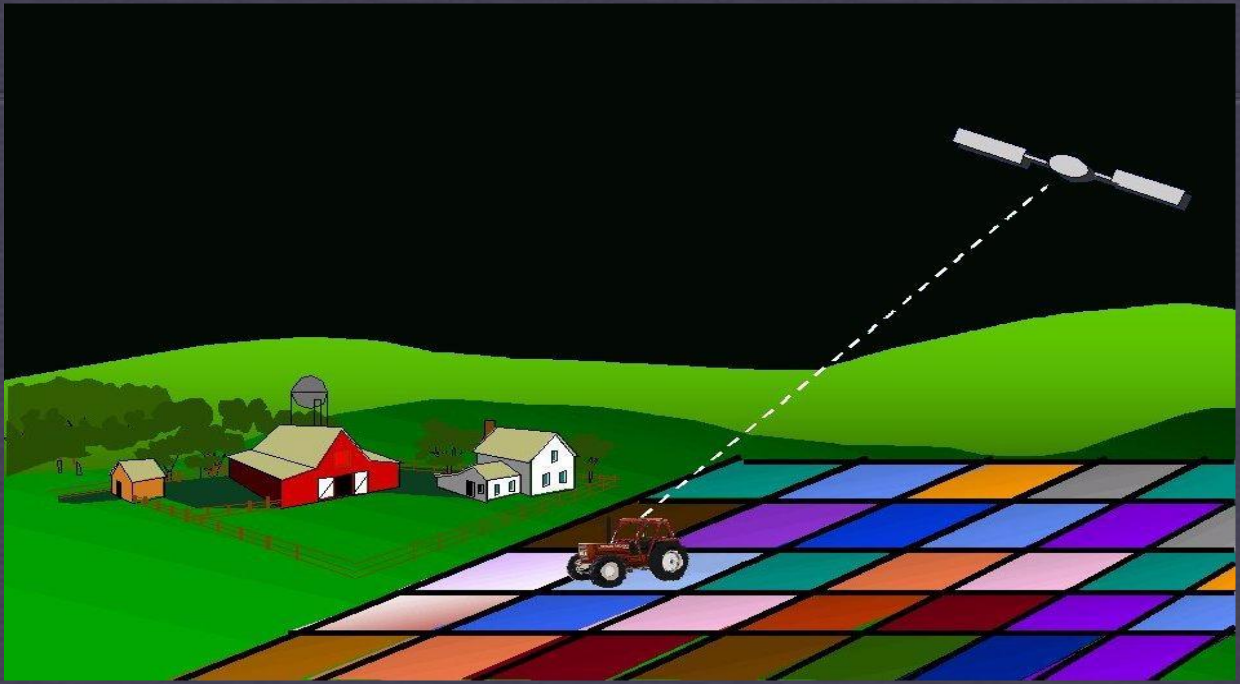
Simulating the Value of Information Generated by On-farm Agronomic Experimentation Using Precision Agriculture Technology

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

Prepared for presentation at the IATRC symposium "Productivity and Its Impacts on Global Trade," Seville, Spain, June 2-4, 2013

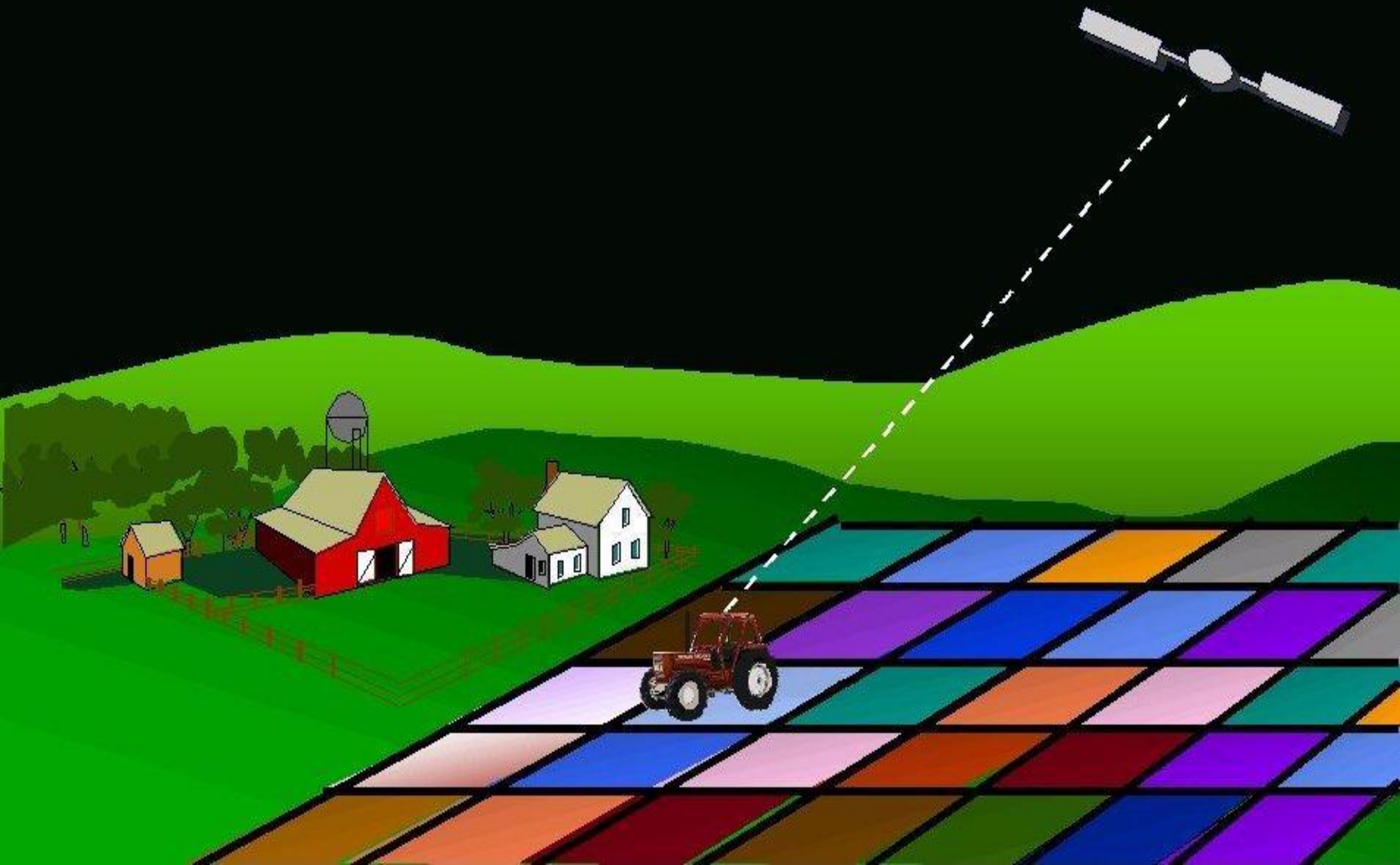
Introduction



In the past:

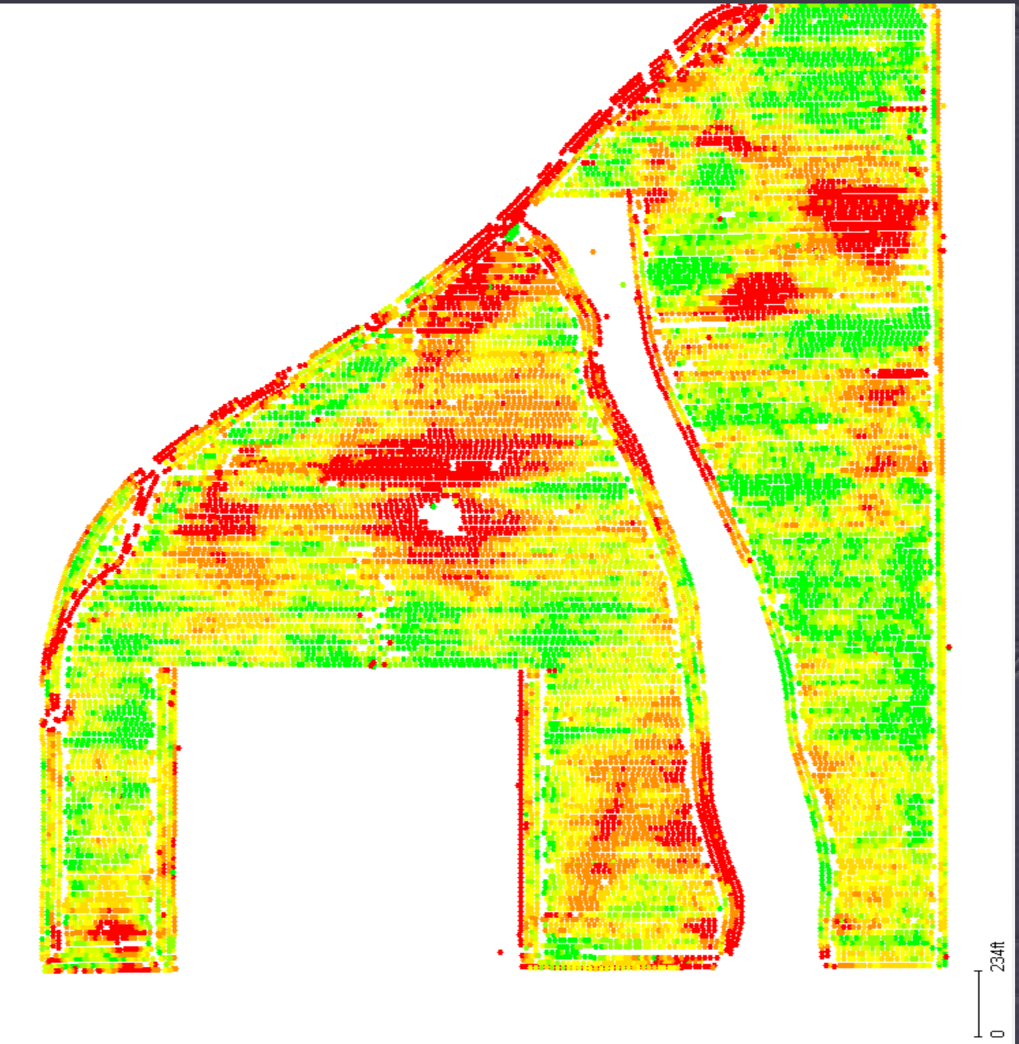
Lots of *excitement* about variable rate technology.

You remember variable rate technology!:



You remember variable rate technology!:

But before long, farmers grain farmers started to say,



"That's a pretty yield map. But how does it make me money?"

And, "that's an exceptional infra-red photo of how green the leaves on my corn plants are. But how does it make me money?"

Fifteen years later, at least for major crops, there's been very little adoption of variable rate technology.

Nobody surfed in on the “wave of the future.” Why not?

Information. For many crops and applications, precision technology has not been profitable because of lack of *information*.

Particularly, information about how yields respond to inputs (fertilizer, seed rate, …).

My paper is about the important interplay between precision agriculture technology and *information*.

Possible good news: VRT can cheaply provide the information needed by VRT to be profitable!

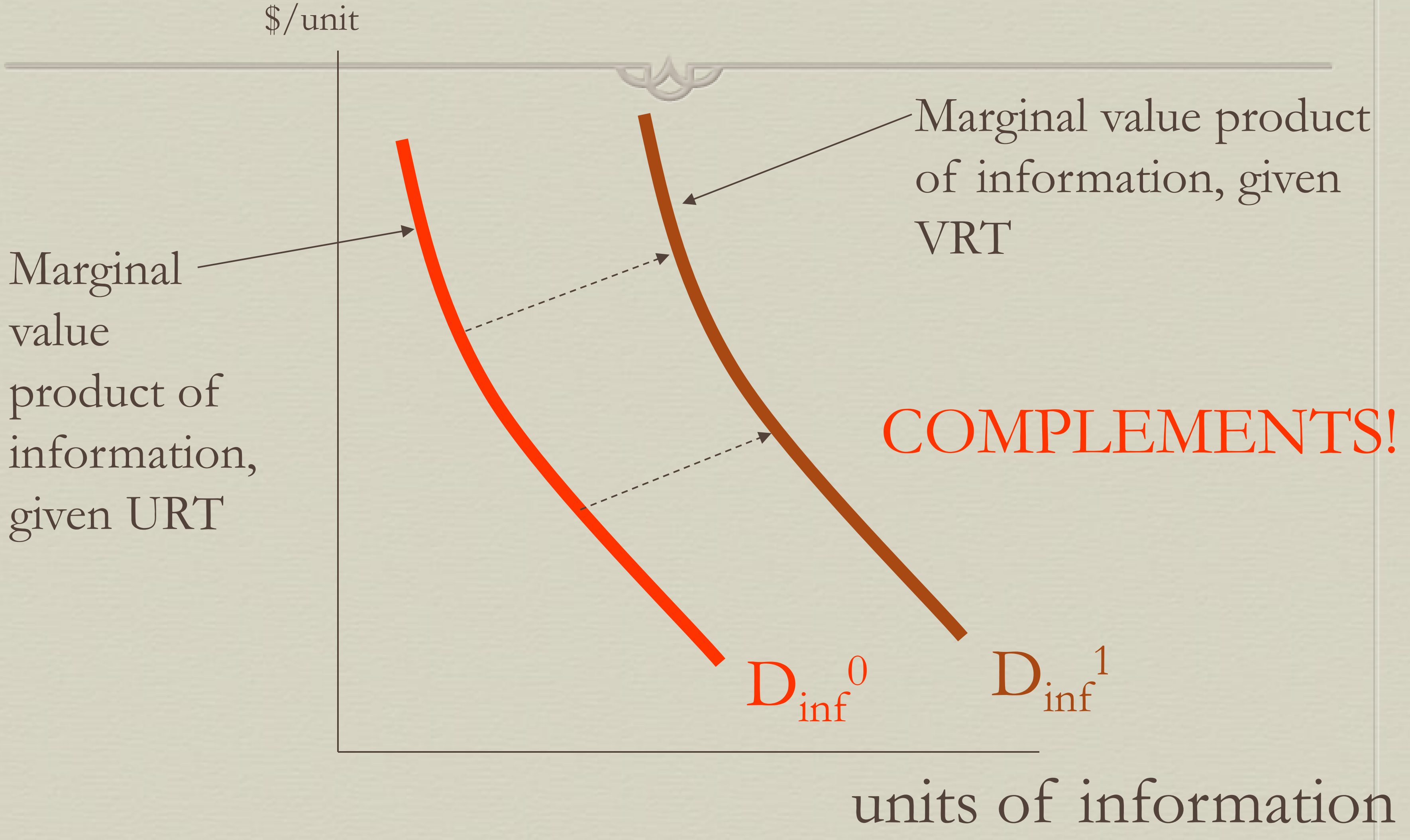
II. A Theory of The (non) Adoption of Precision Ag Technology

VRT and information are *complements.*

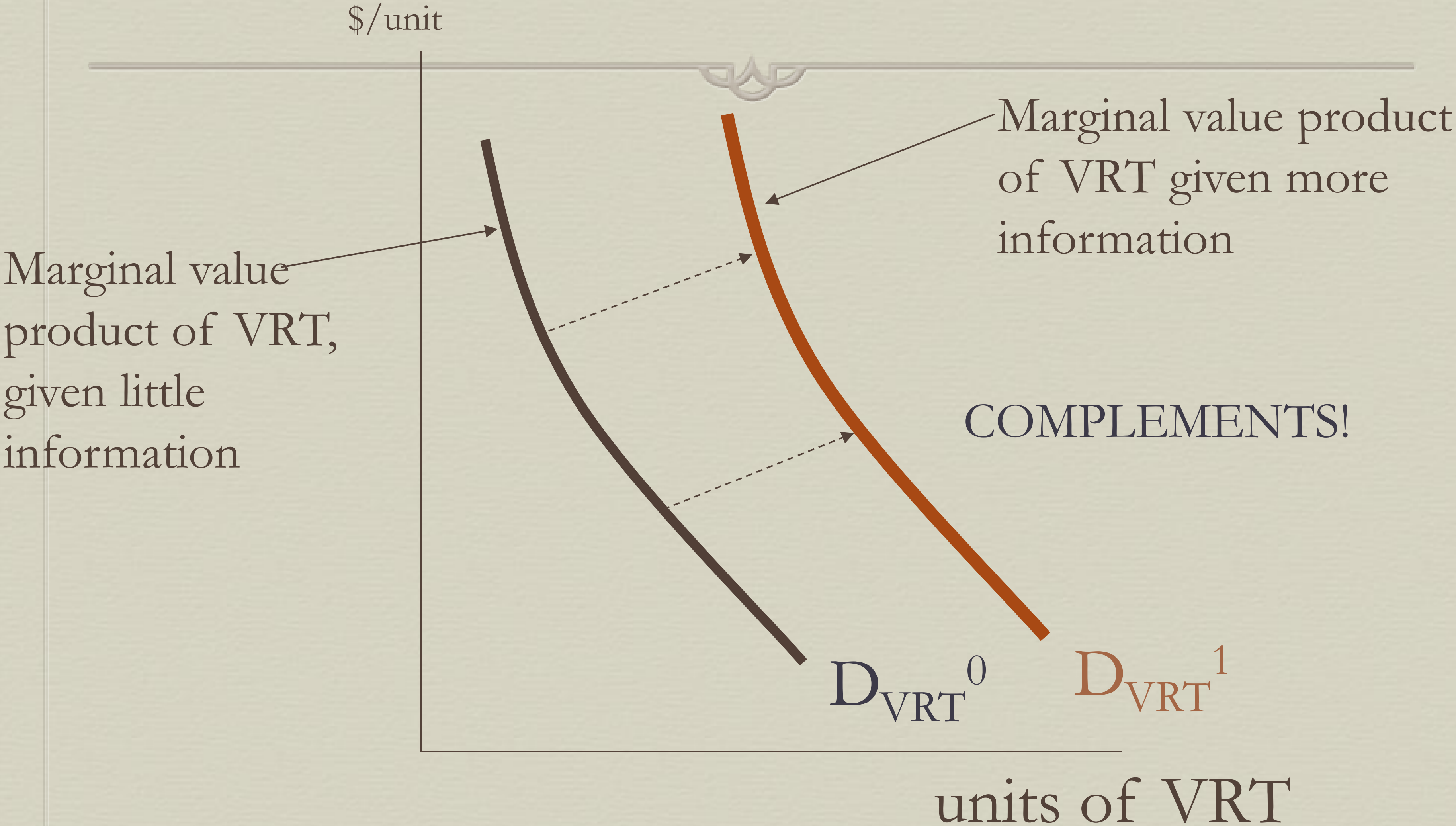
Intuition:

To farm site-specifically, you need site-specific information.

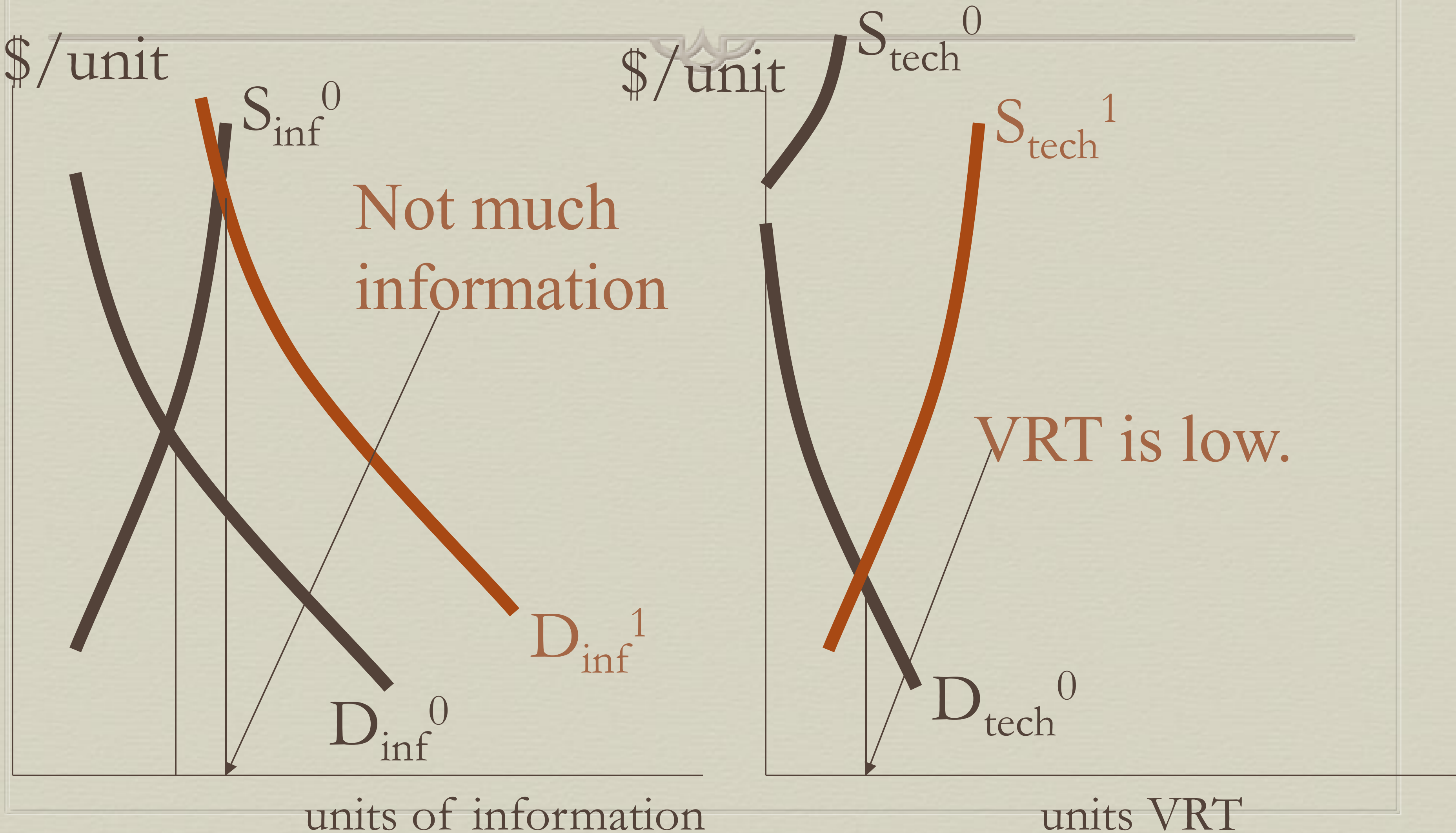
Demand for information



Demand for VRT



Result of the invention of VRT when there is a lack of information:



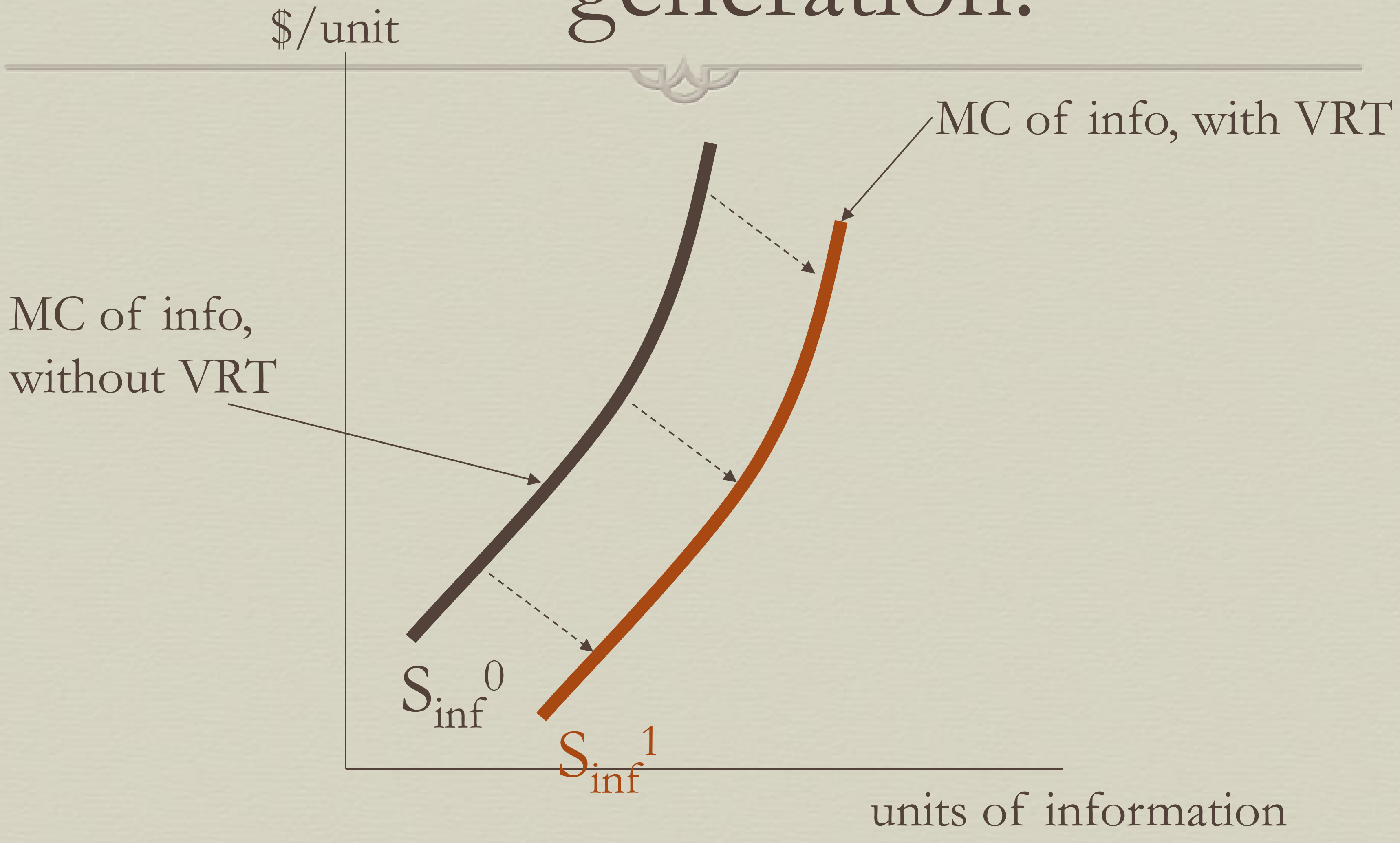
Difficult truth:

Nobody knows much about how yields respond to inputs. So there is little VRT. Equilibrium.

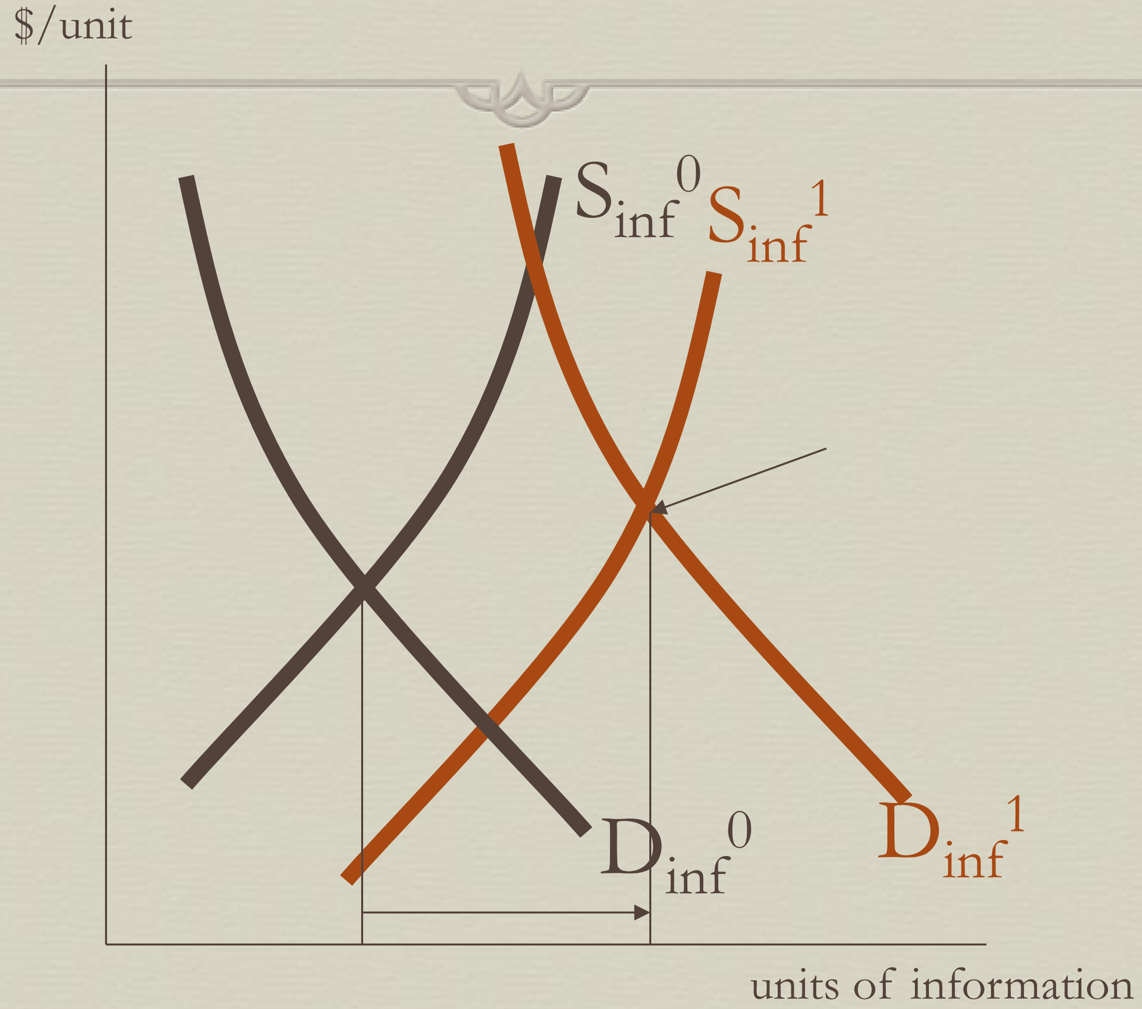
Fortunately….

With precision technology, we can inexpensively gather the kinds of information we need to learn how yields respond to inputs.

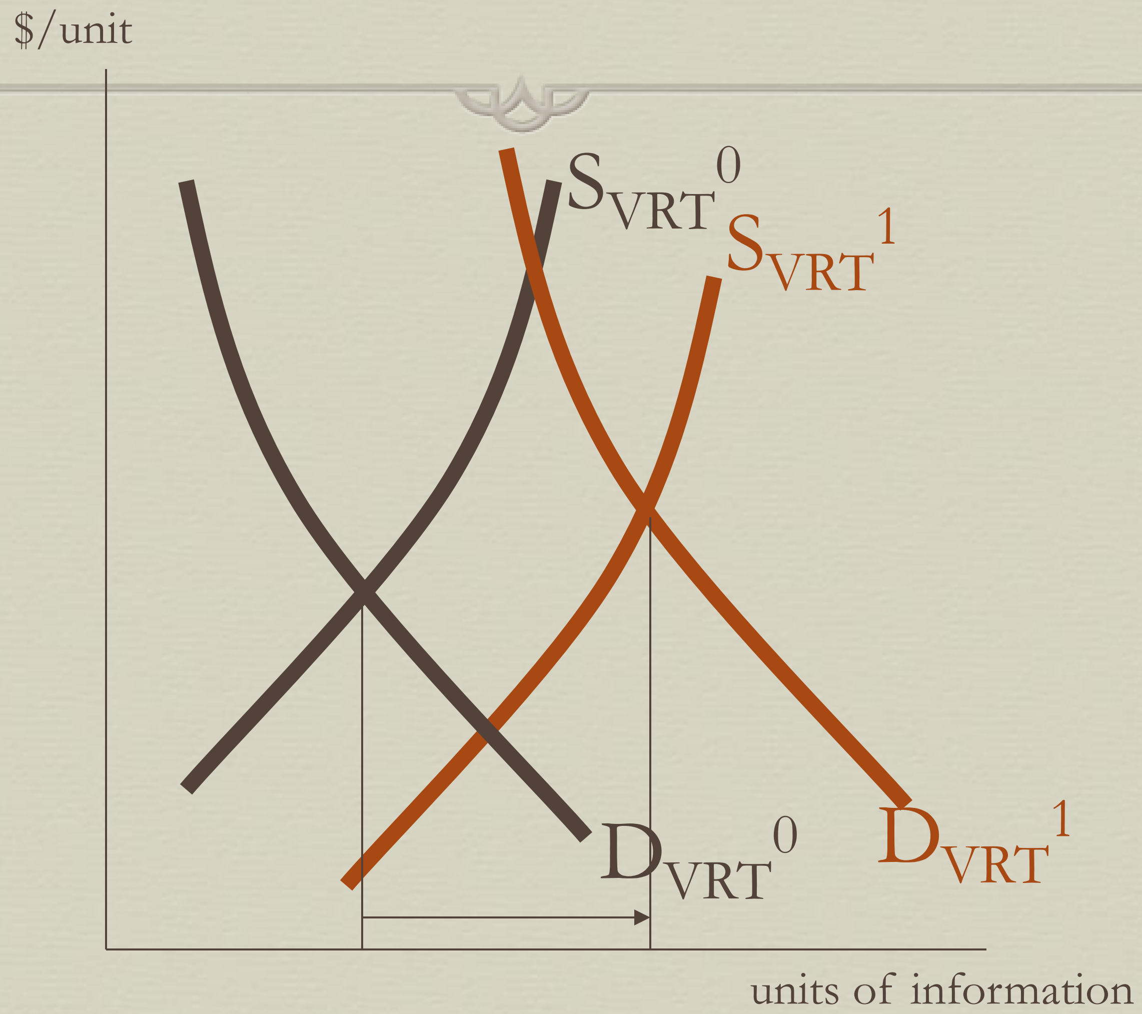
VRT can lower costs of info generation:



If S and D of info both shift, get lots of info:



And get lots of VRT:

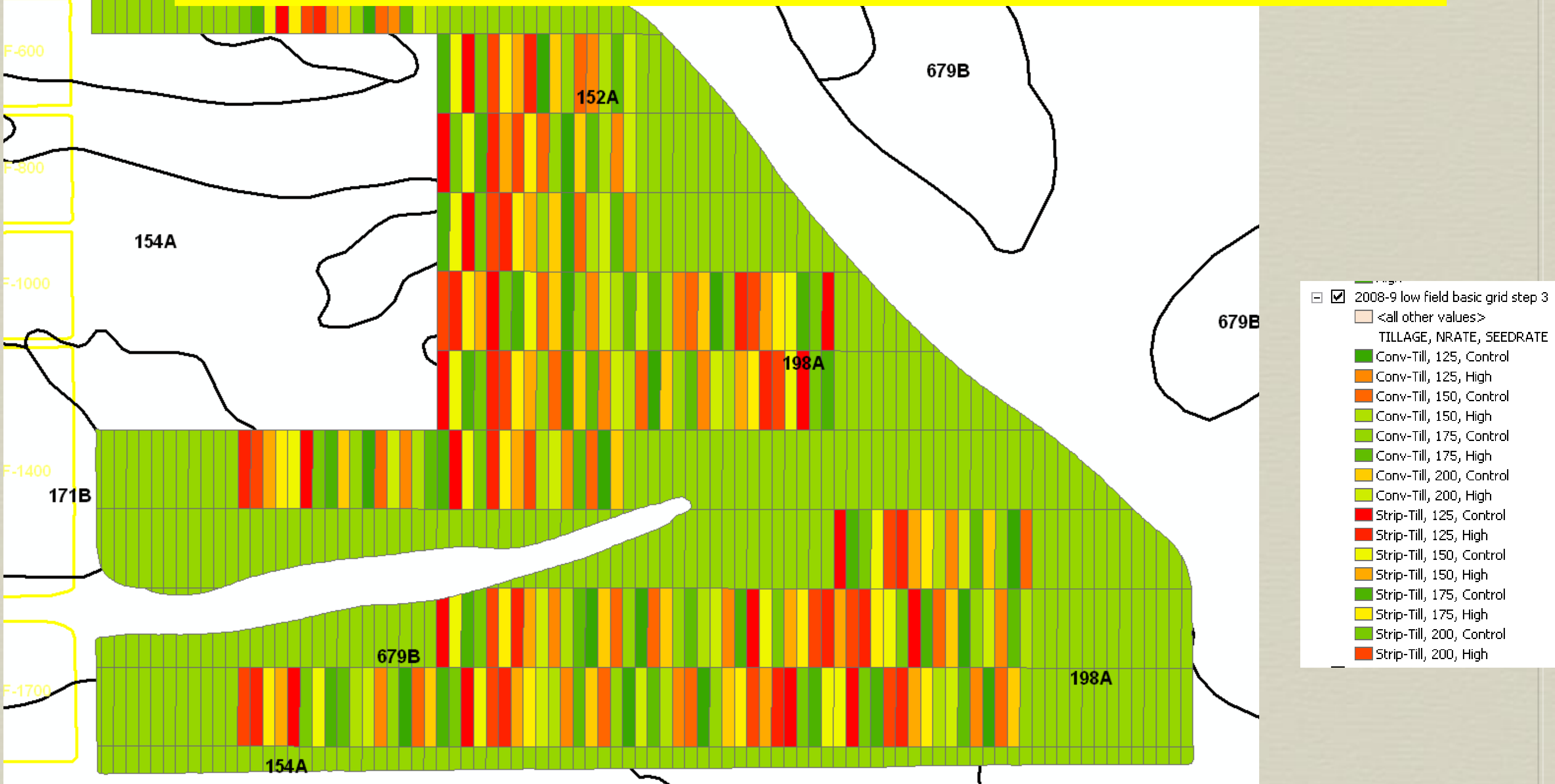


ON-FARM EXPERIMENTATION AND DATA ANALYSIS

These days: we have computer software that makes VRT equipment implement agronomic experiments *cheap*:

Actual Design for Lo Farm: Tillage, N Rate & Seeding Rate Map

COOL! Controlled experiments in economics!



- Farmer implements experiment, listens to Van Halen.

- Farmer harvests crop as usual, and yield monitor collects yield data:

Hasn't this been done before? Didn't Earl Heady do this in the 50s? Where did the data for Quirino Paris's papers come from?

Experiments only run on a much smaller scale, with grad students, orange flags, and measuring tapes. In the rain.

(Typical highly-paid grad student)

These days:

It's cheap to get lots and lots of data from lots and lots of places under lots and lots of weather, soil, topography situations!

We have begun such a research project, with a real farmer, on a real 40-acre Illinois farm field, varying N fertilizer rate and seed rate using a random block design.

We plan to do on-farm agronomic experiments on many different farms over many different years.

Ultimate goal: We want to be the "Moneyball"* guys of farm management!!!!

*American baseball book and movie (starring Brad Pitt!!) in which statisticians are the true heroes.

What kind of money are we talking about?

Pennies per hectare? Tens of dollars per hectare?

What will be the opportunity costs of the experiments?

How valuable will the information we get be, and how long will it take to get it?

To find out,

Monte Carlo simulations of agronomic experiments,

Methodology

Assumed the “true” response function was one we estimated from data on an Illinois cornfield (Bullock, et al., AJAE 2009):

Nitrogen fertilizer

Illinois Soil Nitrogen Test

May Precipitation

Stream Power Index

$$f\left(N, M, S, I, e\right) \equiv b_0 + b_N N + b_I I + b_M M + b_S S$$

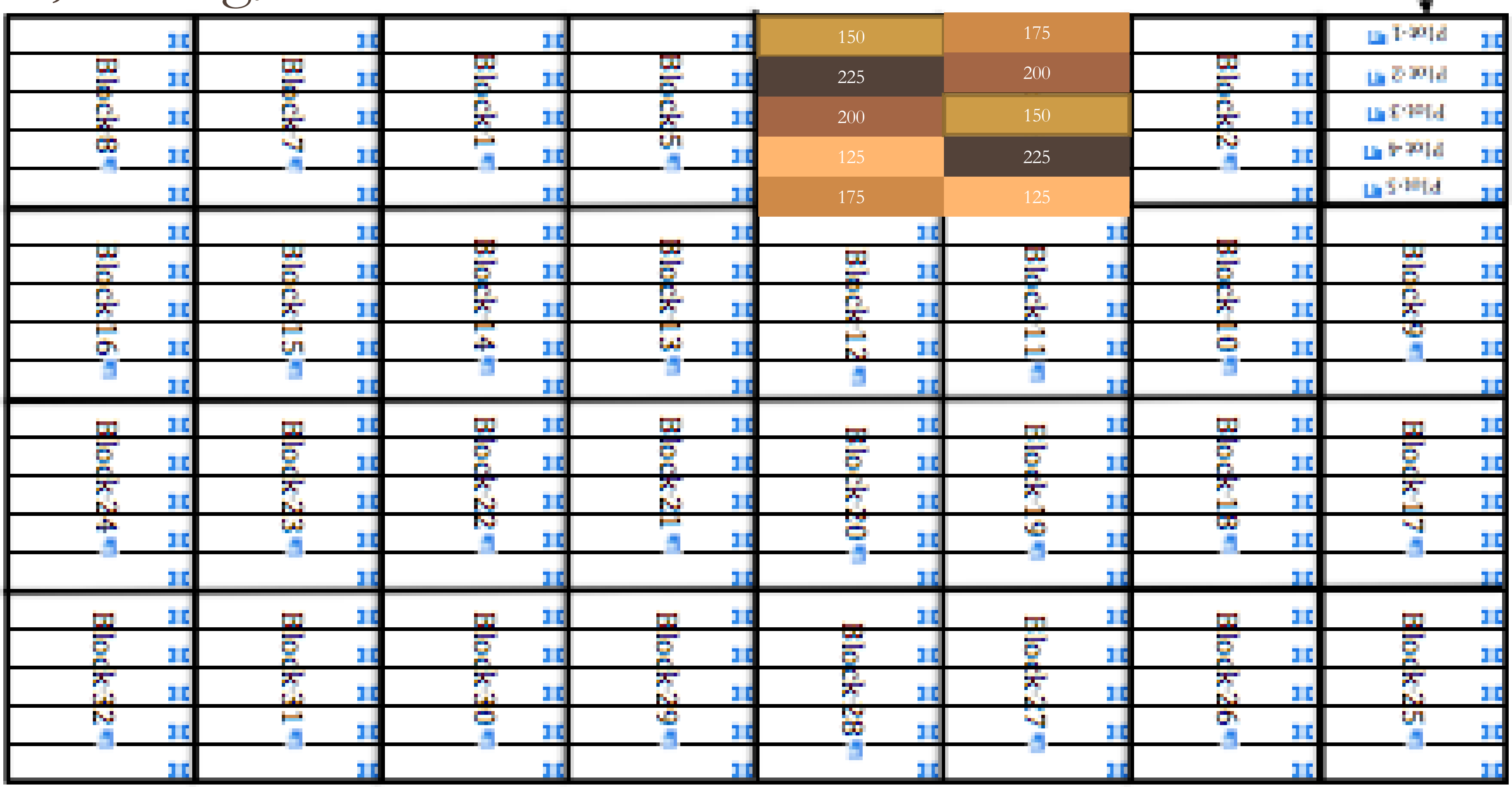
$$+ b_{NN} N^2 + b_{NI} NI + b_{NM} NM + b_{NS} NS + b_{NNI} N^2 I + e$$

Simulated a corn field, agronomic experiments for 30 years:

Randomized block design

32 blocks, 5 plots per block

In each year, each plot in a block randomly assigned one of five levels of N fertilizer rate: N_{bjt} =125, 150, 175, 200, 225 kg/ha:



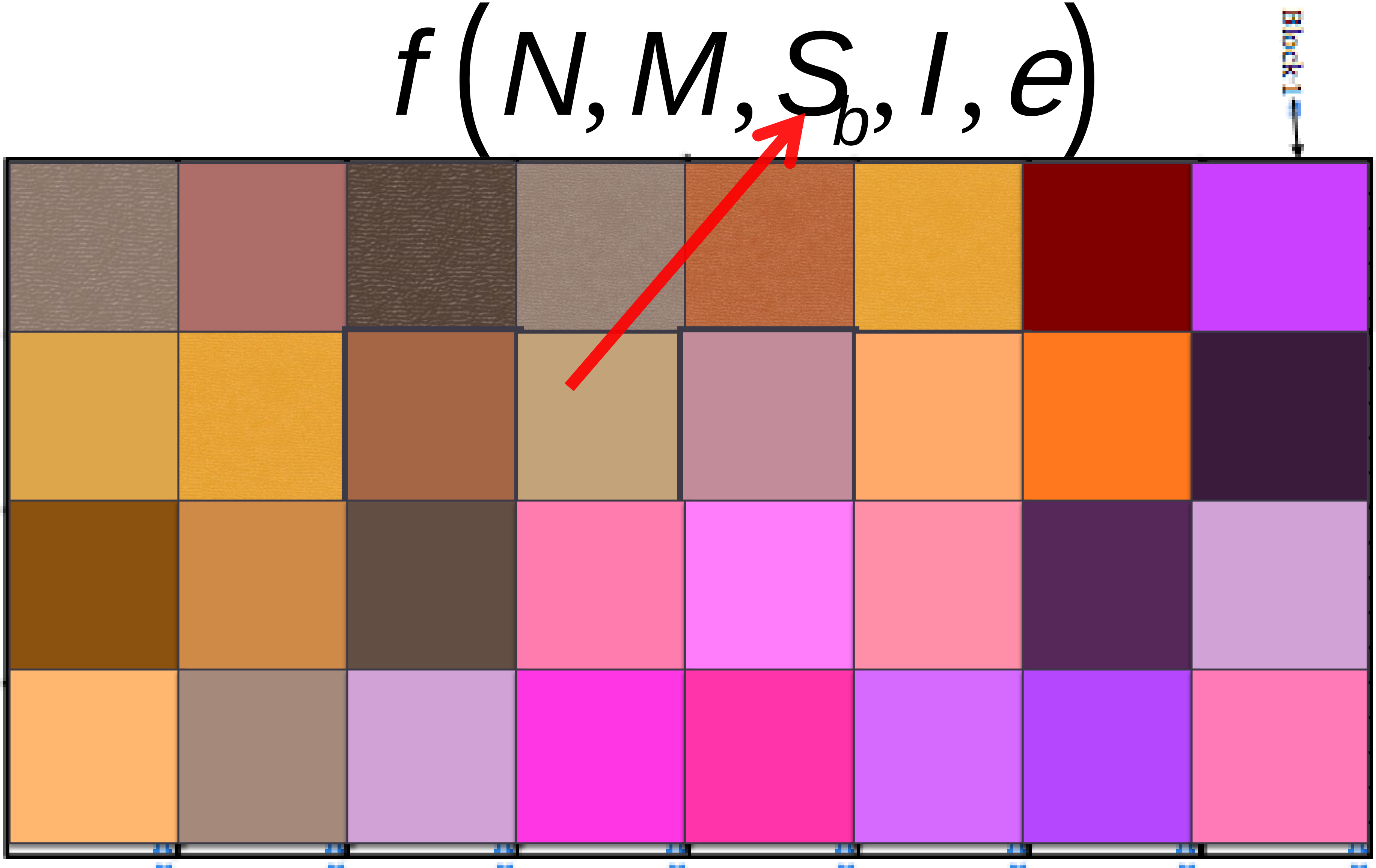
Characterized the field

Each block is given a value of characteristics I and S.

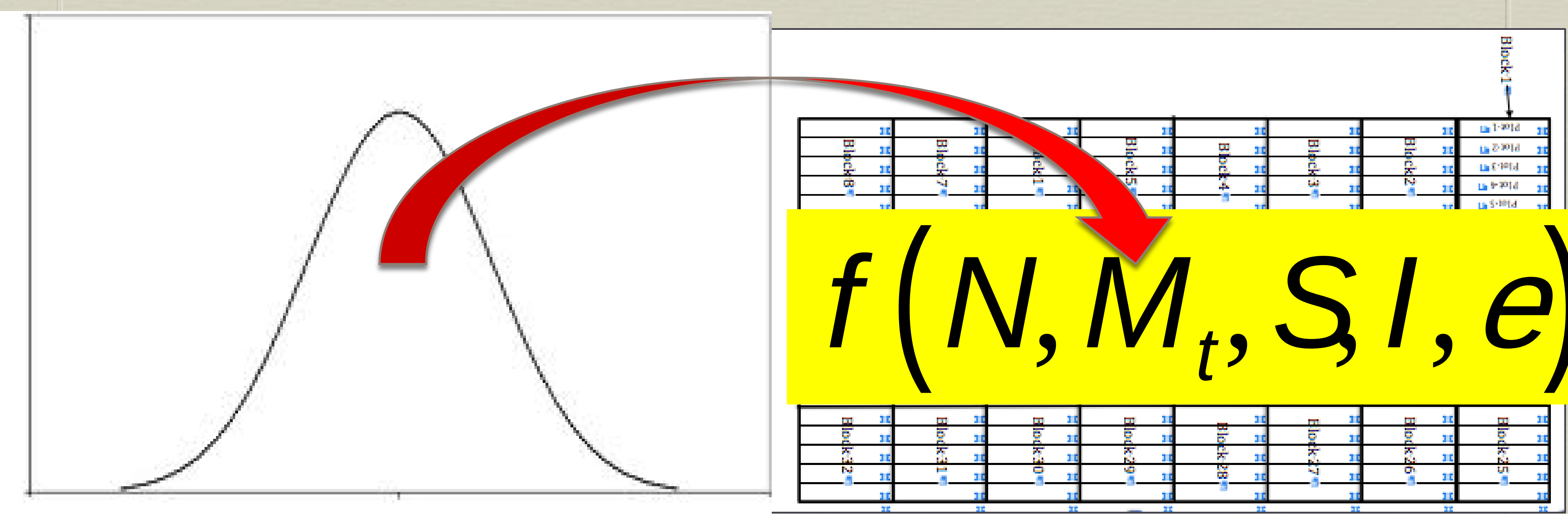
This gave us an "*T*" map and an "*S*" map for the field, each spatially autocorrelated:

Then, each block's assigned *S* and *I* levels is put into the response function for that block:

$$f\left(N, M, S_b, I, e\right)$$

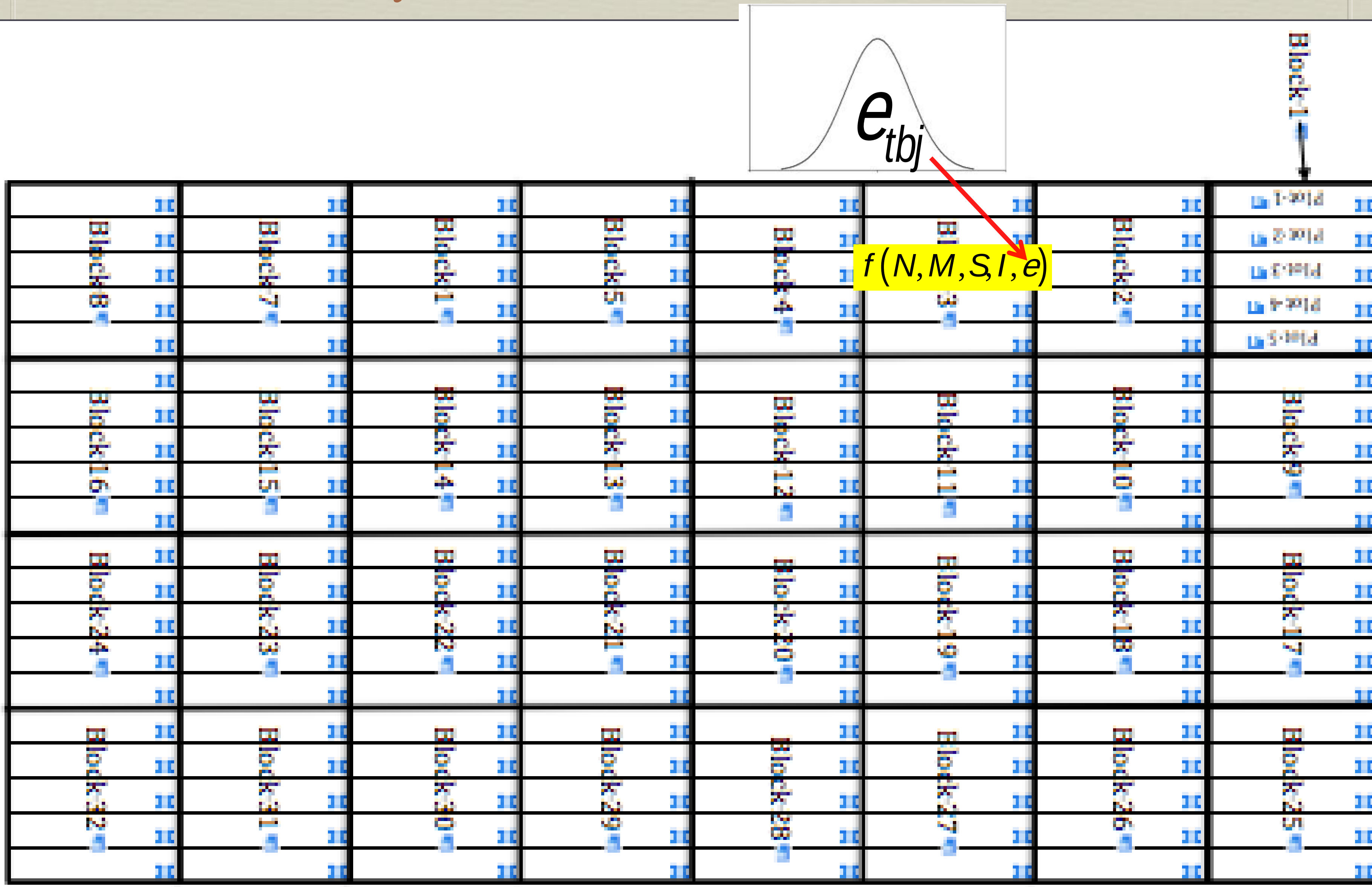


In each year, a random draw for May precipitation for the whole field:



May precipitation, M_t

For each year, each of the 160 plots gets its own random yield disturbance term;



Simulated experimental yield in every plot in every year:

$$q_{tbj} = f\left(N_{tbj}, M_t, S_b, I_b, e_{tbj}\right)$$

This gives us a 30-year data set:

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}	$q_{30.32.5}$

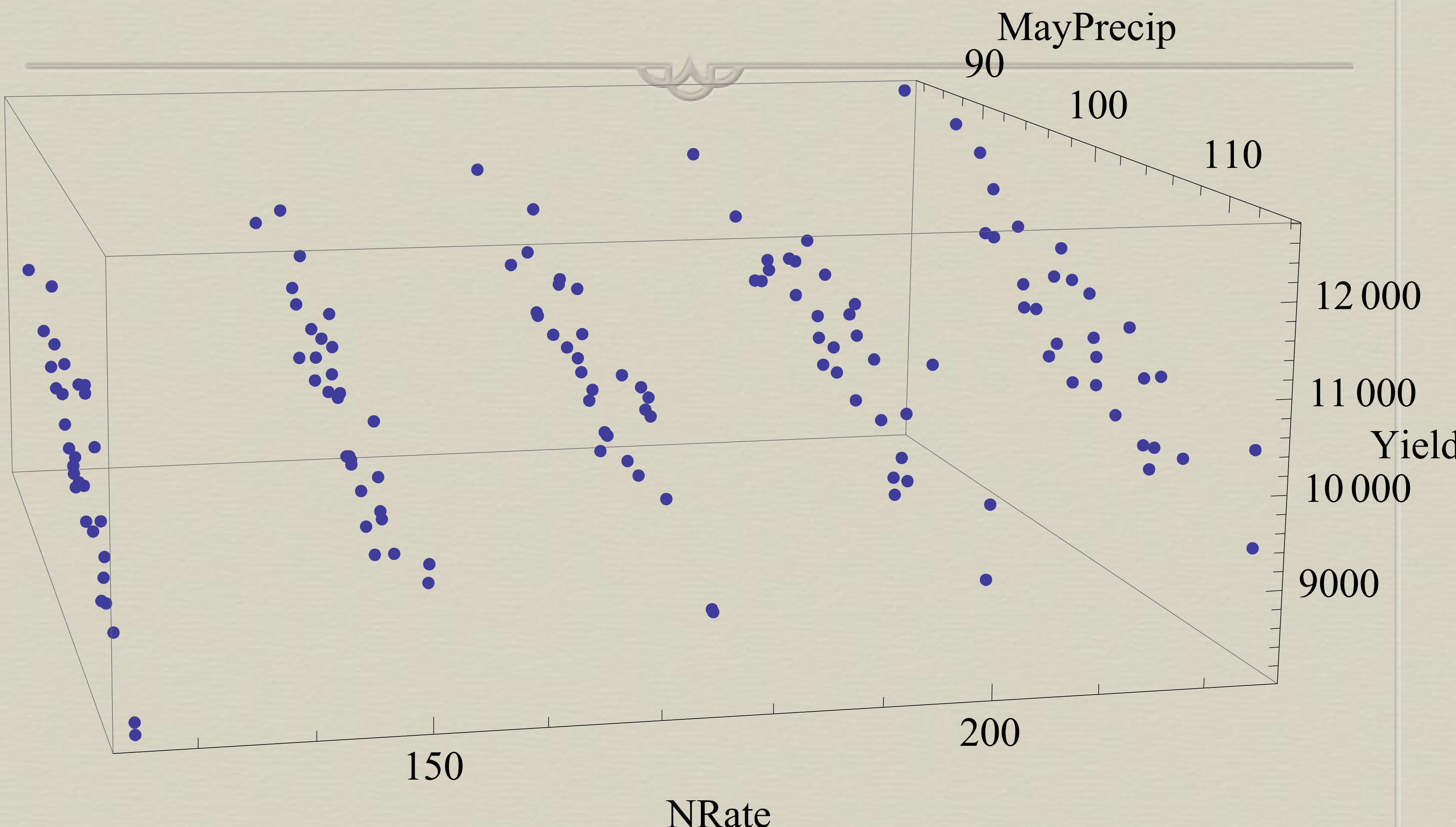
Pretend like economists have the data on yield and N, not S and I, and don't know true response function.

With 30 years of data for a block, here's our data set:

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}	$q_{30,7,5}$

A scatter plot of a block's 150 observations:



Use data to estimate block-specific response function:

$$\hat{f}_b^{30}(N,M) = \hat{b}_{b0}^{30} + \hat{b}_{bN}^{30}N + \hat{b}_{M}^{30}M + \hat{b}_{bMN}^{30}MN + \hat{b}_{bNN}^{30}NN$$

Estimated expected-profit-maximizing block-specific N rate from 30 years of data:

$$\hat{N}_b^{30*} = \frac{\frac{w}{p} - \hat{b}_{bN}^{30} - \hat{b}_{bMN}^{30} E\{M\}}{2\hat{b}_{bNN}^{30}}$$

Site-specific profits:

$$p_b^{30*} = pf\left(N_b^{30*}, M_t, S_b, I_b, 0\right) - wN_b^{30*}$$

Expected profits on the whole field with *t* years of data when using site-specific technology:

$$\Pi_{ss}^{t*} = \frac{1}{32}\sum_{b=1}^{32}\left(pf\left(N_b^{t*}, M_t, S_b, I_c, 0\right) - wN_b^{t*}\right)$$

URT: would use all the blocks' data together.

30 years data gives us 4800 observations.

True whole-field response function under uniform-rate management:

$$f_{wf}\left(N, M_t, \boldsymbol{\varepsilon}\right) \equiv \sum_{b=1}^{B} \sum_{j=1}^{5} f_b\left(N, M_t, \varepsilon_{bjt}\right)$$

Estimate the whole-field response function using all the 30 years of data (4800 observations):

$$\tilde{f}_{wf}^{30}\left(N, M\right) = \tilde{b}_{0}^{30} + \tilde{b}_{N}^{30} N + \tilde{b}_{M}^{30} M + \tilde{b}_{MN}^{30} MN + \tilde{b}_{NN}^{30} NN$$

Estimated expected-profit-maximizing uniform N rate from 30 years of data:

$$N_{wf}^{30*} = \frac{\frac{w}{p} - \tilde{b}_N^{30} - \tilde{b}_{MN}^{30} E\{M\}}{2\tilde{b}_{NN}^{30}}$$

Resultant ex-ante expected profits:

$$\mathrm{P}_{un}^{30*}\left(p,w\right) \equiv$$

$$E\left\{ p \sum_{b=1}^{B} \sum_{j=1}^{5} \left[f_b\left(N_{wf}^{30*}\left(p,w\right), M_t, e_{bjt}\right) - wN_{wf}^{30*}\left(p,w\right) \right] \right\}$$

Results of 100 Monte Carlo Runs:

Marginal value of a year's experiment very small for uniform management.

Can get most of what you need with a few years of experiments.

$$MVI_{un}^{t} = P_{un}^{t*} - P_{un}^{t-1*}$$

Less than $0.20/ha.

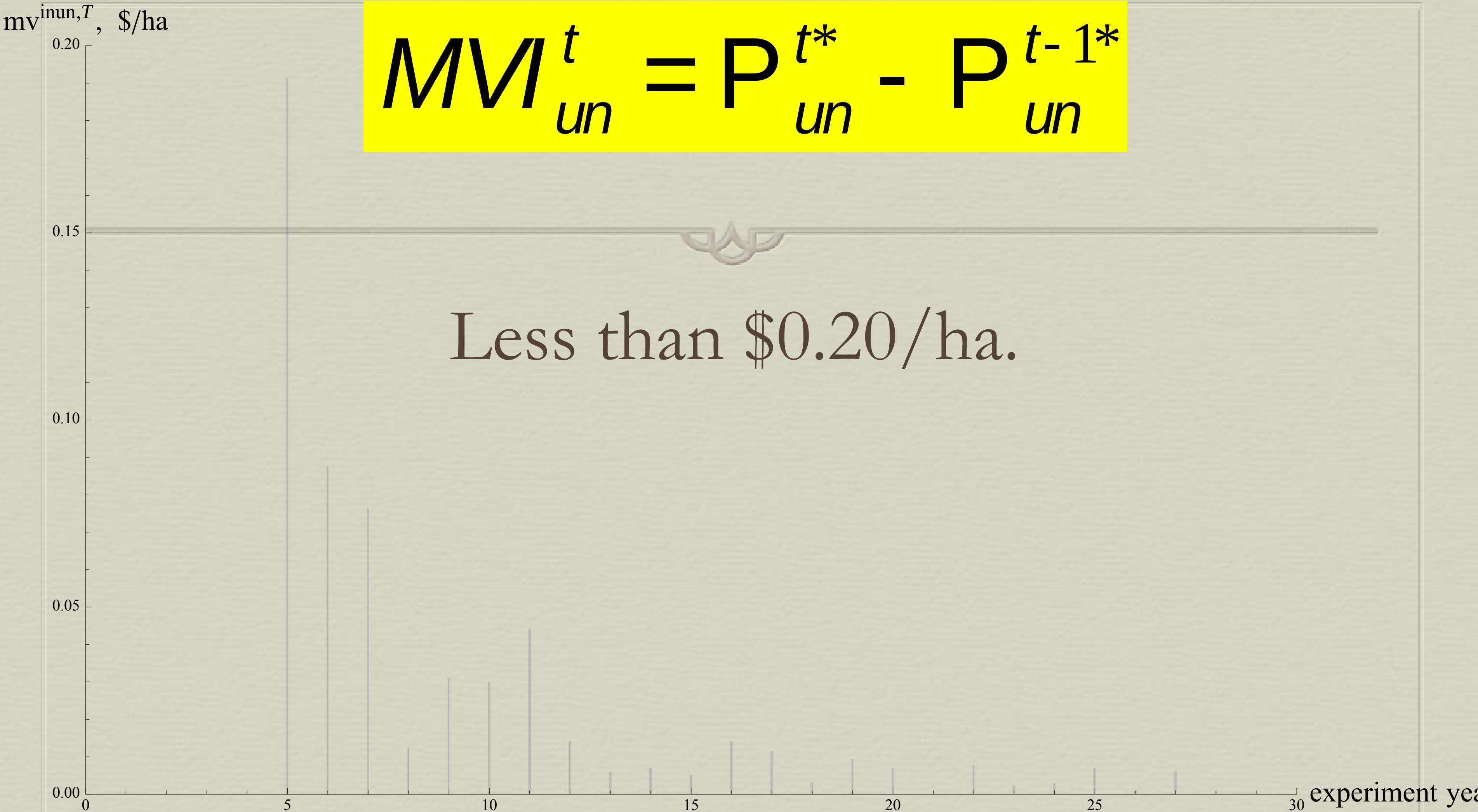


Figure 9. Value to the uniform-rate farmer of the information from an additional year's experiment

Marginal value of a year's experiment bigger for s-s management.

$$MVI_{ss}^{t} = P_{ss}^{t*} - P_{ss}^{t-1*}$$

Marginal value of fifth experimental year: about $1.20/ha.

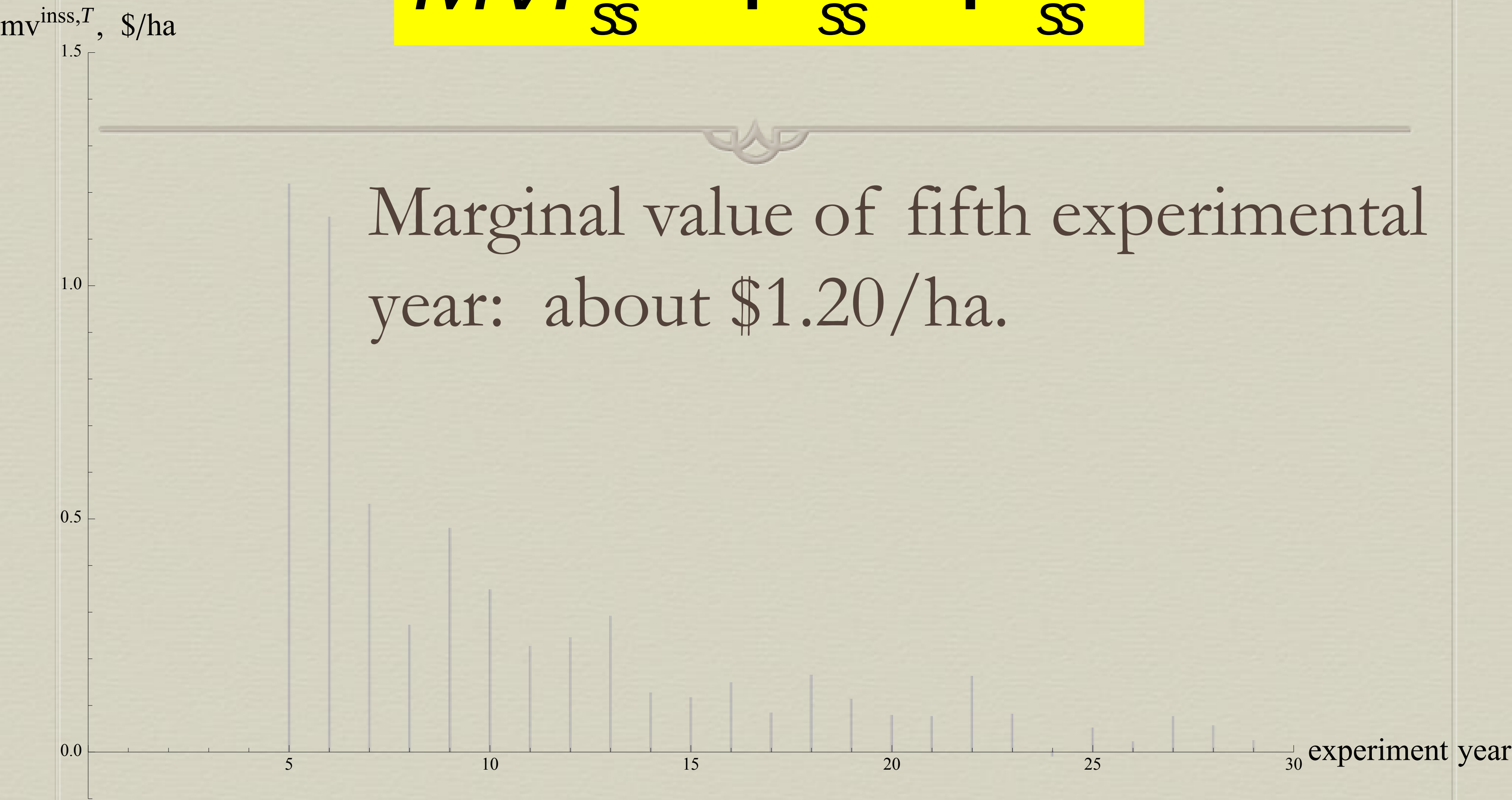
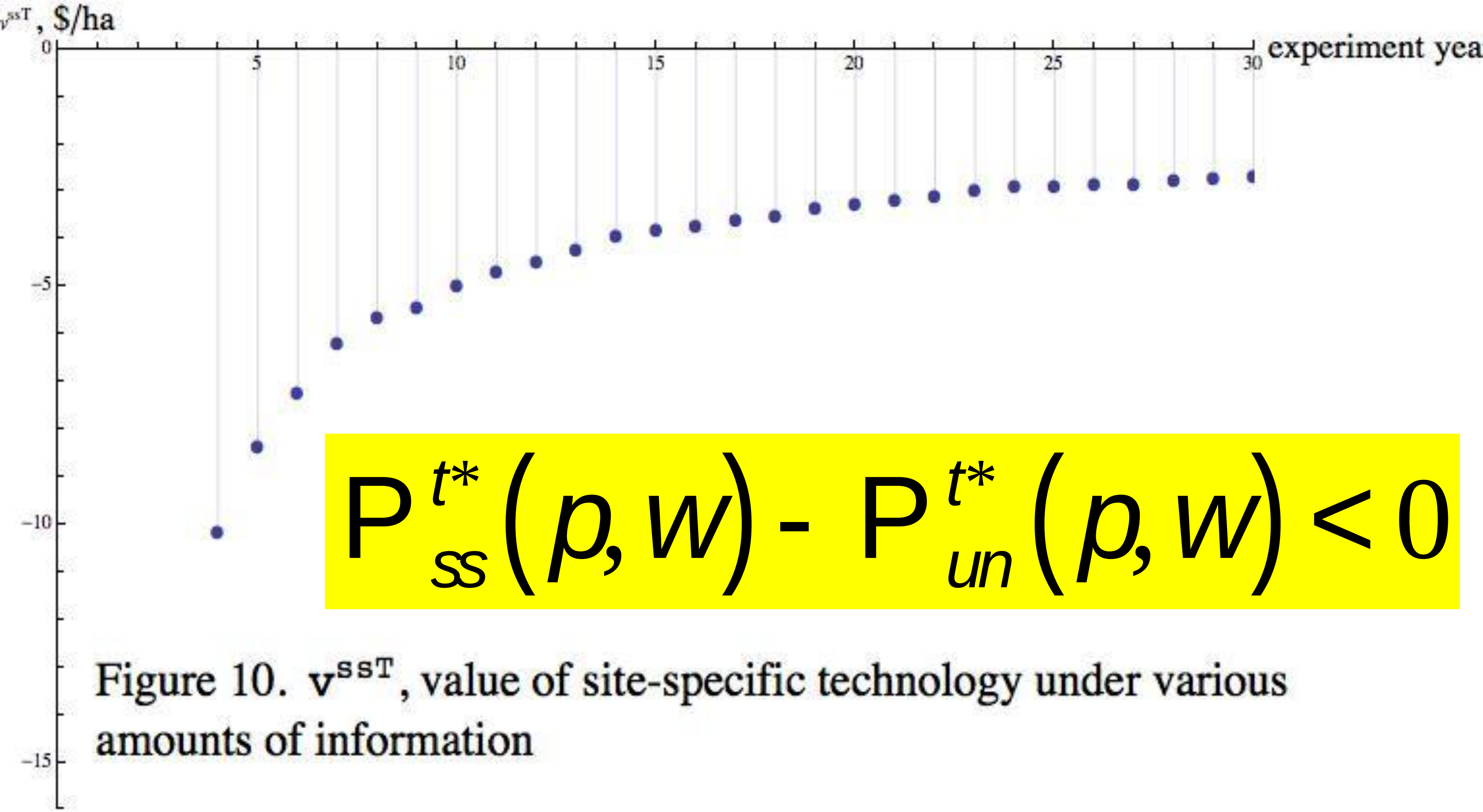


Figure 8. Value to the site-specific farmer of the information from an additional year's experiment

By using VRT, a producer **who knew every block's true response function**, $f_b(N, M, e)$ could expect net revenues **$2.19/ha** greater than a uniform-rate producer who knew every block's true response function.

But without full info, site-specific management is a <u>loser</u>.



$$P_{ss}^{t*}(p,w) - P_{un}^{t*}(p,w) < 0$$

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

In our simulations, it just isn’t possible to get enough information from the experiments for precision agriculture to pay for itself.

Caveat: Only used OLS.

What happens when we do the econometrics the right way, with spatial econometrics?

Note:

This was a flat, black Illinois cornfield. Very homogenous spatial characteristics. VRT worth more on more spatially varied field.