farmer faces a dynamic problem of optimal sequential planting decisions that aim to manage nitrogen in the ground and the corresponding rate of fertilizer application. Thomas (2002) developed a structural dynamic model to account for this trade off and estimated a restricted version of the model using generalized method of moments (GMM).

Another area of research includes dynamic multi-phase optimal control models. These models represent a relatively new approach to modeling crop rotations and focus on the field-level decisions and switch points between crops. The method specifies a set of controls that choose a set of optimal switching times between regimes. Doole (2009) provided an algorithm for solving these problems with transition costs and Doole (2008) provided an application of the algorithm. He modeled the optimal rotation in a lucerne (alfalfa)-wheat phase rotation for managing ryegrass weeds. In the multi-phase problem, he exogenously specifies the set of regimes and the optimal controls in each stage give the solution, as well as the switch point. The number of stages must be exogenously specified in this approach since large state spaces render these models intractable. This seems to limit the usefulness of the model for more complicated rotations involving several crops (stages).
In contrast to previous approaches which consider the dynamics of rotations in terms of aggregate land use proportions we explicitly model the discrete switching of field level decisions, subject to a continuous underlying state. We model the state of the field as “fertility” which represents both rotational and field-specific effects. Rotation effects include pest and disease management and soil fertility. Field-specific effects include soil quality and shallow groundwater salinity. In our dynamic programming model, we are able to represent rotations as dynamic cycles which are determined using empirical data. We allow for both yield and cost rotational carry-over effects, consistent with the pests and soil organic content management benefits of rotations. Cost and yield carry-over effects of rotation are modeled as deviations from the mean and are estimated using our geo-referenced panel data. We show how this model responds to changes in prices and other shocks and highlight the relevance for important questions including dynamic supply elasticities, agricultural-environmental policies, and spatial variation in production.

**Motivating A Model of Field Rotations**

Consider a farmer managing a specific field within the farm which can be planted to annual or multi-year crops on a seasonal (yearly) basis. The field has a fixed size and is not subdivided in any given year. For simplicity, and without changing the main conclusions, we normalize to a unit field size. The farmer seeks to maximize the present discounted value of a future stream of profits by choosing the sequence of crops planted every season, i.e., the crop rotation. Initially consider only the rotation effects. The crop planted in the current season will have different effects on the pests, disease, fertility, and other field characteristics in subsequent years. We represent the fertility of each field with an aggregate measure of pests, disease, weeds, and other field specific characteristics. Fertility changes due to rotation, depending on the current crop and previous crop history. Changes in fertility affect both costs and yields and will differ by crop.

As an example, consider a two-crop rotation of alfalfa (fertility increasing) and grain (fertility depleting). Soil organic content increases when growing alfalfa, but pests and diseases also increase and will eventually outweigh the improvements in soil quality and require rotation to a non-host crop. A rotation with grain breaks the pest and disease cycle, but also depletes soil organic content. Specifically, the crop choice, $c_t$, as measured by the proportion of the field
planted to the respective crops at time $t$, where $c_i \in (0,1)$, represents the control variable. The case where $c_i = 0$ corresponds to the entire field in grain, specifically, $i \in (\text{grain, alfalfa})$ as the crops corresponding to $c_i \in (0,1)$. Where $y_i$ and $p_i$ denote the average crop yield and price, respectively. The fertility of the field, $f_i$, at time $t$ given $f(0) = f_0$, characterizes the state of the system. Fertility changes over time depending on the current crop planted according to (E1).

$$\dot{f} = \alpha_a c_i - (1-c_i)\alpha_g f_i.$$ 

Additionally, $\alpha_a > 0$ and $\alpha_g > 0$ represent the marginal effects of planting a crop (alfalfa or grain) in the current period on the rate of change in fertility of the field. Fertility affects current and future profitability through marginal changes to average crop yields, as captured by the parameters $\beta > 0$ and $\delta > 0$, which represent the marginal effect of current fertility on the yield of grain and alfalfa, respectively. The key difference with this model is that we rarely observe fields being sub-divided and proportionally allocated to two different crops in any given growing season. The size of a field is determined based on the need to efficiently use machinery and economies of scale. The entire field is planted to one crop or the other in any given period (year). The dynamically optimizing farmer rotation profit maximization problem with no discounting is defined by (P1).

$$\max \int_0^T c_i(y_g + \delta f_i) p_a + (1-c_i)(y_g + \beta f_i) p_g \, dt$$ 

st.

$$\dot{f} = \alpha_a c_i - (1-c_i)\alpha_g f_i + \gamma \quad f(0) = f_0$$

In addition to rotation effects on fertility, field-level decisions also depend on physical capital. Some of the more important features include soil quality, salinity, slope, and micro-climate. Since we are considering a small agricultural region (Kern County in our data) we can ignore the effects of micro-climate. Additionally, slope is essentially zero across the entire region. However, we do see significant heterogeneity in soil quality and salinity levels across the region. As soil quality decreases, or salinity increases, crop yields decrease and the rate of yield decline
varies by crop (VanGenuchten and Gupta 1993). We let $s\ell$ and $e\ell$ denote soil quality and shallow groundwater salinity with marginal yield effects of $\eta_s$ and $\eta_e$, respectively. This will also affect the rate of change of fertility, uniformly across crops, by $\rho_s$ and $\rho_e$. The modified problem with no discounting is defined by (P2).

\[
\max \int_0^T c_t(y_a + \delta f_t - \eta_{s,a} s\ell - \eta_{e,a} e\ell) p_a + (1 - c_t)(y_g + \beta f_t - \eta_{s,g} s\ell - \eta_{e,g} e\ell) p_g dt
\]

\[
\text{s.t.}
\]

\[
\dot{f} = \alpha_s c_t (1-c_t) f_t + \gamma + \rho_s s\ell + \rho_e e\ell
\]

\[
f(0) = f_0
\]

\[
c_t (1-c_t) = 0
\]

The farmer chooses the optimal crop, at any point in time, by balancing the returns from current period profits with the anticipated future flow of profits as reflected in the stock value of fertility in the ground. Intuitively, when fertility is low the co-state is high, representing a high value on the stock of fertility. If fertility is initially low, alfalfa will be initially be optimal and fertility will increase (co-state decreases) up to the point that it is more profitable to plant grain when considering current returns and future fertility depletions. After switching to grain, fertility will be depleted (co-state increases) up to a point that it becomes optimal to switch back to alfalfa. The process repeats infinitely in the absence of external shocks. The piece-wise continuous control and, correspondingly, piece-wise continuous Hamiltonian make the solution to the optimal control problem non-trivial. We estimate the dynamic programming equivalent formulation of (P2), in the following section, and corresponding solution paths. We solve for the conditions for a repeating series of switches in the theoretical optimal control model in a subsequent paper.

**A Dynamic Programming Model of Field Rotations**

We observe average data on agricultural production in terms of prices, yields, and costs. We also observe marginal rotation decisions in terms of switching between crops on a field. Specifically, we observe the field-level rotation, as a sequence of discrete switches, from which we infer the parameter values that farmers are responding to. We use the data to model farmer rotation decisions in terms of deviations from the mean yield and costs, depending on the
dynamic sequence of crops planted. First, we identify the most common rotations using an Optimal Matching algorithm from the Bioinformatics literature. Next, we describe a field-level model that captures rotation decisions as deviations from the mean yield and cost. Finally, we estimate the parameters for the model based on observed farmer behavior using Generalized Maximum Entropy and solve an empirical dynamic model that yields observed rotations.

Data

Kern County California is located at the southern end of the San Joaquin Valley. Agriculture in the region is irrigated with water coming from State and Federal surface water projects and groundwater in addition to local surface supplies. Our data include all irrigated agricultural land in Kern County between 1997 and 2009. On each field and year we observe the crop grown, field size in acres, farm owner, and farm manager of the field. We are able to uniquely identify and track fields across time using a geo-referenced dataset provided by the Kern County Agricultural Commissioner's Office.

We observe physical characteristics of each field from data that we aggregate up to the field level. Soil data is obtained from the United States Department of Agriculture (USDA) Soil Survey Geographic Database (SSURGO).\(^2\) Since soil type is unchanged from year to year over the time horizon of the data we take a cross section from 2002. The data are geo-referenced and include 7 classifications for agricultural uses, developed by USDA. Shallow groundwater salinity data are obtained from a 2002 analysis completed by the California Department of Water Resources. Salinity is measured in electrical conductivity (EC) in mS/cm which is aggregated up to the field level. Finally, we observe actual evapotranspiration (ET) and dry biomass production on a 30 by 30 meter scale from satellite data provided by SEBAL North America for 2002, 2005, 2008, and 2009.\(^3\) These data are collected using remote sensing techniques and a proprietary software/algorithm for determining land cover and water use which is accurate at over 90 percent (Thoreson et al. 2009). We couple the spatial data with yield, price, and cost data from the Kern County Agricultural Commissioner’s Office and compiled by USDA’s National Agricultural Statistics Service (NASS).\(^4\) For all of these data, we aggregate up to the field level such that

\(^2\) Available at: http://soils.usda.gov/survey/geography/ssurgo/
\(^3\) From: http://www.de-water.com/
\(^4\) Available at: http://www.nass.usda.gov/Statistics_by_State/California/Publications/AgComm/Detail/index.asp
each individual field in the Kern County land use dataset has a unique soil type, salinity, biomass, and ET. Figure 1 shows land use in 2005 by aggregate crop group.

Figure 1. Agricultural Land Use in Kern County 2005

Rotation Identification – Optimal Matching

Sequence alignment is a branch of research within bio-informatics which deals with identifying and comparing similar sequences of amino acids and DNA base pairs. When scientists discovered that common gene sequences were a proxy for common ancestry and evolution there was a need to identify common parts of sequences across many individuals. The goal of sequence alignment is to identify commonalities between individuals which can be used to create a “distance” weighting between other individuals in order to identify common genes. The full problem is computationally intensive and doesn’t guarantee a unique solution. With the development of better computers, the initial algorithm for analyzing multiple (two initially) sequences was published by Needleman and Wunsch (1970). This has since been expanded significantly to encompass multiple sequence alignment, with a variety of algorithms that can
quickly identify common sequences across many observations. Other fields that use these types of algorithms include finance, string editing, and language processing.

We employ a simple version of a sequence alignment algorithm called Optimal Matching in order to empirically identify crop rotations. Optimal Matching is another method for identifying commonalities across sequences which has a range of applications (Abbott and Tsay 2000). We employ a package, SQ-Ados, developed in Stata by Brzinsky-Fay, Kohler, and Luniak (2006). We use a sequence “suppression” option that condenses multiple sequential crops, of the same type, into a single observation and identifies commonalities across reduced form sequences. For example, AABC is the same as ABBC and ABCC, etc. We justify this by noting that rotations are a dynamic process, subject to external shocks, and we intend to formally model the underlying process. We anticipate that price expectations and heterogeneity in land characteristics will induce farmers to grow crops for sequential years. The goal of the sequence analysis is solely to identify potential, base, rotations.

Table 1 summarizes the aggregate data and Table 2 summarizes the results. We show the 20 most commonly observed sequences. We select alfalfa-cotton-grain-fallow as the rotation that we will reproduce and simulate in what follows in the paper. This is a combination of the top two rotations observed in Kern County, which we observe over 963 fields. This rotation is selected because it offers an interesting rotation of crops with different salt tolerance, profitability, and includes fallow. Fallowing a field is a zero profit (absent of rotational effects) event and is an interesting addition to the dynamic model.

<table>
<thead>
<tr>
<th>Table 1. Summary of Rotations, by field, in Kern County</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fields in Top 20 Rotation</td>
</tr>
<tr>
<td>Total Fields in Annual Crops (plus Alfalfa)</td>
</tr>
<tr>
<td>Total Fields in Perennials</td>
</tr>
<tr>
<td>Total Fields</td>
</tr>
</tbody>
</table>
Since alfalfa is a perennial crop we use satellite data to estimate the mean yield in any given year and treat different years of alfalfa as different crops. Specifically, we allow for four years of alfalfa and estimate the mean yield of a field at any point in the four year sequence. Using SEBAL satellite data we identify the mean alfalfa yield by field for 2002 and using the Kern geo-referenced data we determine the age of the stand. Figure 2 summarizes the distribution of alfalfa yields across fields and Table 3 summarizes the corresponding mean alfalfa yields.
Figure 2. Alfalfa Yield Distribution in Kern County 2002 by Fields

<table>
<thead>
<tr>
<th>Yield in tons/ac</th>
<th>Number of Fields</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;5</td>
<td>20</td>
</tr>
<tr>
<td>5-6</td>
<td>40</td>
</tr>
<tr>
<td>6-7</td>
<td>80</td>
</tr>
<tr>
<td>7-8</td>
<td>60</td>
</tr>
<tr>
<td>8-9</td>
<td>50</td>
</tr>
<tr>
<td>9-10</td>
<td>100</td>
</tr>
<tr>
<td>10-11</td>
<td>20</td>
</tr>
<tr>
<td>&gt;11</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 3. Mean Alfalfa Yield by Year (tons/ac)

<table>
<thead>
<tr>
<th>Year</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>8.1</td>
<td>8.3</td>
<td>7.9</td>
<td>7.4</td>
</tr>
</tbody>
</table>

**Model Formulation and Parameter Estimation**

We define the dynamic programming, four-crop rotation model as follows. We extend the basics of the previous model to allow for cost carry over effects and allow the farmer to have basic expectations over future prices. The farmer has expectations about future prices and has direct knowledge of the soil type and water quality associated with the individual field. The farmers’ problem is to determine an infinitely repeating cycle for optimal management of crop plantings on the field.

Let $c_t$ denote planting crop $c$ in period $t$ for $t = 1, 2, ..., T$ to a unit field size. We write $p_t(c)$ as the price of a unit of output (yield in tons) of crop $c$ in period $t$ which yields $\bar{y}(c)$ average yield in tons per acre. Assume that average yield, $\bar{y}(c)$, is constant over time. The variable costs of production are constant over time and denoted $A(c)$, we do not model fixed costs in this version of the model. Thus, the average profits generated from the field, in the absence of rotational effects, are defined by (D1).
We introduce rotational effects, as represented by the effect of crop rotation on the state of the field, in terms of adjustments to mean costs and yields. Let $s_t$ be the state variable that represents the underlying “fertility” of the field, which depends solely on a function of the crop planted in the previous period, $s_t = g(c_{t-1})$. At any point in time yields, and thus profits, are affected by the fertility of the field. Following the suggestions by Hennessey (2006), crop rotations result in yield increases and input savings. To capture yield effects (output boost) we introduce a function that adjusts average yield by crop (positive or negative) depending on the state of the system (a function of previous crop choice). Let $\Gamma(c_t, s_t)$ denote the yield adjustment function for crop $c$ in period $t$ given state $s_t$. Assume that this functional form is stationary. Input savings are included in a similar fashion, denoted $\Psi(c_t, s_t)$.

In addition to rotational effects, yields vary with soil quality and water quality. Thus, we introduce two coefficients to capture the effect of salt and soil on yields. Let these be denoted by $\beta_1(c)$ and $\beta_2(c)$ representing the marginal effect of salt and soil on crop yields, respectively. We assume that these are stationary and unaffected (directly) by crop rotation.

Farmers form expectations about future prices in order to make current production decisions. To model prices that are uncertain at any point in time we allow prices to follow a first-order Markov process, with five states per crop, to represent farmer price expectations. We estimate the corresponding transition matrix based on County Agricultural Commissioner time series of price data. This simplification allows us to model prices as being time independent. Taking this into account and letting the discount factor be $\delta$, the farmers’ optimal rotation problem is written in (D2).

\[
(D2) \quad \max_{c_t} \sum_{t=0}^{T} \delta^t \left( p_t(c_t) \left( \bar{y}(c_t) - \Gamma(c_t, s_t) - \beta_1(c_t) * ec - \beta_2(c_t) * sl \right) - (A(c_t) - \Psi(c_t, s_t)) \right)
\]

subject to:

- $s_{t+1} = g(c_t)$
- $p_t$ follows a first order Markov Process
Following Bellman (1957) we can express this problem using the Bellman Equation. Let $V_t(s)$ be the value function, which is the maximum attainable sum of current and future profits from cropping activities given that the field is in state $s$ at time $t$.

\[ (D3) \quad V_t(s_{t+1}, p_{t+1}(c_{t+1}), \bar{y}(c_{t+1})) = \]
\[ \max_{c_t} \mathbb{E}_{p_t} \left\{ p_t(c_t)(\bar{y}(c_t) - \Gamma(c_t, s_t) - \beta_1(c_t)*ec - \beta_2(c_t)*sl)\right\} \]
\[ -\left( A(c_t) - \Psi(c_t, s_t) \right) + \beta V_t(s_t, p_t(c_t), \bar{y}(c_t)) \]

This formulation satisfies the contraction mapping theorem and lends itself to the solution methods used in Dynamic Programming. Specifically, we assume an infinite time horizon for the problem and use value function iteration to find a fixed point of the Bellman Equation. From there we determine the optimal policy function, which gives the optimal crop planting decision, at any point in time, given the state of the system. We simulate the alfalfa-cotton-grain-fallow rotation and assume a one year crop lag as detailed in Equation (D2). Instead of specifying functions for the soil, salt, yield carryover, and cost carryover effects we specify these as individual crop specific parameters, constant across time. Given the above model definition, there are 106 parameters. After imposing restrictions on second, third, and fourth year alfalfa, the dynamic model reduces to 61 parameters. These parameters are as follows.

$\beta_1 \quad \text{soil}$
$\beta_2 \quad \text{salinity}$
$\Gamma \quad \text{yield carryover effect}$
$\Psi \quad \text{cost carryover effect}$

Where $\beta_1$ and $\beta_2$ are 4 by 1 vectors of crop specific soil and salinity yield adjustment parameters, respectively. We anticipate that these are positive, reflecting a negative effect of decreasing soil quality and increasing shallow groundwater salinity. The parameters $\psi$ and $\gamma$ represent cost and yield carryover effects due to crop rotation, respectively. The $i,j$ entry of each matrix represents the yield or cost adjustment from planting crop $i$ today given crop $j$ was planted
in the previous year. We anticipate that these parameters can take any sign, representing both the positive and negative agronomic effects from rotating crops.

We observe the base alfalfa-cotton-grain-fallow rotation on 963 fields in the data. We define crop \( i \) and \( j \in \{a1, a2, a3, a4, c, g, f\} \), prices as \( p_i \), and average yields as \( \bar{y}_i \). Additionally, there is variation in yield, unobserved in the panel dataset, but observed in County level data across years. Yield variation is due to non-rotation factors such as weather shocks and water supply.

We estimate the yield variance, \( \sigma_i^2 \), using County Agricultural Commissioner data for Kern County between 2000 and 2009. We define the crop specific profits on a single field from growing crop \( i \) following crop \( j \) in any year \( t \) in (E2)

\[
\pi_{t,ij} = p_i(i)\left([\bar{y}(i) + \varepsilon_i] - \Gamma(i,j) - \beta_i(i)\cdot ec - \beta_j(i)\cdot sl\right) - \left(F(i) - \Psi(i, j)\right)
\]

Where \( \varepsilon_i \sim N(0, \sigma_i^2) \) and \( sl \) and \( ec \) measure shallow groundwater salinity and soil quality, as defined previously.

Assuming farmers are profit maximizing agents, behaving according to the model specified above, there are a set of 42 dynamic first-order conditions that must hold in order for the base alfalfa-cotton-grain-fallow rotation to be observed. For example, if alfalfa is observed on a field in the current year then the field will rotate into cotton in the subsequent period and was fallowed in the previous period (prior to year 1 alfalfa). This implies six conditions that must be satisfied for each year of the alfalfa crop (four years). Similar logic holds for fields that, at a point in time, are observed in cotton, grain, or fallow. These conditions will not hold exactly in all cases, thus we specify an error term and use this system of equations for the estimation procedure. These are defined, without error term, by equations (E3) to (E6)

\[
\pi_{cja} \geq \pi_{cia} \text{ for all } i \neq c
\]

\[
\pi_{gic} \geq \pi_{circ} \text{ for all } i \neq g
\]

\[
\pi_{f|g} \geq \pi_{f|g} \text{ for all } i \neq f
\]
(E6) \[ \pi_{af} \geq \pi_{if} \text{ for all } i \neq a \]

There are 42 first-order conditions and 61 parameters, the problem is underdetermined, which may explain why there is a dearth of empirical estimations in the literature. The problem is ill-posed and one solution is to use Generalized Maximum Entropy (GME) (Jaynes 1963) (Shannon 1948) (Mittelhammer, Judge and Miller 2003). GME is based on the Kullback-Leibler criteria and is based on the idea that, given that you have incomplete observations about a statistical process, the best way to recover parameters is to impose probabilistic structure on the model in such a way that it is consistent with observed data and imposes as little additional information as possible.

The GME problem is specify support values over the unknown parameters, and estimate the probability weights over each parameter support distribution by imposing minimal additional information and satisfying the known data constraints. For each field we observe the average price per ton of yield, soil quality, shallow groundwater salinity levels, and the average yield, where the actual yield is stochastic between fields. We define a truncated uniform support space for all the parameters: yield rotation-carryover, cost rotation-carryover, soil, salinity, and error terms. The GME estimation procedure is to maximize the entropy measure by choosing the probability weights over the support values subject to the data constraints. The data constraints include the 42 first order conditions, and the requirement that each of the probability supports sum to one, as well as the typical non-negativity restrictions.

The support space for the parameters in the GME program is a truncated uniform distribution. The yield rotation-carryover parameters support is +/- 35 percent of average crop yield. The cost rotation-carryover parameters support is +/- 50 percent of average variable production costs. Finally, the salt and soil yield adjustment parameters, corresponding to decreasing soil quality or increasing salinity, support is 0 to -40 percent of average crop yield. In general, we impose a loose probabilistic structure on the model, consistent with agronomic literature on rotation, salinity, and soil effects.
**Parameter Estimation Results**

We solve the GME program in the General Algebraic Modeling Software (GAMS) using the conopt3 non-linear solver. An iteration of the program solves in just over 20 minutes. We bootstrap standard errors for the parameters with 500 boot-strap iterations. Results of the GME estimation are shown in Tables 4, 5, and 6, which we discuss in turn below. Standard errors are reported in parentheses.

Table 4 summarizes the marginal effect of salinity and soil on average yields. Parameters are interpreted as the marginal adjustment in tons per acre to mean yield due to a one unit change in salinity or soil quality. Salinity is measured in dS/m and soil is by SSURGO definitions, as discussed previously. The estimated marginal effect of salinity on crop yield is consistent with the literature. Namely, alfalfa is relatively salt-intolerant and cotton and grain are more salt tolerant. The estimated marginal effects, in percentage terms, reflect this agronomic information with alfalfa realizing the largest yield decrease.

|      | ALF1 | ALF2 | ALF3 | ALF4 | COT  | GRN    | FAL |
|------|------|------|------|------|------|--------|-----|------|
| **Salinity Yield Parameters** |      |      |      |      |      |        |     |      |
| ALF1 | 0.6750 | 0.6750 | 0.6750 | 0.6750 | 0.0576 | 0.2305 | n/a |      |
|      | (0.000) | (0.00) | (0.00) | (0.00) | (0.030) | (0.110) |     |      |

<table>
<thead>
<tr>
<th></th>
<th>ALF1</th>
<th>ALF2</th>
<th>ALF3</th>
<th>ALF4</th>
<th>COT</th>
<th>GRN</th>
<th>FAL</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Soil Yield Parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALF1</td>
<td>0.1197</td>
<td>0.1197</td>
<td>0.1197</td>
<td>0.1197</td>
<td>0.0120</td>
<td>0.0386</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.011)</td>
<td>(0.002)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Parameter estimates for the rotation adjustment effects for costs and yield are reported in Tables 5 and 6. An entry in the matrix is interpreted as given that crop (column) was planted last period the marginal change in costs/yield relative to the average if crop (row) is planted this period. Parameter estimates are based on farmer behavior and, as such, should be interpreted as the implied yield and cost rotational adjustments based on observed farmer behavior. Entries denoted with “n/a” represent imposed restrictions.
Table 5. Estimated Effect of Rotation on Costs

<table>
<thead>
<tr>
<th></th>
<th>ALF1</th>
<th>ALF2</th>
<th>ALF3</th>
<th>ALF4</th>
<th>COT</th>
<th>GRN</th>
<th>FAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE</td>
<td>(0.113)</td>
<td>(0.024)</td>
<td>(0.011)</td>
<td>(0.000)</td>
<td>(0.393)</td>
<td>(0.020)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>ALF2</td>
<td>-17.6089</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>SE</td>
<td>(0.005)</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALF3</td>
<td>n/a</td>
<td>-21.8552</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>SE</td>
<td>.</td>
<td>(0.711)</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALF4</td>
<td>n/a</td>
<td>n/a</td>
<td>-26.8385</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>SE</td>
<td>.</td>
<td>.</td>
<td>(0.006)</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>COT</td>
<td>51.6447</td>
<td>61.0621</td>
<td>77.6275</td>
<td>-58.3097</td>
<td>88.2000</td>
<td>94.8150</td>
<td>64.5438</td>
</tr>
<tr>
<td>SE</td>
<td>(0.000)</td>
<td>(0.130)</td>
<td>(0.045)</td>
<td>(0.056)</td>
<td>(0.000)</td>
<td>(0.002)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>GRN</td>
<td>29.5550</td>
<td>29.5550</td>
<td>29.5550</td>
<td>29.5550</td>
<td>-51.4000</td>
<td>55.2550</td>
<td>30.7217</td>
</tr>
<tr>
<td>SE</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>FAL</td>
<td>23.0000</td>
<td>23.0000</td>
<td>23.0000</td>
<td>23.0000</td>
<td>23.0000</td>
<td>-200.0000</td>
<td>23.0000</td>
</tr>
<tr>
<td>SE</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

Table 6. Estimated Effect of Rotation on Yields

<table>
<thead>
<tr>
<th></th>
<th>ALF1</th>
<th>ALF2</th>
<th>ALF3</th>
<th>ALF4</th>
<th>COT</th>
<th>GRN</th>
<th>FAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALF1</td>
<td>-0.9020</td>
<td>-1.1002</td>
<td>-1.1621</td>
<td>-0.7201</td>
<td>-1.2000</td>
<td>-0.6160</td>
<td>1.5788</td>
</tr>
<tr>
<td>SE</td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.007)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>ALF2</td>
<td>1.0242</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>SE</td>
<td>(0.004)</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALF3</td>
<td>n/a</td>
<td>1.1952</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>SE</td>
<td>.</td>
<td>(0.001)</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALF4</td>
<td>n/a</td>
<td>n/a</td>
<td>1.1996</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>SE</td>
<td>.</td>
<td>.</td>
<td>(0.004)</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>COT</td>
<td>-0.0755</td>
<td>-0.1209</td>
<td>-0.1347</td>
<td>0.0937</td>
<td>-0.1350</td>
<td>-0.3393</td>
<td>-0.1277</td>
</tr>
<tr>
<td>SE</td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>GRN</td>
<td>-0.2233</td>
<td>-0.2233</td>
<td>-0.2233</td>
<td>-0.2233</td>
<td>0.5800</td>
<td>-1.0933</td>
<td>-0.2442</td>
</tr>
<tr>
<td>SE</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>FAL</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>SE</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

The elements below the main diagonal (and in the top right corner) in the matrices of Table 5 and 6 are the key parameters for the rotation problem. The base sequence of crops in the rotation is
alfalfa year 1 through alfalfa year 4, cotton, grain, and then fallow. For example, when grain follows cotton this translates into an average cost savings of $51.40 per acre and an average grain yield increase of 0.58 tons per acre. If, instead, cotton were to follow cotton this would imply an increase in variable costs of $88 per acre and a decrease in average cotton yield of 0.135 tons per acre. In other words, farmers are behaving as if average costs and yields were adjusted according to the matrices in Tables 5 and 6.

**Simulation of the Model**

We substitute the parameter estimates from the GME program back into the dynamic program (D2), and define the Bellman Equation (D3). We solve the program using Value Function Iteration which solves for the fixed point of the Bellman equation. The program converges in slightly over 200 iterations.

First we demonstrate the base results of the model in Figure 3, below. We assume a field starting in fallow, over a 50 year time horizon with average salinity and soil conditions. As shown, we can reproduce the base alfalfa-cotton-grain-fallow rotation. The field is in an infinite cycle of alfalfa-cotton-grain-fallow. We show this cycle over fifty years in Figure 3.

![Figure 3. Average Conditions - Base Results](image)

Parameters of the average conditions, base results, are based on the fundamental alfalfa-cotton-grain-fallow rotation. On any given field this rotation is observed over some sequence of years.
in the data. However, other crops may be grown or the rotation may be of different length. The extent of this difference depends on relative prices, salinity, and soil quality.

**Policy Simulation**

In the data we also observe rotations that deviate from the base (alfalfa-cotton-grain-fallow) as relative prices change. To demonstrate this, we impose price shocks and show how different fields respond. To demonstrate the model for these situations we create a grain price shock from $195/ton to $295/ton in year 15 lasting 10 years. We consider a 30 year horizon for this example. This is a stylized example of the grain price spike in 2007/2008.

Figure 4 shows the field described above, over average salt and soil, except this field is initially fallowed. As shown, in year 15 the farmer has already made the decision to plant first year alfalfa. Thus the optimal decision is to cycle through 4 years of alfalfa, into cotton, and then into a grain monoculture for the duration of the price spike. Note that if the grain price spike is of shorter duration (say only 3 years) then this field never rotates into grain during the price spike. In other words, it is more profitable to grow grain during the price spike, but it does not immediately outweigh the dynamic cost of switching out of the rotation. Alternatively, if the grain price spike is higher, the field rotates directly into grain. When the spike ends, the farmer shifts the field back into first year alfalfa and continues the base observed rotation.

![Figure 4. Average Conditions – Grain Price Spike](image)

Figure 5 shows a field situated over high salinity or poor soil quality, the results are the same for both, subject to the same grain price shock. The dynamically optimizing farmer initially plants
the field to a continuous alfalfa-fallow rotation. This is a result of the effect of poor soil or high salinity on crop yield, it is not profitable at initial prices to plant the field into other crops. Alfalfa increases soil organic content but, at the same time, pests and disease build up. Every four years the field is cleared by rotation into fallow. When the grain price spike hits, in year 15, the farmer has just finished a 4 year alfalfa rotation. It is optimal to shift the field into grain monoculture for the duration of the price spike. After the spike, the field is put back into an alfalfa-fallow rotation. Again, we would see different effects if the field was in first, second, or third year alfalfa at the time of the price spike.

Figure 5. Poor Soil or High Salinity – Grain Price Spike

Figure 6 shows a field over very marginal land, with high salinity and poor soil quality. At average prices, this field is never optimal to plant. With a grain price spike in year 15, grain monoculture becomes profitable for the duration of the price increase, at which point the field is fallowed again. This case highlights the potential for fallow (marginal) land to be brought into production during price spikes. Note that average grain yield on this field is much lower than other fields, given the high salinity and poor soil quality.
We motivated this paper with two examples of why field-level decisions are important: spatially-dependent agricultural-environmental policies and dynamic supply response. We illustrate the latter with an example. Fields in the exact same rotation, but at different points in the rotation cycle, will respond differently to relative price changes. In this case, we highlight the same grain price spike described above. Figure 7 shows two fields that are over the same salinity level and the same soil type. Specifically, soil is average quality and salinity is high. This is comparable to the field in Figure 3 with high salinity. The difference between the two fields shown in Figure 7 is that they are at different points in the rotation when the grain price spike hits. This induces different responses across fields, as shown in Figure 7. The top field, in third year alfalfa in year 15, cuts the alfalfa rotation short at three years and shifts into a cotton-grain rotation for the duration of the price spike. In contrast the bottom field is in fourth year alfalfa, plants cotton in the following year, and then shifts into a cotton-grain rotation for the duration of the price spike. The difference is the rotation adjustment costs, given the current state of the field, are too high to skip cotton in the rotation.

The importance of this result is that supply response to price shocks (or, more generally, policy response to any policy) will be dependent on both the price shock and the dynamic effects of rotation on a given field. Additionally, across sets of fields within a region, different sets of fields may be planted to the same crop but are part of different rotations. For example, as shown in Table 1, alfalfa-grain-fallow and vegetable-grain-fallow are two different rotations although both include grain. Each is governed by a different underlying dynamic rotation process, and
resulting cycle, thus we anticipate different responses to price shocks. In other words, the costs of rotation adjustment vary both within and across rotation systems. We leave this for future work.

Figure 7. Average Soil and Moderate Salinity Grain Price Spike; Demonstrating the Importance of Rotation Cycle

Conclusion

In this paper we determine empirically observed rotations using an Optimal Matching algorithm from the bioinformatics literature. We formulate a dynamic model to estimate observed rotations, estimate the parameters for the model based on observed farmer behavior, and solve the dynamic rotation problem. Finally, we apply the model to an example grain price spike and demonstrate the effects of price spikes on rotational cycles. The solution of the rotation problem, over average conditions, is an infinitely repeating cycle of the observed rotation: alfalfa-cotton-grain-fallow. As relative prices, salinity, and soil quality change (across fields or due to
exogenous shocks) the optimal rotation changes accordingly. In addition to the novelty of the data, model, and estimation, this type of model and analysis is relevant for estimating dynamic supply response elasticities and evaluating agricultural-environmental policies.

The dynamic costs of switching a rotation are important when agriculture is facing exogenous shocks such as price spikes. As the relative profitability of one crop increases the farmer may switch into this crop or simply shift the rotation system. There is a future cost, in terms of profits forgone by breaking the pest and disease management cycle, which we capture in our dynamic model. Treating agricultural production as a static process, or modeling supply in aggregate proportions, may omit important dynamic features of production.

As agriculture is increasingly constrained by environmental concerns, agricultural-environmental policies are more important and need to realize spatial specialization in agricultural production. Policies that limit pesticide application or reduce nitrates are spatially dependent on individual field locations. For example, modeling in terms of aggregate land use proportions offers limited insights into a policy that restricts pesticide application near surface water (including irrigation ditches). If any part of a field is affected, the farmer would likely adjust management practices on the entire field. We expect the rotation system to adjust accordingly and only field-level analysis can capture this type of response. Additionally, aggregate, regional analysis of agricultural production overlooks heterogeneity in human and physical capital within a region. As geo-referenced land use and satellite production data become more widely available field modeling and accounting for dynamic rotation effects become more important. In other words, where is as important as what.
References


