Simulating the Value of Information Generated by On-farm Agronomic Experimentation Using Precision Agriculture Technology

David S. Bullock
University of Illinois
Department of Agricultural and Consumer Economics

Paper to be presented at the
IATRC Symposium, *Productivity and Its Impacts on Global Trade*
June 2-4, 2013
Seville, Spain

Simulating the Value of Information Generated by On-farm Agronomic Experimentation Using Precision Agriculture Technology

Site-specific agricultural technology has been commercially available for over fifteen years. The technology, in particular variable-rate input application technology, was first met with great excitement in the farm and popular presses. The idea of using truly space-age methods to customize input management to small parts of farm fields was exciting. However actual adoption of site-specific agricultural technology remains scant,—far below projections of an enthusiastic farm press of the 1990s (Whipker and Akridge 2006).

Addressing this subject amidst the early excitement, Bullock, et al. (1998), Bullock and Bullock (2000), and Bullock, Lowenberg-DeBoer and Swinton (2002) predicted that variable rate technology would be neither profitable nor widely adopted in the then-foreseeable future. They argued that the complementarity between information about yield response and precision technology made it infeasible to use the new technology without much more information than was available:[1]

> *... while much has been made in the popular press about the future possibilities of VRS* [variable rate seeding]*, given our current very limited knowledge of the effects of characteristics on yield response, this*

[1] The need for variable rate management technology to be accompanied by increased information about how yields respond to input was largely unanticipated by the agricultural community. Especially in terms of fertilizer application, a principle reason that this information and technology complementarity was not well understood was because university- and industry-guidelines for fertilizer application management were largely based on a "yield-goal" method. It was believed that the economically optimal nitrogen fertilizer application rate for corn was a linear function of the site's "yield potential." Therefore, because precision technology provided farmers with yield maps, it was commonly believed that site-specific economically optimal nitrogen fertilizer application rates could be calculated straightforwardly from yield maps. Subsequent research has made clear that there has never been a scientific basis for the "yield goal" approach to fertilizer management (Rodriguez 2013).

> *technology on its own is of no economic benefit to farmers. If VRS is ever to be profitable on a wide scale and adopted, farmers will have to know much more about how the characteristics of their fields' sections vary, and how this variance causes yield response … to vary across sections.*
> (Bullock, et al. 1998, p. 836)

Bullock, et al. (2009) provided empirical support for this hypothesis.

The economic complementarity of precision technology equipment and yield response information has important research implications and ultimately important practical implications, which are both bad news and good news to precision technology enthusiasts. The bad news is that unless more information can be produced about yield response, precision technology may never be profitable on a wide scale. Similarly, because precision technology is available, the value of information produced by yield response research has increased. However the good news, as proposed by Bullock, et al. (2002) and Bullock and Lowenberg-deBoer, (2007) is that precision technology itself might be used to inexpensively supply the information needed to make precision technology profitable:

> *Whether there is much of a future for variable rate technology … depends heavily on how much information can be generated on how crop yields respond to inputs. An encouraging aspect of the problem is that the same technology that has increased the demand for site-specific information can also increase its supply… . That is, GPS systems and variable rate spreaders have greatly cut the cost of agronomic experimentation. Agronomic experimentation, which used to be a very labor-intensive*

endeavor (graduate student labor, that is) can now with precision technology be capital intensive. Small plot experimentation previously required a good deal of labor time in a field, setting out orange flags to demarcate separate plots, fertilizing the small plots at different rates with small equipment, etc. But with VRT and GPS, now these same tasks can be performed "on the fly" on whole fields. It's not hard to imagine a future in which many farmers, consultants, and extension personnel collaborate in running on-farm agronomic experiments. The amount of data produced would dwarf what is now available, and the accuracy of response function estimates would increase many fold. Since this is exactly the type of information needed to make VRT profitable, VRT itself might be the key to supplying the information needed to increase VRT demand. (Bullock and Lowenberg-DeBoer 2007).

The objective of our research is to examine the practicality of this hypothesis that precision technology could be used to sufficiently lower the costs of acquiring the information necessary to make precision agriculture profitable.

A Simple Theory of Precision Technology Adoption (and the Lack Thereof)

In this section we present a theory to explain the potential social value of using precision technology to conduct increased on-farm agronomic experimentation. Figure 1 features a partial equilibrium model of two markets. The right-hand panel shows a market for precision technology equipment. The curve labeled S_E' represents the supply that equipment in some initial period. The demand for precision technology equipment is initially $D_E(P_E, P_I^0)$, where P_I^0 is the initial price of information about how crop yields

respond to input application rates in various parts of the crop producer's fields.[2] We have drawn these initial curves such that the vertical intercept of S_E' exceeds that of $D_E(P_E, P_I^0)$, implying that in the initial equilibrium precision technology equipment is neither bought nor sold. This is meant to represent a state in which precision technology has not yet been invented, or else would be so expensive to produce that no buyers are willing to pay enough for any to be produced in equilibrium. In the initial state, even without precision technology, farmers do have some use for information about yield responses to inputs on their fields, because these can be helpful as farmers make management decisions about how to use conventional uniform rate technology. Therefore, in the initial equilibrium a positive amount of information, $\underline{Q_I}^0$, is bought and sold, and the price per unit is P_I^0.

Following the discussions of Bullock and Bullock (2000) and Bullock, Lowenberg-DeBoer, and Swinton (2002), we assume that precision technology equipment and site-specific information about yield response are economic complements: roughly stated, the equipment is more valuable if the information is available, and the information is more valuable if the equipment is available.

A subsequent equilibrium is brought about when precision technology equipment is brought into the market ("invented"), which we model as a shifting down of the supply curve from S_E' to S_E''. This shift results in the emergence of the market for precision technology equipment. The subsequent equilibrium is shown with prices and quantities $\underline{P_E}^1$, $\underline{P_I}^1$, $\underline{Q_E}^1$, $\underline{Q_I}^1$. Note that in this subsequent equilibrium precision equipment is bought

[2] At this level of abstraction, we make no attempt to define in exact quantitative terms what we mean by "information." For a more rigorous treatment of the subject, see Bullock, et al. (2009).

and sold, and the market for site-specific information also exists. But both quantities $\underline{Q_E}^1$ and $\underline{Q_I}^1$ are shown as small. Basically, because the costs of producing information, as represented by S_I', have remained high, the demand for precision equipment and site-specific information are meager. The resultant equilibrium is much like the one observed today in actual markets for precision equipment and site-specific information: the markets exist, but to very limited extents.

The purpose of our study is to begin to think about whether the subsequent equilibrium illustrated in figure 1 has to be the long-run equilibrium for the precision technology and site-specific information markets. We believe that it may not be, because it might be possible to use precision technology itself to lower the costs of generating information about site-specific yield response information. Generation of cheaper and more site-specific information could affect the markets as shown in figure 2, where the use of precision technology to cheaply generate site-specific information is represented by a shift down and out in the information supply curve, from S_I' to S_I''. This drops the

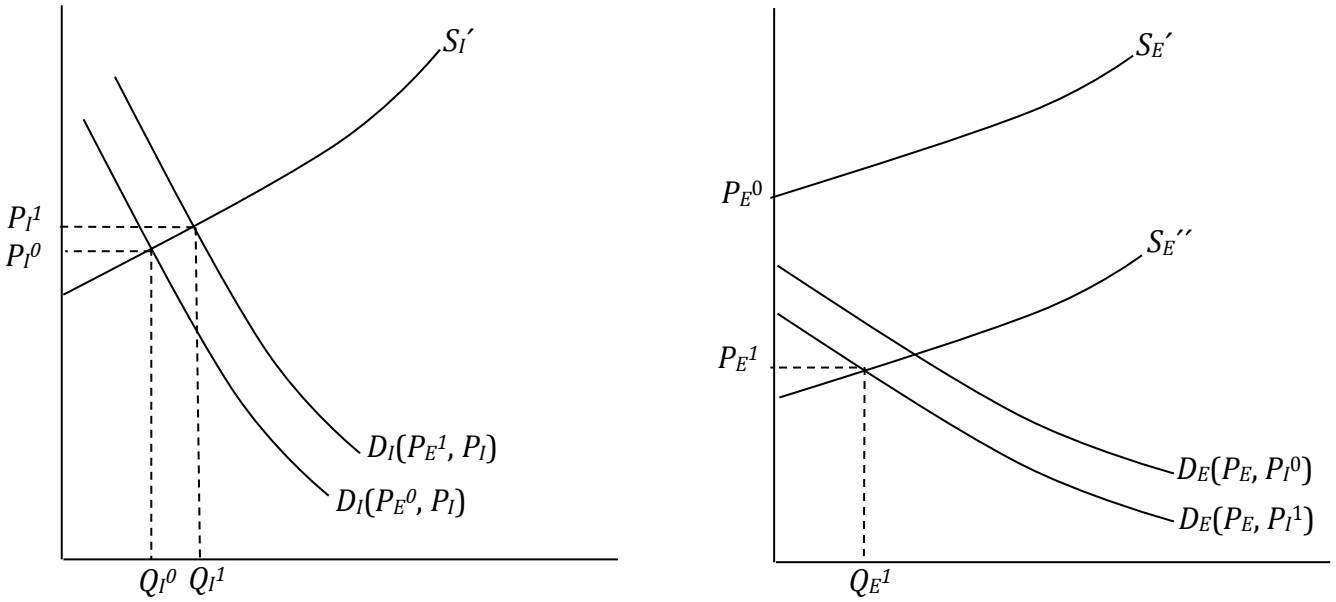

Figure 1. Markets for precision technology equipment and information

price of information, which in turn shifts out the demand for precision equipment. In the new equilibrium, quantities have risen to Q_E^2, Q_I^2. That is, precision technology has been used to supply the information needed to increase the value of precision technology in crop production. That is, precision technology has been used to create demand for itself.

Of course, in competitive markets for private goods, we might expect for (Q_I^1, P_I^1, Q_E^1, P_E^1) in figure 2 to represent the long-run equilibrium. Once precision technology equipment is brought to market, producers of information would increase their supply. But, as we will argue, much of the information pertinent to site-specific farm management is actually a public good. The policy implication is that, in the presence of precision technology, the public payoffs from government policies that encourage the generation of information about how crop yields respond to managed inputs are higher than in the absence of precision technology.

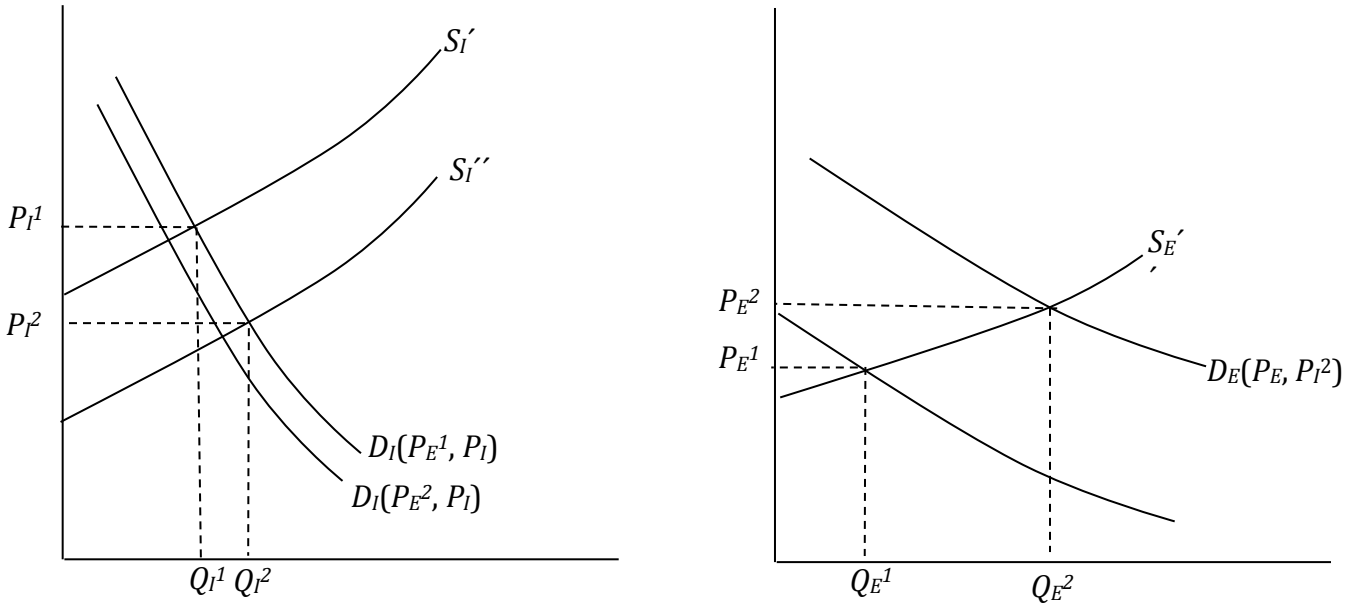

Figure 2. Costs of information production fall, and demand for precision equipment rises as a result.

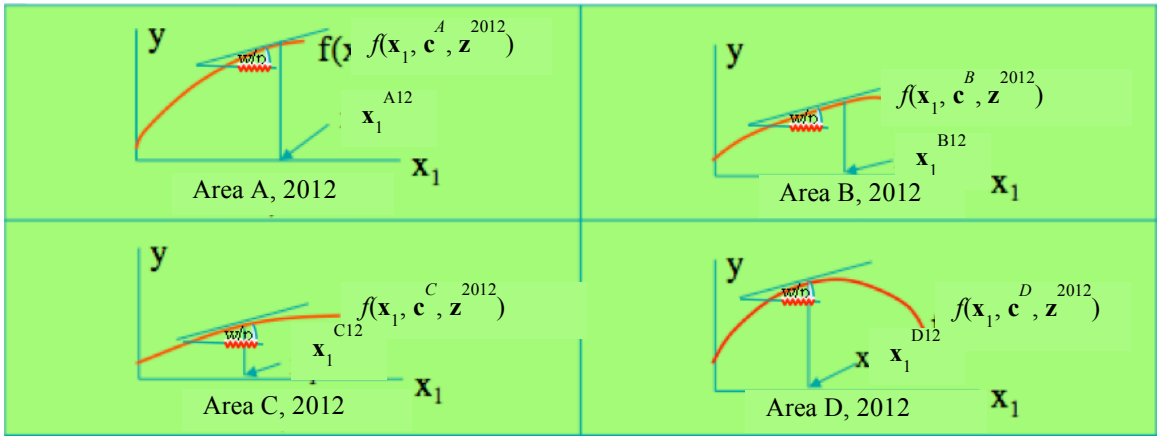

Figure 3. A "map" of site-and-year-specific production functions and profit-maximizing input application rates

Using Precision Technology to Lower the Cost of Generating Information about Yield Response

We maintain that given the current existence of precision technology, farmers possess insufficient knowledge of how their crop yields respond to efficiently manage fertilizer site-specifically. Next we discuss the types of information that farmers would need to use variable rate technology profitably.

An Illustration of What Farmers Would Need to Know to Use Variable Rate Technology Profitably

Figures 3 illustrates what we mean when we say that farmers have insufficient knowledge about yield response. Following Bullock, et al. (2000), we call yield y, and we suppose that yield depends on three classes of variables: $y = f(\mathbf{x}, \mathbf{c}, \mathbf{z})$, where the arguments of the function are described as follows.

1. The elements of the vector of variables **x** are managed inputs. These might include the quantities of various types of fertilizer applied during the year, the seeding rate, the tillage system used, etc.

2. The elements of the vector of variables **c** are field characteristics. These might describe the plot's topography (like the SCA and CTI variables that were found to be empirically significant in Bullock, et al. (2009)), or its soil type, organic matter content, or electro-conductivity.
3. The elements of the vector of variables **z** are temporally stochastic, non-managed factors of production. For the most part, these would be weather variables.[3]

In figure 3, we show "slices" of four "production functions," like those typically featured in undergraduate agricultural economics textbooks. For the purposes of illustration, we only consider the effects of one managed input, called x_1. Because the diagrams are two-dimensional, the values of the **c** and **z** vectors are held constant along the curves showing relationships between x_1 and y. Also for the purposes of illustration, we assume that the field under consideration has four distinct areas or "sites," labeled *A*, *B*, *C*, and *D*. We use $\mathbf{c}^i$ to denote the field characteristics of area $i = A, B, C, D$. We assume that each site's characteristics differ from each other site's: $\mathbf{c}^A \neq \mathbf{c}^B$, etc. These inequalities mean that even if every site receives the exact same weather, the sites' yield responses to the input x_1 differ. Thus, the four curves labeled $f(x_1, \mathbf{c}^A, \mathbf{z}^{2012})$, $f(x_1, \mathbf{c}^B, \mathbf{z}^{2012})$, $f(x_1, \mathbf{c}^C, \mathbf{z}^{2012})$, and $f(x_1, \mathbf{c}^D, \mathbf{z}^{2012})$ in figure 3 can be called *site-and-year-specific response functions*. If production functions did not vary spatially, there would be no need for precision agriculture technology, since the economically optimal rates of managed input application would be equal throughout a field in any given year. In contrast, because response functions do differ spatially, precision agriculture has the potential to increase farm profits. *But the profitability of precision agriculture technology depends crucially*

[3] Some factors of production are both temporally and spatially stochastic. To maintain simplicity, we gloss over this complication in our current treatment of the subject.

on the farmer's knowledge of the many site-specific production functions in his field. Farmers who do not know how yields respond to managed inputs, and how those responses differ spatially, cannot make good decisions about those inputs' application rates and how to vary them spatially.[4] That is, farmers need to know the response function "map," as illustrated in figure 3.

Current Methods of Research Provide Farmers with Little Knowledge of Their Response Function "Maps"

Unfortunately, at least for major production crops, farmers rarely, if ever, know enough about their yield response map to take much advantage of precision agriculture technology. One might suppose that such knowledge had already been provided by research in agronomics and agricultural production economics. Yield response function research has been conducted and reported for many decades now. But that research has been insufficient to provide farmers with the detailed knowledge needed, especially now that precision technology is available. The reason so little is known about how crops respond to inputs is that yield response research over the past century has relied on inadequate agronomic experimentation. Nearly all experiments producing yield-response data have been conducted over a short run (one or two years), at only a few locations, and mostly on very small plots. Partially as a consequence of the resulting data shortage, not enough is currently known about how crops respond to inputs, in

[4] As is explained in many undergraduate agricultural economics texts, if the farmer had complete knowledge of his four production functions, complete knowledge of the input and output prices w and p, and perfect foresight about the crop season's weather, the farmer working the field shown in figure 3 could maximize his fields profits by applying the input at rates x_1^{A13}, x_1^{B13}, x_1^{C13} and x_1^{D13}.

particular fertilizer, nor about how managed inputs interact with non-managed inputs (topology, soil type, weather variables) to allow researchers to make helpful site-specific management recommendations to farmers.[5]

Figure 4 helps illustrate why empirical studies of yield response have so far provided only very limited useful information. The studies were commonly subject to omitted variable bias and insufficient variation in the characteristics (**c**) and weather (**z**) data. The insufficient characteristics data arose because the empirical studies' agronomic experiments were generally conducted on small plots in an isolated geographic region. Since field characteristics tend to be spatially correlated, conducting a small-plot study in limited area is not likely to provide much variation in the **c** data. Insufficient variation in the observed weather data occurred because the agronomic experiments were almost always very short-run (one or two years) in nature, and covered a relatively small geographic area.

The data points shown in figure 4 are a fair (not particularly exaggerated) representation of the data often used in applied yield response studies. For the purposes of the illustration, we assume in figure 4 that an agronomic experiment was run in four areas of a field over two years. The curve labeled $f(x_1, \mathbf{c}^i, \mathbf{z}^j)$ represents the true site-and-year specific response function for area i in year j. In each year 2012 and 2013 and on each block A, B, C, and D, it is assumed that an agronomic experiment was run, where each block had five plots, on which quantities of x_1 of 50, 100, 150, 200, and 250 units

[5] It is not surprising that insufficient work was done on yield response. Since site-specific information and precision technology are economic complements, site-specific information simply was of relatively little worth before the advent of precision technology. That is, just as the technology is not worth much without the information, the information was not worth much without the technology.

were applied. Yield data was recorded for each area in each year. The resulting (x_1, y) data points from each experiment-year are shown. Various difficulties must be confronted when using data like that displayed in figure 4 to develop effective management guidelines.

1. It is possible to take the data and estimate site-and-year specific response functions. Since the **c** and **z** variables are constant in a given location and a given year, conceptually the response of a site-and-year specific response function to managed inputs can be measured without omitted variable bias. That is, defining the site i's response function in year j as $f_{ij}(\mathbf{x}) \equiv f(\mathbf{x}, \mathbf{c}^i, \mathbf{z}^j)$, as long as the managed input variables **x** are included in the regression, then $\mathbf{c}^i$ and $\mathbf{z}^j$ need not be observed to estimate $f_{ij}(\mathbf{x})$. The estimated function, however, does not necessarily provide information needed to make good recommendations about the management of site i in other years, or of other sites in any year. Basically, an estimate of $f_{ij}(\mathbf{x})$ can provide information on how site i should be managed if the farmer were to know with certainty that in the future the weather variables would take on the same value that they took on in year j. Since the farmer cannot know the weather with certainty, and since different weather in a different year would result in a different site-and-year-specific response function, the estimate of $f_{ij}(\mathbf{x})$ is of limited use.
2. Even if the weather variables are included in the estimation, any attempt to estimate a site-specific response function $f_i(\mathbf{x}, \mathbf{z}) \equiv f(\mathbf{x}, \mathbf{c}^i, \mathbf{z})$, if the agronomic experiments run are of short duration and are conducted in nearby geographic areas , there may be insufficient variation in the **z** data to permit accurate estimation of $f_i(\mathbf{x}, \mathbf{z})$. And even

if good estimation of this site-specific response function is possible, analysis of it may help little in the development management recommendations for other sites.

3. Similarly, even if the characteristics variables are included in the estimation, in attempt to estimate a year-specific response function $g_j(\mathbf{x}, \mathbf{c}) \equiv f(\mathbf{x}, \mathbf{c}, \mathbf{z}^j)$, if the agronomic experiments run are of short duration and are conducted in nearby geographic areas , there may be insufficient variation in the **z** data to permit accurate estimation of $f_i(\mathbf{x}, \mathbf{c})$. And even if good estimation of this site-specific response function is possible, analysis of it may help little in the development management recommendations for other years. Thus, the estimation can answer the *ex post* management question, "What management plan would have been best in the year of the experiment?", but can help to answer the more pertinent *ex ante* question, "Which management plan is best for the upcoming year?" if the statistical properties of the vector random weather variables are known, along with producer risk preferences.

More data than just **x** data and yield data are needed to derive decent management recommendations. Ideally, management recommendations would come form experiments run at many different places over many different years, and therefore provide a good deal of data and wide variation in **c** and **z**.

Following Bullock, et al. (2009), we propose two methods by which precision technology could be used to inexpensively generate information about site-specific yield response: a "meta-function approach" and a "same-farm approach." Both approaches rely on on-farm, large-scale, long-term agronomic experimentation.

Precision Technology Has Been Proved to Greatly the Cost of Implementing Agronomic Experiments

We call the function $f(\mathbf{x}, \mathbf{c}, \mathbf{z})$ discussed in the previous section "Nature's meta-response function." This function describes how a crop's yield depends on everything that it depends on. Conceptually, if we have included all relevant factors in **x, c**, and **z**, f is universal. It describes how yield responds to factors of production in all places and at all times. If the function $f(\mathbf{x}, \mathbf{c}, \mathbf{z})$ were known, the spatial and temporal distributions of farm's c and z variables were known, the distributions of relevant input and output prices were known, and a farm household's preferences were known, then in principal the farm's maximization problem could be written out and solved. In this sense, optimal input application rates could be calculated for every site on the farm.

In general, of course, information about the meta-response function $f(\mathbf{x}, \mathbf{c}, \mathbf{z})$ is limited. But generation of information about is $f(\mathbf{x}, \mathbf{c}, \mathbf{z})$ valuable. One strategy for

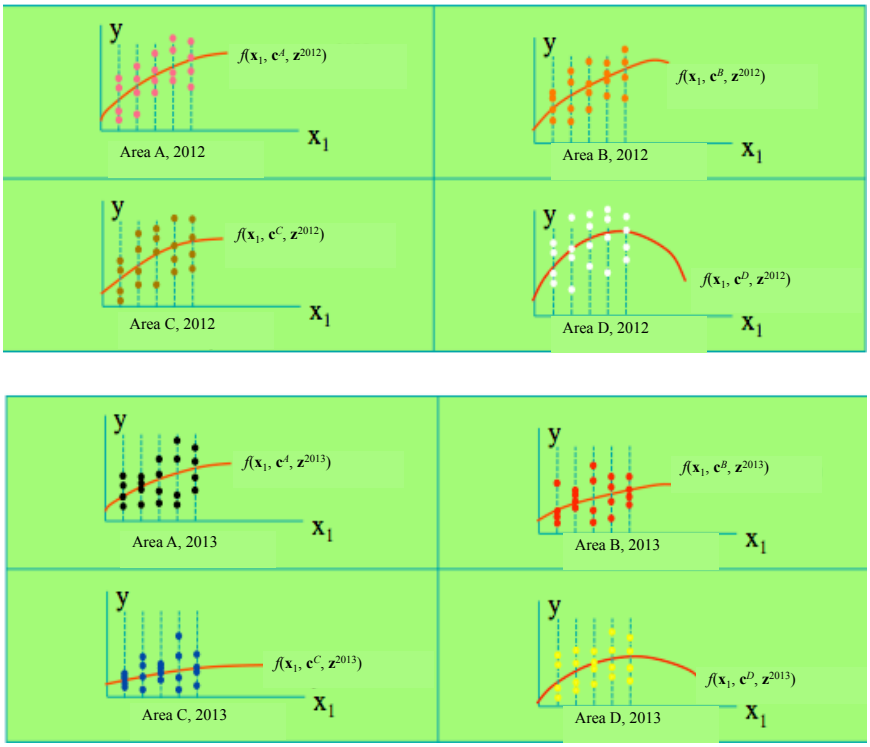

Figure 4. The type of data often used in response function estimations.

generating information about $f(\mathbf{x}, \mathbf{c}, \mathbf{z})$ would be to run agronomic experiments in many places for many years, in order to generate the variation is all three types of variables necessary for good estimation of the function.

Ultimately, estimating $f(\mathbf{x}, \mathbf{c}, \mathbf{z})$ would require producing and making available to researchers a central data bank, conceptually like the one illustrated in table 1. As is to be explained, such a data bank could be generated both by experts *and by farmers*. Farmers would be motivated to participate because by doing so they would gain access to interpretation of the data from their farms, and to the managerial recommendations that can be derived from that data. The part that farmers play in generating the data would occur totally in the "background" of their usual production practices. That is, farmers would do very little different from what they usually do in the production and monitoring of their crops. Experts (either USDA personnel, university researchers and their assistants, or private consultants) could be easily trained to use currently available GIS-related software to design agronomic experiments, which would be transferred to the farmer's computer, and then implemented using existing GIS software to guide the variable rate application of managed inputs. Basically, the farmer would simply do his planting, fertilizing, and harvesting as usual. But a computer program would control the application rates of his seed, fertilizer, herbicides, etc., and as the farmer drove his planter, tractor, or combine through his field, the computer program would use GPS equipment to determine the farmer's (nearly) exact location in the field at every point in time. The computer program would control variable rate applicators to vary the rates of application according to the agronomic experiments design. (The experiment would not need to be performed everywhere on a field, but rather could be confined to spatially diverse areas

making up some fraction of the field's area.) While the farmer followed his usual harvesting practices, yield monitors paired with GPS equipment would provide data on yields for plots all over the farmer's field. All the data could be transferred to the central data bank in real time.

No new technology or farm practices need to be developed to implement the plan described above. The necessary technology already exists; at the farm-level the whole procedure has been performed before with cooperating farmers with much success. Bullock, et al. (2009) and Ruffo, et al. (2008) report their generation and analysis of the kind of data showed in table 1. They designed agronomic experiments using the University of Illinois's EFRA software (an extension of ESRI's ArcInfo GIS software, and cooperating farmers "gathered" the data while planting, fertilizing, and harvesting in their usual ways. Cooperating farmers were motivated to participate in the data generation because they learned valuable information about crop yield response to management inputs, and how the yield response varied spatially in their fields. This information allows them to more profitably manage their fields on a site-specific basis.

A Monte Carlo Simulation of the Relationships between Application Rate Technologies and Information from Agronomic Experimentation

Using Mathematica (Wolfram 2013), we conducted Monte Carlo simulations to initiate an examination of the relationships between precision input application rate technologies and information about yield response.

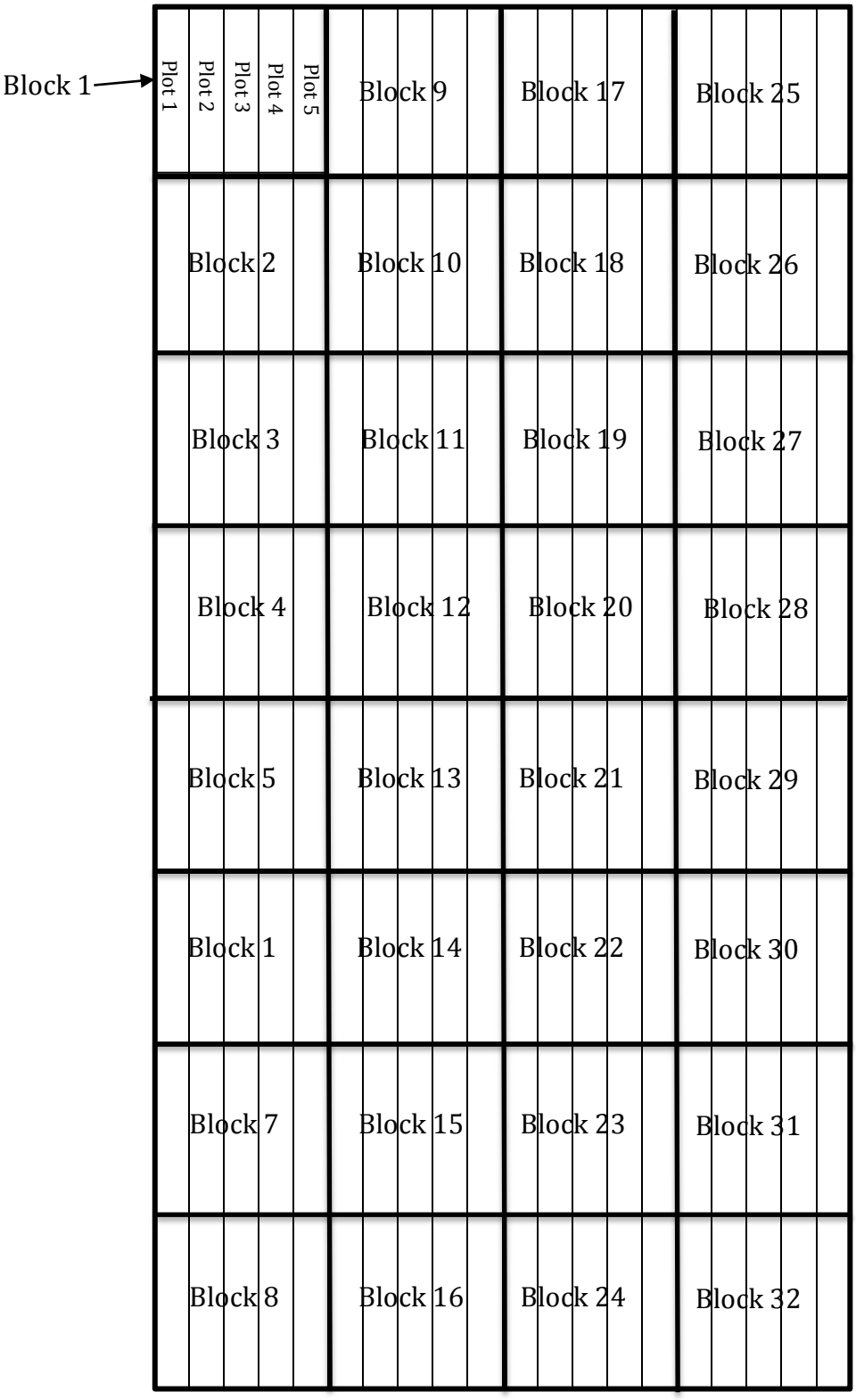

Figure 5. Experimental Design of thirty-two 300 ft × 300 ft blocks, each with five 60 ft × 300 ft plots

Experiment Design

Our experimental design involved a "true" meta-response function, a field with spatially autocorrelated characteristics, random assigments of nitrogen fertilizer rates to field blocks, and resultant experimental yields.

The Simulations' "True" Meta-response Function *f*

Throughout our analysis, we used as our "true" yield response function the estimated response function (in kg/ha) reported in table 1 of Bullock, et al. (2009, p. 218) for Farmer *D* in Champaign County, Illinois:

$$(1)\quad f(N,M,S,I,\varepsilon) \equiv \beta_0 + \beta_N N + \beta_D D + \beta_I I + \beta_M M + \beta_R R + \beta_S S + \beta_{NN} N^2 + \beta_{NI} NI + \beta_{NM} NM + \beta_{NS} NS + \beta_{NNI} N^2 I + \varepsilon,$$

where $\beta_0 = 5524$, $\beta_N = 154.97$, $\beta_{NN} = -0.4461$, $\beta_M = -108.5$, $\beta_S = -194.7$, $\beta_I = 52.33$, $\beta_D = -2440$, $\beta_R = 0$, $\beta_{NM} = 0.1624$, $\beta_{NS} = 0.7537$, $\beta_{NI} = -0.4866$, $\beta_{NNI} = 0.001318$, and ε is a normally-distributed error term with mean zero a standard deviation of 200 kg/ha.

The variables in (1) are explained in detail in Ruffo (2006), and are *N* (nitrogen fertilization rate, in kg/ha), *D* (a dummy variable, in our case always equal to one), *M* (May precipitation, in mm), *R* (a dummy variable, in our case always equal to zero), *S* (the value of the stream power index (*SPI*)), and *I* (the value of the Illinois Nitrogen Soil Test (ISNT) result). In reality, the *ISNT* is an index of soil-borne nitrogen, and changes over time. For the purposes of our simulations, we treat the *ISNT* values as exogenous and constant over time.

The Field

In each Monte Carlo run, we simulated a farm field with spatially autocorrelated production characteristics. As illustrated in figure 5, we assumed field dimensions of

1200 ft $\times$ 2400 ft. For the agronomic experiments, we divided the field into thirty-two 300 ft $\times$ 300 ft square blocks, each containing five 60 ft $\times$ 300 ft plots. We assumed that the model's geospatial characteristics (the SPI and ISNT values) varied by block, but did not vary within any block. For each Monte Carlo round separately, we generated a spatially autocorrelated ISNT map and a spatially autocorrelated SPI map as follows. First, in reference to Ruffo, et al. (2006), we made thirty-two independent draws from a normal distribution with mean zero and a standard deviation of 40.0. We placed these thirty-two values in a vector called δ. We then generated the ISNT map through the autoregressive process in equation (2) below. The associated data generating process statement is in (3). Equations (2) and (3) are versions of equations (1.10) and (1.11) of LeSage and Pace (2009, p. 13).

(2) $v_{ISNT} = \alpha\iota_{32} + \rho W \cdot v_{ISNT} + \delta$

(3) $v_{ISNT} = \left[I_{32} - \rho W\right]^{-1} \iota_{32}\alpha + \left[I_{32} - \rho W\right]^{-1} \delta$,

In (2) and (3), ι_{32} is a vector of thirty-two values of 1.0, I_{32} is a 32×32 identity matrix, ρ = 0.5 is the dependence parameter, α = 235(1-ρ) = 117.5 is a constant. W is a 32×32 normalized matrix of spatial weights, which has zeroes on the diagonal, and the distance between block centroids, raised to the power of -4, elsewhere. The SPI map was generated similarly, but assuming a mean of 1.5 and a standard deviation of 0.3.

Simulated Data Sets

Given the characteristics map generated in a Monte Carlo run, we then created a 30-year series of agronomic experiments on the field, which in every plot in every block in every year resulted in a simulated experimental yield. For every year of modeled experimentation, we obtained a value for May precipitation, in millimeters, a random

variable drawn from a normal distribution with mean $\bar{M}$ = 100 and standard deviation σ_M = 10. We assumed that in a given year *t*, the same amount of rain per ha, M_t, fell on every plot in the field. This amount of rain was assumed to fall on every plot of the field in the given year. Then for each of the 160 plots in the field, a random yield disturbance term was drawn from a normal distribution with mean 0 and standard deviation 200 kg/ha. Then, in every block, the five plots were randomly assigned one of five nitrogen fertilization rates: 125, 150, 175, 200, and 225 kg/ha. In each Monte Carlo run, for any given block *b*, all the plots in the block were assigned the same ISNT value I_b, and the same SPI value S_b, and these values do not change over time. Every plot was subject to a random yield disturbance every year, ε_{bjt}, which is a value of the random variable ε, drawn from a normal distribution mean $\bar{\varepsilon}$ = 0 Mg/ha, and standard deviation σ_e = 200: $\varepsilon \sim N(\bar{\varepsilon}, \sigma_e)$. Given the plot's block's assigned levels of ISNT and SPI values, for every plot in every block in every year of the experiment, a yield value was generated according to equation (1). Thus, every observation in the resultant simulated data set had a value for the applied nitrogen fertilization rate, a value for May precipitation, a value for its random disturbance term, and a yield. Since the experiment was simulated for five plots on each of 32 blocks for 30 years, the resultant number of observations was 5×32×30 = 4800. The data set took on the appearance shown in table 1. We call this data set the "all-block, 30-year" data set.

To simulate varying lengths of agronomic experimentation, we simply omitted years from the all-block, 30-year data set. We thus created a 29-year data set, which contained the observations represented in Table 1, but from the last 29 years. In this manner we generated data sets simulating experiments of 4, 5, …. , 29, and 30 years of length.

To farm block-specifically, a farmer would use information generated from the experiments conducted on every block separately. The block-specific data was used estimate block-specific response functions. For example, block 7's 30-year data set would have 5 observations per year for 30 years, and take on the form of table 2.

Information

The data produced by the agronomic experimentation can provide information relevant to an expected-profit-maximizing producer's choices. But "how much" information is generated would depend on the length of the experimentation. We say that a producer has *full information* when he knows the distributions from which the yield disturbance ε will be drawn and from which May precipitation M will be drawn, and each block's "true" block-specific response function. That is, when for $b = 1, \ldots, 32$ he knows the values of $\gamma_{0,b}$, $\gamma_{N,b}$, $\gamma_{M,b}$, $\gamma_{NM,b}$, and $\gamma_{NN,b}$ in the following:

$$(2)\quad f_b(N,M,\varepsilon) \equiv f(N,M,S_b,I_b,\varepsilon) \equiv \beta_0 + \beta_N N + \beta_D + \beta_I I_b + \beta_M M + \beta_S S_b + \beta_{NN} N^2 + \beta_{NI} N I_b + \beta_{NM} NM + \beta_{NS} N S_b + \beta_{NNI} N^2 I_b + \varepsilon$$

$$\equiv \underbrace{\left[\beta_0 + \beta_D + \beta_I I_b + \beta_S S_b\right]}_{\gamma_{0,b}} + \underbrace{\left[\beta_N + \beta_{NI} I_b + \beta_{NS} S_b\right]}_{\gamma_{N,b}} N + \underbrace{\beta_M}_{\gamma_{M,b}} M + \underbrace{\beta_{NM}}_{\gamma_{NM,b}} NM + \underbrace{\left[\beta_{NN} + \beta_{NNI} I_b\right]}_{\gamma_{NN,b}} N^2 + \varepsilon .$$

Table 1. Form of the simulated "all-block, 30-year" data set

Obs.	year, t	block, b	plot, j	experimental fertilization rate, N_{tbj}	May precipitation, M_t	yield, $q_{t.b,j}$
1	1	1	1	N_4=200	M_1	$q_{1,1,1}$
2	1	1	2	N_5=225	M_1	$q_{1,1,2}$
3	1	1	3	N_3=175	M_1	$q_{1,1,3}$
4	1	1	4	N_2=150	M_1	$q_{1,1,4}$
5	1	1	5	N_1=125	M_1	$q_{1,1,5}$
6	1	2	1	N_3=175	M_1	$q_{1,2,1}$
7	1	2	2	N_2=150	M_1	$q_{1,2,2}$
8	1	2	3	N_5=225	M_1	$q_{1,2,3}$
9	1	2	4	N_1=150	M_1	$q_{1,2,4}$
10	1	2	5	N_4=200	M_1	$q_{1,2,5}$
11	1	3	1	N_4=200	M_1	$q_{1,3,1}$
	⋮	⋮	⋮	⋮	⋮	⋮
4796	30	32	1	N_1=125	M_{30}	$q_{30,32,1}$
4797	30	32	2	N_4=200	M_{30}	$q_{T,32,2}$
4798	30	32	3	N_2=150	M_{30}	$q_{30,32,3}$
4799	30	32	4	N_5=225	M_{30}	$q_{30,32,4}$
4800	30	32	5	N_3=175	M_{30}	$q_{30,32,5}$

Table 2. Form of block 7's simulated 30-year data set

Obs	year, t	block, b	plot, j	experimental fertilization rate, N_{tbj}	May precipitation, M_t	yield, $q_{t.b,j}$
1	1	7	1	N_3=175	M_1	$q_{1,7,1}$
2	1	7	2	N_5=225	M_1	$q_{1,7,2}$
3	1	7	3	N_4=200	M_1	$q_{1,7,3}$
4	1	7	4	N_1=125	M_1	$q_{1,7,4}$
5	1	7	5	N_2=150	M_1	$q_{1,7,5}$
6	2	7	1	N_4=200	M_1	$q_{2,7,1}$
7	2	7	2	N_1=125	M_1	$q_{2,7,2}$
8	2	7	3	N_2=150	M_1	$q_{2,7,3}$
9	2	7	4	N_5=225	M_1	$q_{2,7,4}$
10	2	7	5	N_3=175	M_1	$q_{2,7,5}$
11	3	7	1	N_2=150	M_1	$q_{3,7,1}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
146	30	7	1	N_5=225	M_{30}	$q_{30,7,1}$
147	30	7	2	N_1=125	M_{30}	$q_{30,7,2}$
148	30	7	3	N_4=200	M_{30}	$q_{30,7,3}$
149	30	7	4	N_3=175	M_{30}	$q_{30,7,4}$
150	30	7	5	N_2=150	M_{30}	$q_{30,7,5}$

To have full information, the producer need not know the values of the β parameters or of I_b and S_b for any of the blocks. It suffices simply to know for each block b the reduced form, $f_b(N, M, \varepsilon)$, of the meta-response function f.

A farmer has *partial information* when he does not know the true block-specific response functions f_b, but has access to the experimental data, with which he (or some economist!) can estimate the f_b functions. The amount of information depends on the number of years of experimental data that the producer has access to.

Use of Full Information with Site-specific Technology

The producer is assumed to maximize expected profits. In an arbitrary year t, under full information, the producer using site-specific technology for fertilization solves,

$$(3)\quad \underset{N_1,\ldots,N_B\geq 0}{Max}\ E\left\{p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[f_b\left(N_b,M_t,\varepsilon_{bjt}\right)-wN_b\right]\right\},$$

where p is the output price and w is the price of N fertilizer. The solution to (3) is the *vector of ex-ante optimal site-specific nitrogen fertilization rates under full information*:

$$(4)\quad \mathbf{N}^{ssfi}\left(p,w\right)\equiv\left(N_1^{ssfi}\left(p,w\right),\ldots,N_B^{ssfi}\left(p,w\right)\right)\equiv\underset{N_1,\ldots,N_B\geq 0}{\arg Max}\,E\left\{p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[f_b\left(N_b,M_t,\varepsilon_{bjt}\right)-wN_b\right]\right\}.$$

Ex-ante maximized profits under site-specific technology with full information is then,

$$(5)\quad \pi^{ssfi}\left(p,w\right)\equiv E\left\{p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[f_b\left(N_b^{ssfi}\left(p,w\right),M_t,\varepsilon_{bjt}\right)-wN_b\right]\right\},$$

which, because M and ε enter the response function linearly, can be rewritten,

$$(6)\quad \pi^{ssfi}\left(p,w\right)\equiv p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[f_b\left(N_b^{ssfi}\left(p,w\right),\bar{M},0\right)-wN_b^{ssfi}\left(p,w\right)\right]$$

Use of Full Information with Uniform Technology

A producer who will use uniform-rate technology under full information has the following demand function for the optimal uniform application rate:

$$(6)\quad N^{unfi}(p,w) \equiv \arg\underset{N\geq 0}{Max}\, E\left\{ \underbrace{p\sum_{b=1}^{B}\sum_{j=1}^{5} f_b\left(N, M_t, \varepsilon_{bjt}\right)}_{f_{wf}\left(N, M_t, \boldsymbol{\varepsilon}\right)} - w\sum_{b=1}^{B}\sum_{j=1}^{5} N \right\},$$

where $f_{wf}(N, M, \boldsymbol{\varepsilon})$ is the whole-field-response-to uniform-N function. Ex-ante maximized profits under uniform technology with full information are then,

$$(7)\quad \pi^{unfi}(p,w) \equiv E\left\{ p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[f_b\left(N^{unfi}(p,w), M_t, \varepsilon_{bjt}\right) - wN^{unfi}(p,w)\right]\right\}$$

$$\equiv p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[f_b\left(N^{unfi}(p,w), \bar{M}, 0\right) - wN^{unfi}(p,w)\right].$$

Because the maximization problem in (4) is less constrained than the one in (6), then under full information, profits under site-specific management cannot be lower, and in general will be higher than under uniform management. (Here we are not considering the costs of the technology itself.) The difference between these, which we call the *ex-ante value of site-specific technology under full information*, is

$$(8)\quad v^{ssfi}(p,w) \equiv \pi^{ssfi}(p,w) - \pi^{unfi}(p,w) \geq 0.$$

Use of Partial Information with Site-specific Technology

A farmer with partial information has access estimated yield response functions. Since the estimates are generated from the data from agronomic experiments, then in some sense the "amount" of information that the farmer has access to depends on the number of years that the agronomic experiment has been run. Since the number of observations

in the data set increases with the number of experiment years, then the trend will be for estimates of the f_b functions to improve as experiments are run for more years.[6]

To model different information settings, we simulate the effects of conducting agronomic experiments (to be explained in more detail below) for different numbers of years. We denote the number of experiment years by the variable T, and we let T vary between 4 and 30, depending on the simulation. Using "experiment T" to denote the experiment run for T years, we estimated parameters of the reduced-form site-specific response functions $\gamma_{0,b}$, $\gamma_{N,b}$, and $\gamma_{NN,b}$ in (2). Call these estimates $\hat{\gamma}^T_{0,b}$, $\hat{\gamma}^T_{N,b}$, and $\hat{\gamma}^T_{NN,b}$ for $b = 1, \ldots, B = 32$ and $T = 4, 5, \ldots. 30$. Then the estimated reduced-form site-specific response for block b with T years of experiments is,

$$(9)\quad \hat{f}^T_b\left(N,M\right) \equiv \hat{\gamma}^T_{0,b} + \hat{\gamma}^T_{N,b}N + \hat{\gamma}^T_{M,b}M + \hat{\gamma}^T_{NM,b}NM + \hat{\gamma}^T_{NN,b}N^2 .$$

Letting $\bar{M}$ denote expected May precipitation, a producer using uniform technology is assumed to maximize expected profits from the whole field. If the only information that he has is the estimated whole-field response function, then his objective is,

$$(10)\quad \underset{N_1,\ldots,N_B\geq 0}{Max}\; E\left\{ p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[\hat{f}^T_b\left(N_b,M_t\right) - wN_b\right]\right\},$$

which can be rewritten

$$(11)\quad \underset{N_1,\ldots,N_B\geq 0}{Max}\; p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[\underbrace{\hat{\gamma}^T_{0,b} + \hat{\gamma}^T_{N,b}N + \hat{\gamma}^T_{M,b}\bar{M} + \hat{\gamma}^T_{NM,b}N\bar{M} + \hat{\gamma}^T_{NN,b}N^2}_{\hat{f}^T_b\left(N_b,\bar{M}\right)} - wN_b\right].$$

[6] Outcomes of an additional year's experiment are random, and if the values of the draws from the random variables' distributions are sufficiently unusual, then the estimates of the f_b functions can be made worse be the additional year's data. Still, as experiment years are added, the trend of the estimates of the f_b functions will be one of increasing accuracy.

The solution to the problem above is the *vector of ex-ante optimal block-specific nitrogen fertilization rates under information from an experiment of length T*:

$$(12)\ \mathbf{N}^{ssT}(p,w) \equiv \left(N_1^{ssT}(p,w),...,N_B^{ssT}(p,w)\right) \equiv \underset{N_1,...,N_B \geq 0}{\arg Max}\ p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[p\hat{f}_b^T\left(N_b,\bar{M}\right) - wN_b\right]$$

Expected ex-ante maximized profits under site-specific technology with information from *T* years of experiments are then,

$$(13)\ \pi^{ssT}(p,w) \equiv E\left\{p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[f_b\left(N_b^{ssT}(p,w),M_t,\varepsilon_{bjt}\right) - wN_b\right]\right\}$$

$$\equiv p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[f_b\left(N_b^{ssT}(p,w),\bar{M},0\right) - wN_b^{ssT}(p,w)\right].$$

(Note that the true profits with incomplete information depend on the true response functions f_b, not the estimated response functions $\hat{f}_b^T$.) The per-ha *value of the information gap between site-specific management with full information and site-specific management with information from T years of experiments* is therefore,

$$(14)\ v^{inss,T}(p,w) \equiv \pi^{ssfi}(p,w) - \pi^{ssT}(p,w).$$

The per-ha *marginal value of information the T^{th} year's experiment given site-specific technology* is,

$$(15)\ mv^{inss,T}(p,w) \equiv \pi^{ss,T}(p,w) - \pi^{ss,T-1}(p,w),\ T = 5,\ \ldots,\ 30.$$

Use of Partial Information with Uniform Technology

The producer operating with uniform rate technology with information from *T* years of the agronomic experiment solves

$$(16)\ \underset{N \geq 0}{Max}\ E\left\{p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[\hat{f}_b^T\left(N,M_t\right) - wN\right]\right\},$$

which can be rewritten

$$(17)\quad \underset{N\geq 0}{Max}\; p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[\underbrace{\hat{\gamma}_{0,b}^{T}+\hat{\gamma}_{N,b}^{T}N+\hat{\gamma}_{M,b}^{T}\bar{M}+\hat{\gamma}_{NM,b}^{T}N\bar{M}+\hat{\gamma}_{NN,b}^{T}N^{2}}_{\hat{f}_{b}^{T}\left(N,\bar{M}\right)}-wN\right].$$

The solution to the problem above is the *ex-ante optimal whole-field nitrogen fertilization rate under information from an experiment of length T*:

$$(18)\quad N^{unT}\left(p,w\right)\equiv \arg\underset{N\geq 0}{Max}\; p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[p\hat{f}_{b}^{T}\left(N,\bar{M}\right)-wN\right]$$

Expected ex-ante maximized profits under site-specific technology with information from T years of experiments are then,

$$(19)\quad \pi^{unT}\left(p,w\right)\equiv E\left\{p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[f_{b}\left(N^{unT}\left(p,w\right),M_{t},\varepsilon_{bjt}\right)-wN^{unT}\right]\right\}.$$

$$\equiv p\sum_{b=1}^{B}\sum_{j=1}^{5}\left[f_{b}\left(N^{unT}\left(p,w\right),\bar{M},0\right)-wN^{unT}\left(p,w\right)\right].$$

(Again, note that the true profits with incomplete information depend on the true response functions f_b, not the estimated response functions $\hat{f}_b^T$.) The per-ha *value of the information gap between uniform management with full information and uniform management with information from T years of experiments* is therefore,

$$(20)\quad v^{inun,T}\left(p,w\right)\equiv \pi^{unfi}\left(p,w\right)-\pi^{unT}\left(p,w\right).$$

The per-ha *marginal value of information the T^{th} year's experiment given uniform technology* is,

$$(21)\quad mv^{inun,T}\left(p,w\right)\equiv \pi^{un,T}\left(p,w\right)-\pi^{un,T-1}\left(p,w\right),\; T=5,\ldots,30.$$

Finally, we compare the net revenues of the site-specific producer with some amount of partial information to those of a uniform-rate producer wth some amount of

information. This is the *ex-ante value of site-specific technology under partial information*:

(22) $v^{ssT}(p,w) \equiv \pi^{ssT}(p,w) - \pi^{unT}(p,w) \geq 0.$

This value is the one most pertinent to producers deciding which technology to employ. If the difference in the costs of the technologies is less than the difference in the net revenues as defined to (22), then the producer's best choice is to adopt site-specific technology.

Results

We conducted 100 Monte Carlo simulations using the steps outlined above. Throughout the analysis we assumed a corn output price of p^0 = \$0.246/kg (\$6.25/bu) and a nitrogen fertilizer price of w^0 = \$\$0.80/kg for anhydrous ammonia. Our objective was to examine the values and marginal values of information from years of experimentation, given uniform and site-specific technologies, and to measure the value of site-specific technology as compared to uniform technology, given varying levels of information. We simply used OLS in all regressions. Obviously, our next step will be to conduct the simulations using spatial econometrics.

We found that from all 100 Monte Carlo runs, the mean of the value of site-specific technology under full information, as defined in (8), was $v^{ssfi}(p^0, w^0)$ = \$2.19/ha. That is, by using site-specific variable rate technology instead of uniform rate technology, a producer who knew every block's true response function, $f_b(N, M, \varepsilon)$ could expect to be able to increase annual net revenues (revenues minus N fertilizer costs) by \$2.19/ha. In the context of our model, in which there is no change in the response functions over time,

this advantage provided by the site-specific technology would be gained every year. The present value of a stream of annual net revenue increases of $2.19 per hectare per annum would be over $40/ha at an assumed 5% discount rate, and over $100/ha at an assumed 2% discount rate. It seems likely that these types of returns would easily cover the capital and maintenance costs of the variable rate equipment. In this sense, given that the producer has full information, site-specific technology "pays" in our model.

The results of our model so far make it less clear that replacing uniform technology with site-specific technology could pay for itself when information is not full and free, but rather must be garnered from agronomic experimentation. Figure 6 shows $v^{inss,T}$ from (14), how much more farmers using site-specific technology make in net revenues than do those using site-specific technology but possessing only information from agronomic experiments of various lengths. The farmer who had full information and used precision technology could expect to enjoy net revenues more than $10/ha greater than the net revenues of the farmer using precision technology having only four years of experimental data. Even for farmers with over twenty years of experimental data, net revenues from precision farming would be about $3/ha less than the net revenues of farmers using precision technology with full information. Clearly, at least in our simulations, information about crop response to managed inputs is key to the profitability of variable rate technology.

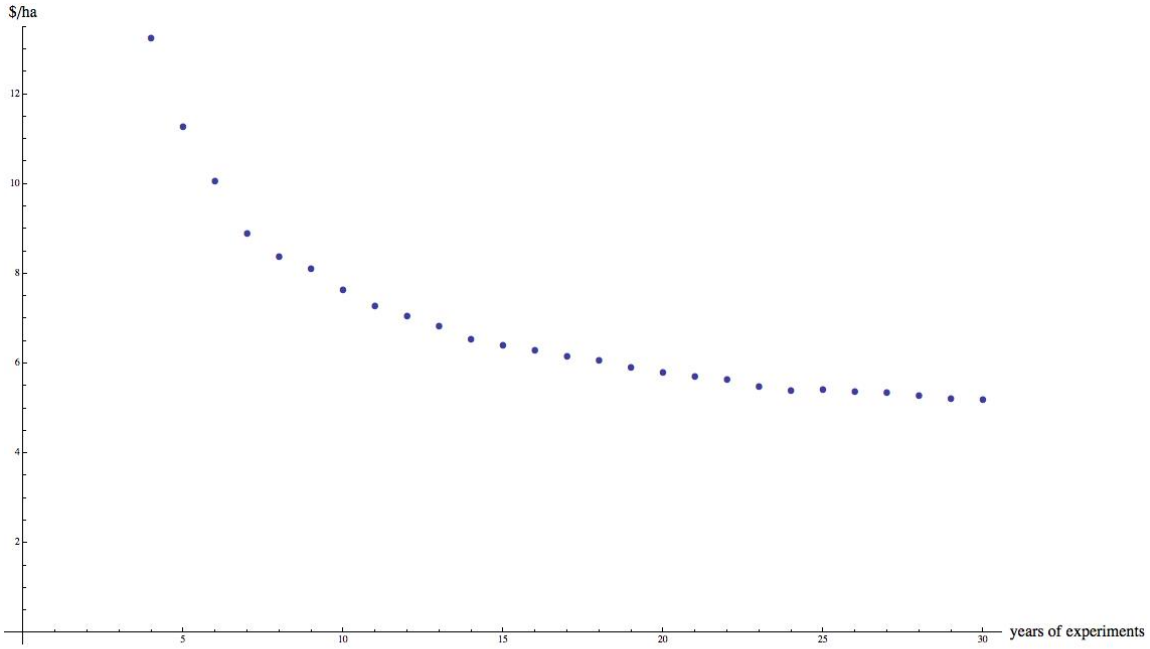

Figure 6. Difference in net revenues between a site-specific producer with full information and a site-specific producer only with information from agronomic experiments

Figure 7 is similar to figure 6, except that it shows $v^{inun,T}$ from (20), the difference in net revenues of farmers who use uniform rate technology with full information compared to farmers who use uniform rate technology with information from agronomic experiments. The information is worth much less to the uniform rate farmer than to a variable rate farmer. The farmer who only has information from an experiment of four years in length only makes net revenues about $0.60/ha less than the uniform rate farmer with full information.

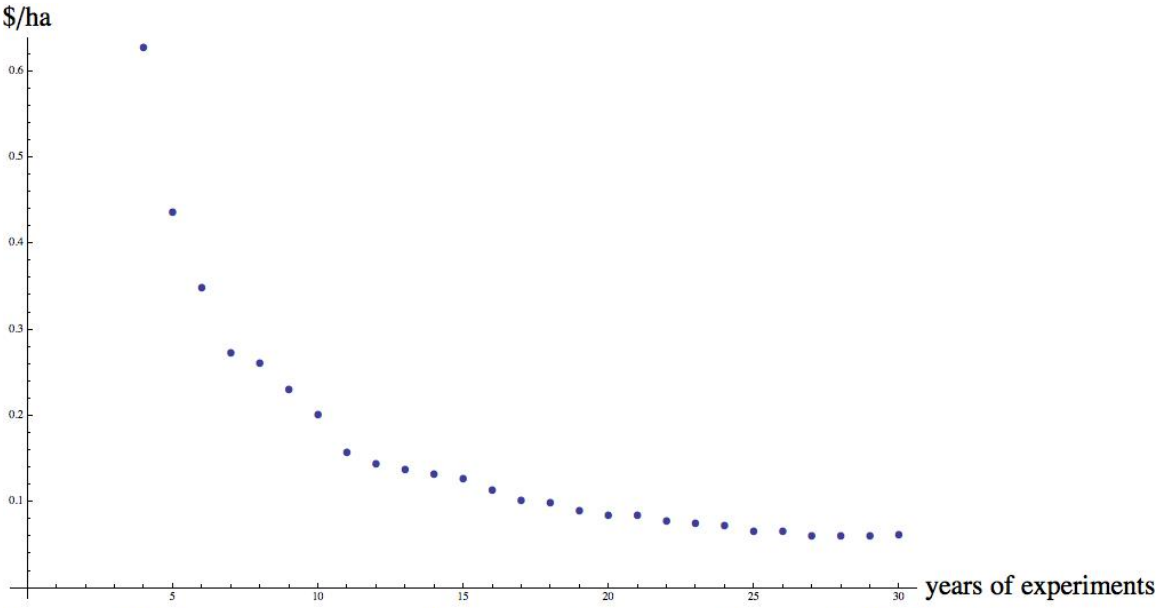

Figure 7. Difference in net revenues between a uniform-rate producer with full information and a uniform-rate producer with information from agronomic experiments

Figure 8 shows $mv^{inss,T}(p^0,w^0)$ in (21), a site-specific producer's marginal change in net revenues provide by the information generated from an additional year of experimentation. The marginal value of the information from the fifth year's experiment (that is the value of the information gained from running five years of experiments instead of four) averaged approximately $1.50/ha over the Monte Carlo runs. After the fifth year, the value of the marginal product of experimentation dropped steeply, and after fifteen years was consistently under $0.25/ha. This result occurs because the accuracy of the estimates of the site-specific N fertilization rates is not much improved by an additional year of experimentation once data from about seven or ten years of experimentation is at hand.

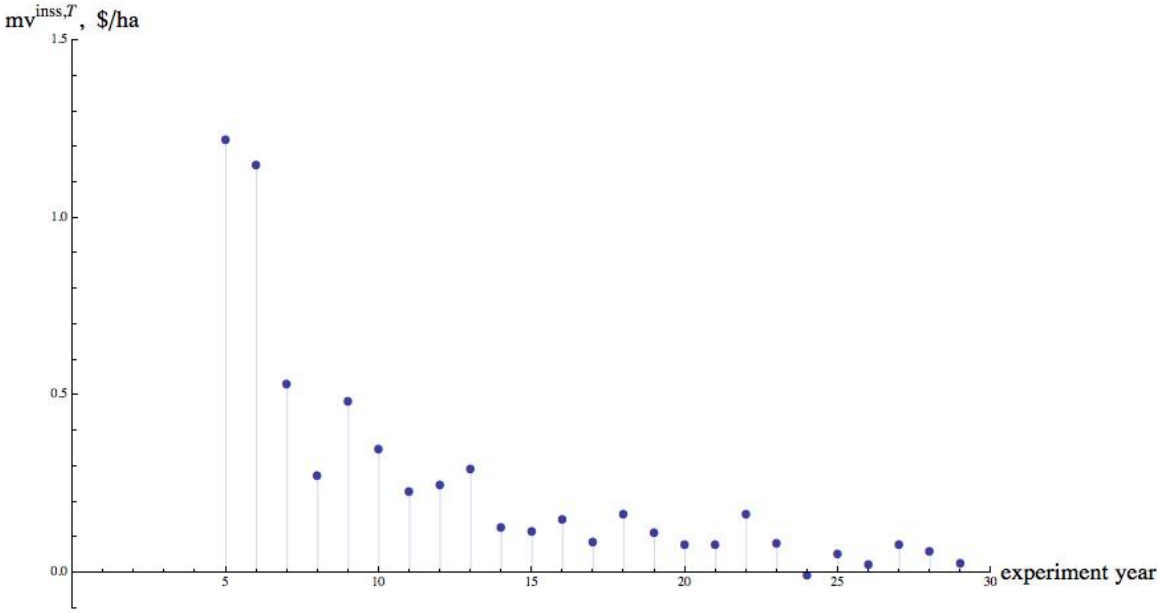

Figure 8. Value to the site-specific farmer of the information from an additional year's experiment

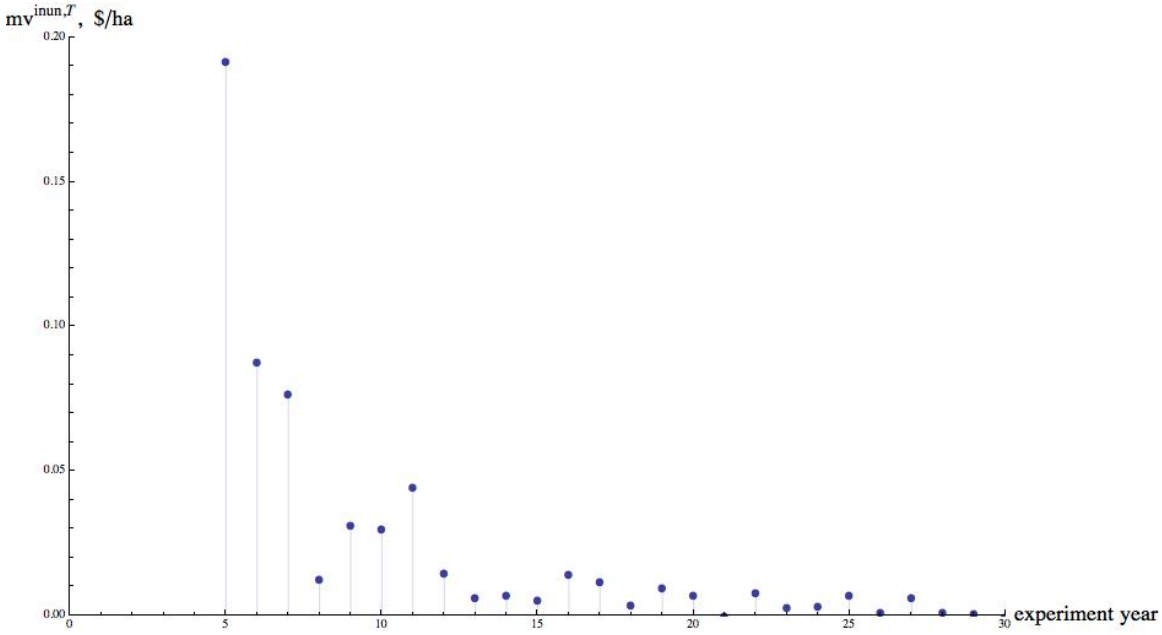

Figure 9. Value to the uniform-rate farmer of the information from an additional year's experiment

The value of the marginal product of experimentation under uniform rate technology, is $mv^{inun,T}(p^0,w^0)$, shown in figure 9. The value of $mv^{inun,T}(p^0,w^0)$, is at all times quite small, well below \$0.20/ha after fifth year. Again, this reflects that additional information is of far less use to producers using uniform-rate technology than to those using site-specific technology.

Finally, figure 10 displays the means of the Monte Carlo runs' ex-ante values of site-specific technology under various amounts of partial information, $v^{ssT}(p^0,w^0)$ from (22). Note that for every amount of experiment length, from four to thirty years, this value is negative. That is, in our model, no matter how many years of agronomic experiments are run, the estimation of the site-specific response functions, $f_b(N, M, \varepsilon)$, and there for the site-specific optimal nitrogen fertilizer application rate, N_b^{ssT}, is always negative. For the farmer with only partial information, net revenues are always higher under uniform technology than under site-specific, even when we do not consider the differences in the costs of producing under these technologies.

The reason for the results above have to do with how we define "partial information," and the advantages of using a particular data set to estimate a whole-field response function instead of many block-specific response functions. In our simulations using the 30-year data set, estimation of the whole-field-response-to-uniform N function $f_{wf}(N, M, \boldsymbol{\varepsilon})$ depended on 4800 observations. But estimations of block-specific response functions $f_b(N, M, \varepsilon)$, were based on only one-thirty second as many observations, that is, 150. As a result, the estimates of the optimal uniform N rate were far more accurate than the estimates of the optimal N rates for the individual blocks. So, here we see a trade-off displayed. By using site-specific technology, the producer has the advantage of

being less constrained in his choices. However, a given number of agronomy experiments (that is, as we have defined it, a given amount of information) can be more useful to the uniform-rate farmer than to the site-specific farmer, because the increased number of observations make for better estimations of the optimal uniform N fertilizer rate than for the optimal site-specific rates. Now, these results will depend on the heterogeneity of the spatial characteristics in the field. The field used in our model was a central Illinois "flat and black" field, and the spatial heterogeneity of the SPI and ISNT values was not great. This leads to similar site-specific inputs rates among sites, and limits the site-specific technology's advantage. Obviously, extending our analysis to examine the values of information and technology types of fields less spatially homogeneous is called for.

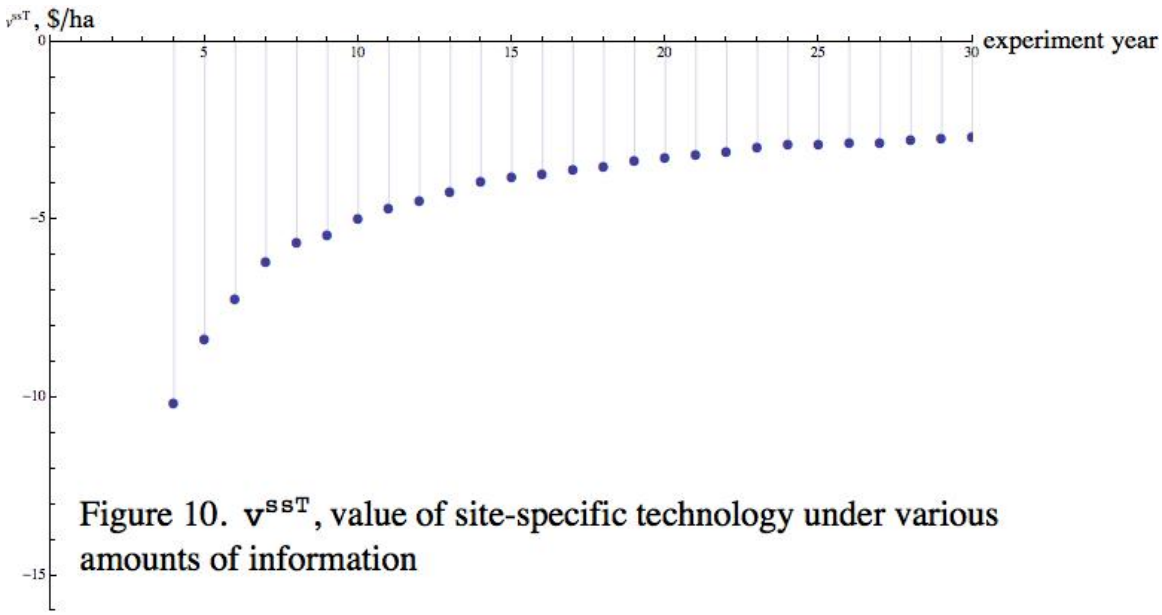

Figure 10. $\mathbf{v^{ssT}}$, value of site-specific technology under various amounts of information

References

Bullock, D.G., D.S. Bullock, E.D. Nafziger, T.A. Doerge, S.R. Paskiewicz, P.R. Carter, and T.A. Peterson. "Does Variable Rate Seeding of Corn Pay?" *Agronomy Journal* 90(1998): 830-836.

Bullock, D.S., and D.G. Bullock. "From Agronomic Research to Farm Management Guidelines: A Primer on the Economics of Information and Precision Agriculture." *Precision Agriculture* 2(2000): 71-101.

Bullock, D.S., J. Lowenberg-deBoer, and S. Swinton. "Adding Value to Spatially Managed Inputs by Understanding Site-specific Yield Response." *Agricultural Economics* 27(2002): 233-245.

Bullock, D.S., and J. Lowenberg-DeBoer. "Using Spatial Analysis to Study the Values of Variable Rate Technology and Information." *Journal of Agricultural Economics* 53(2007): 517-535.

Bullock, D.S., M.L. Ruffo, D.G. Bullock, and G.A. Bollero. "The Value of Variable Rate Technology: An Information-Theoretic Approach." *American Journal of Agricultural Economics* 91(2009): 209-223.

Laffont, J.J. *The Economics of Uncertainty and Information*. Cambridge, MA : The MIT Press, 1989.

LeSage, J., and R.K. Pace. *Introduction to Spatial Econometrics.* Boca Raton, Florida: Chapman & Hall/CRC, 2009.

Rodriguez, D.G.P. The Stanford "1.2 Rule" for Nitrogen Fertilizer: A Long-term Failure of University Extension. Working paper, University of Illinois Department of Agricultural and Consumer Economics, 2013.

Ruffo, M.L., G.A. Bollero, D.S. Bullock, and D.G. Bullock. "Site Specific Production Functions for Variable Rate Corn Nitrogen Fertilization." *Precision Agriculture* 7(2006): 327-342.

Wolfram Research, Inc., *Mathematica*, Version 9.0.1.0, Champaign, IL, 2013.