Space and Agriculture

NSF-AERC-IGC Technical Session on Agriculture and Development

3 December 2010

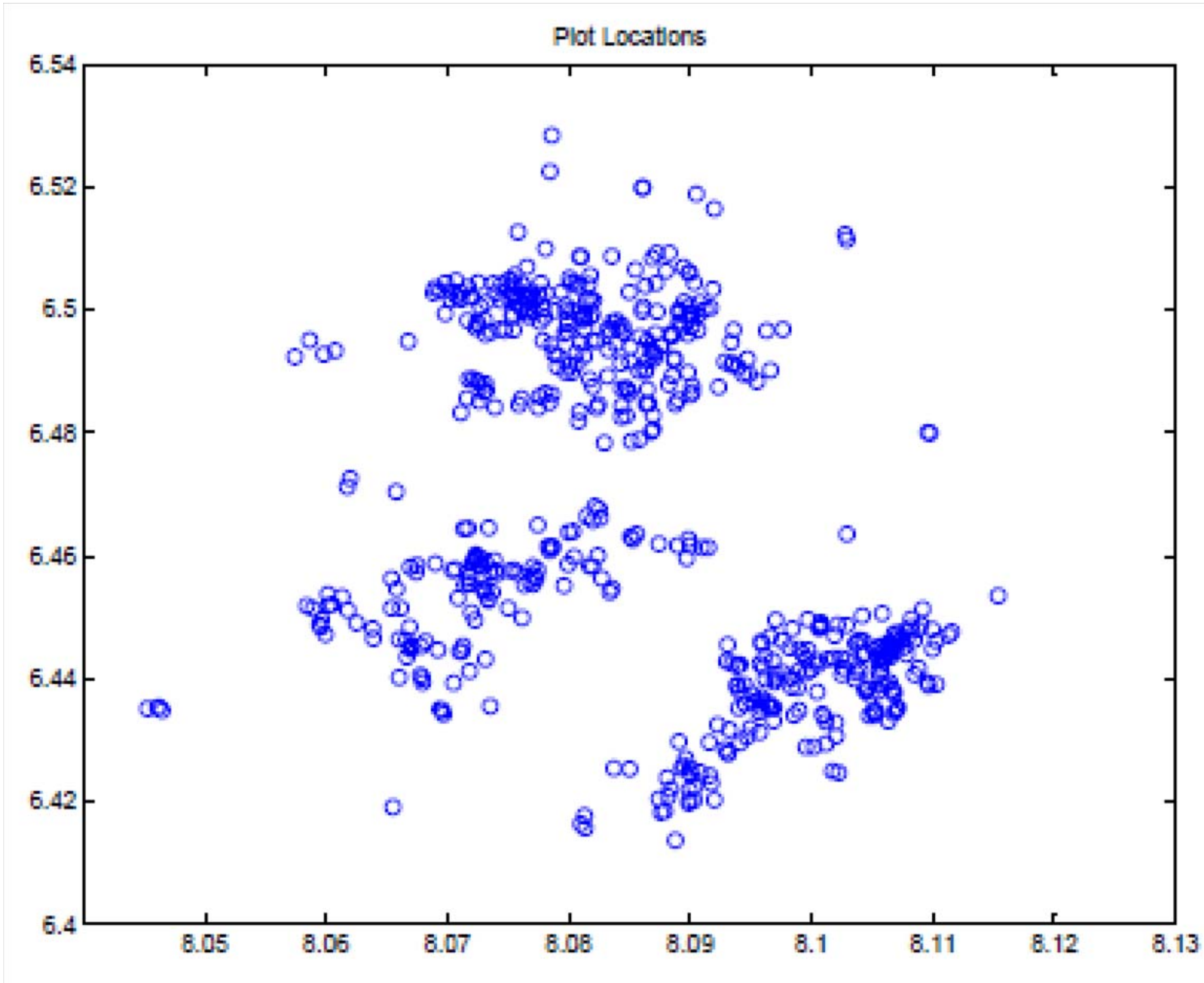
Chris Udry, Yale

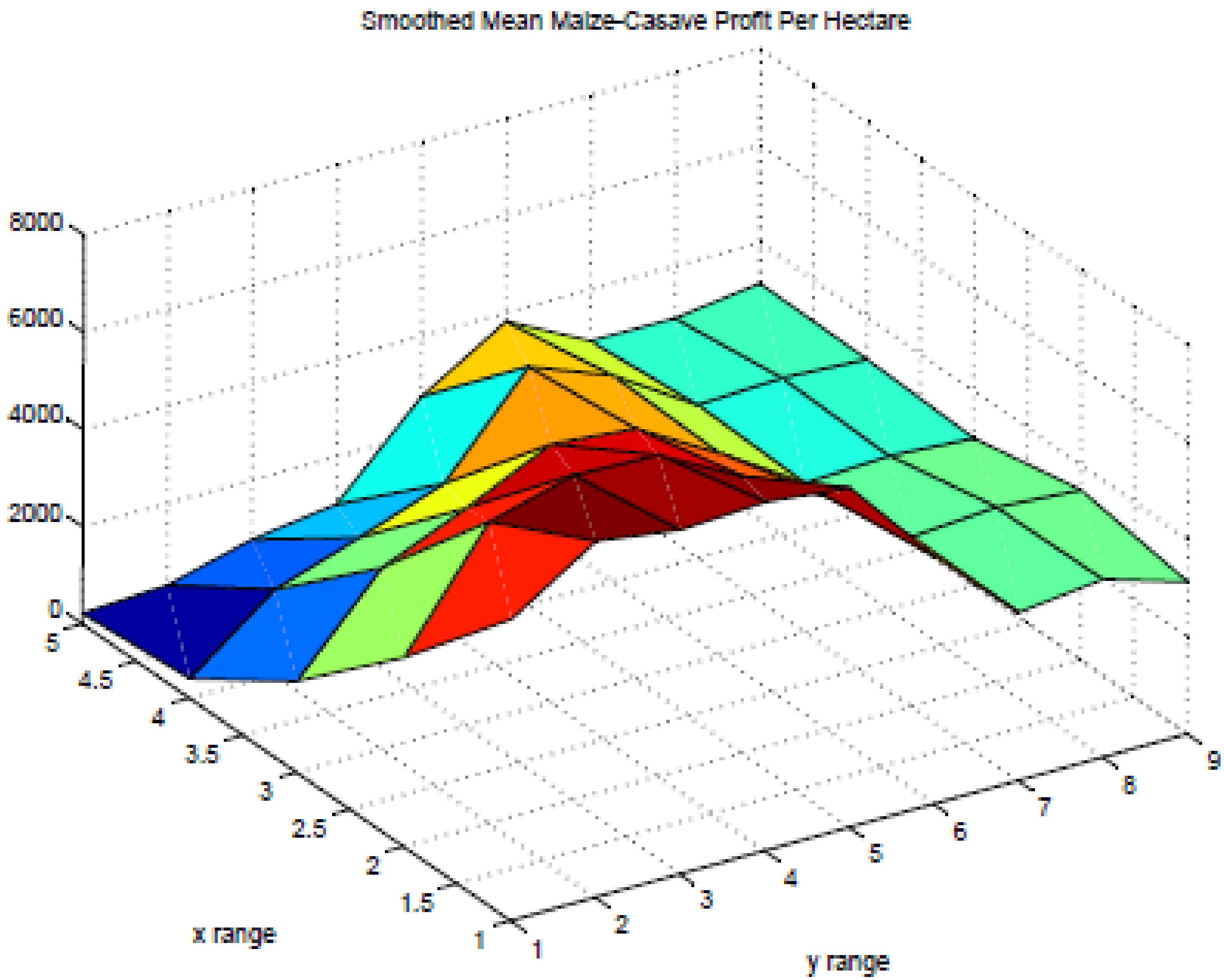
notes available online at www.econ.yale.edu/~cru2/papers.html

- Evenson and Westphal, Handbook of Development Economics, "Strong interaction between the environment and biological material makes the productivity of agricultural techniques ... highly dependent on local soil, climatic, and ecological characteristics."

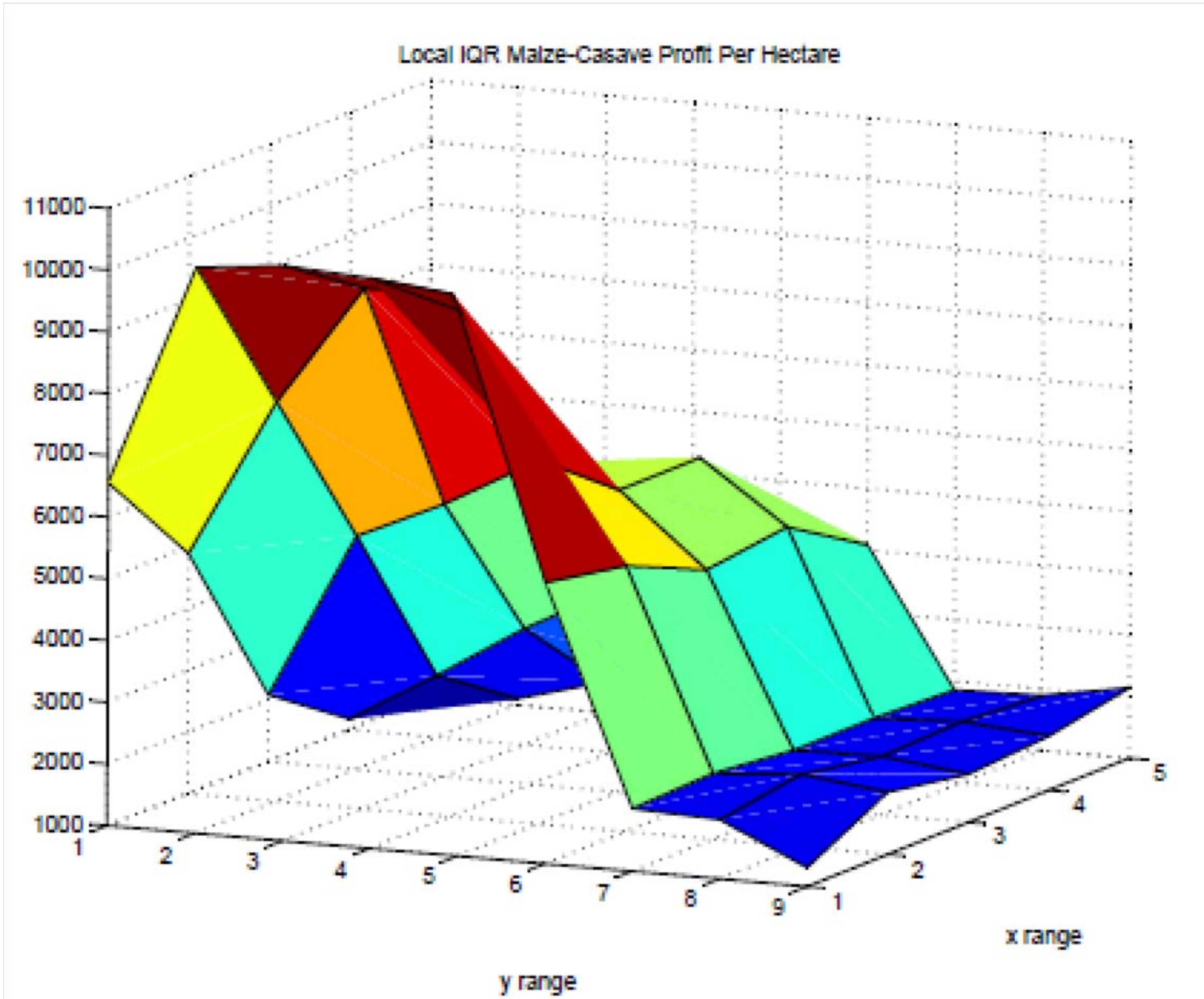
- This occurs, at least in West Africa, at the scale of kilometers

- Dramatic consequences for understanding agricultural production functions, profitability of new technologies, even social organization,


Plot Locations
6.54
6.52
6.5
6.48
6.46
6.44
6.42
6.4
8.05
8.06
8.07
8.08
8.09
8.1
8.11
8.12
8.13



Smoothed Mean Maize-Casave Profit Per Hectare
8000
6000
4000
2000
0
5
4.5
4
3.5
3
2.5
2
1.5
1
x range
1
2
3
4
5
6
7
8
9
y range


Local IQR Maize-Casave Profit Per Hectare
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
1
2
3
4
5
6
7
8
9
1
2
3
4
5
x range
y range

Moreover, ‘space’ is correlated with other things that matter:

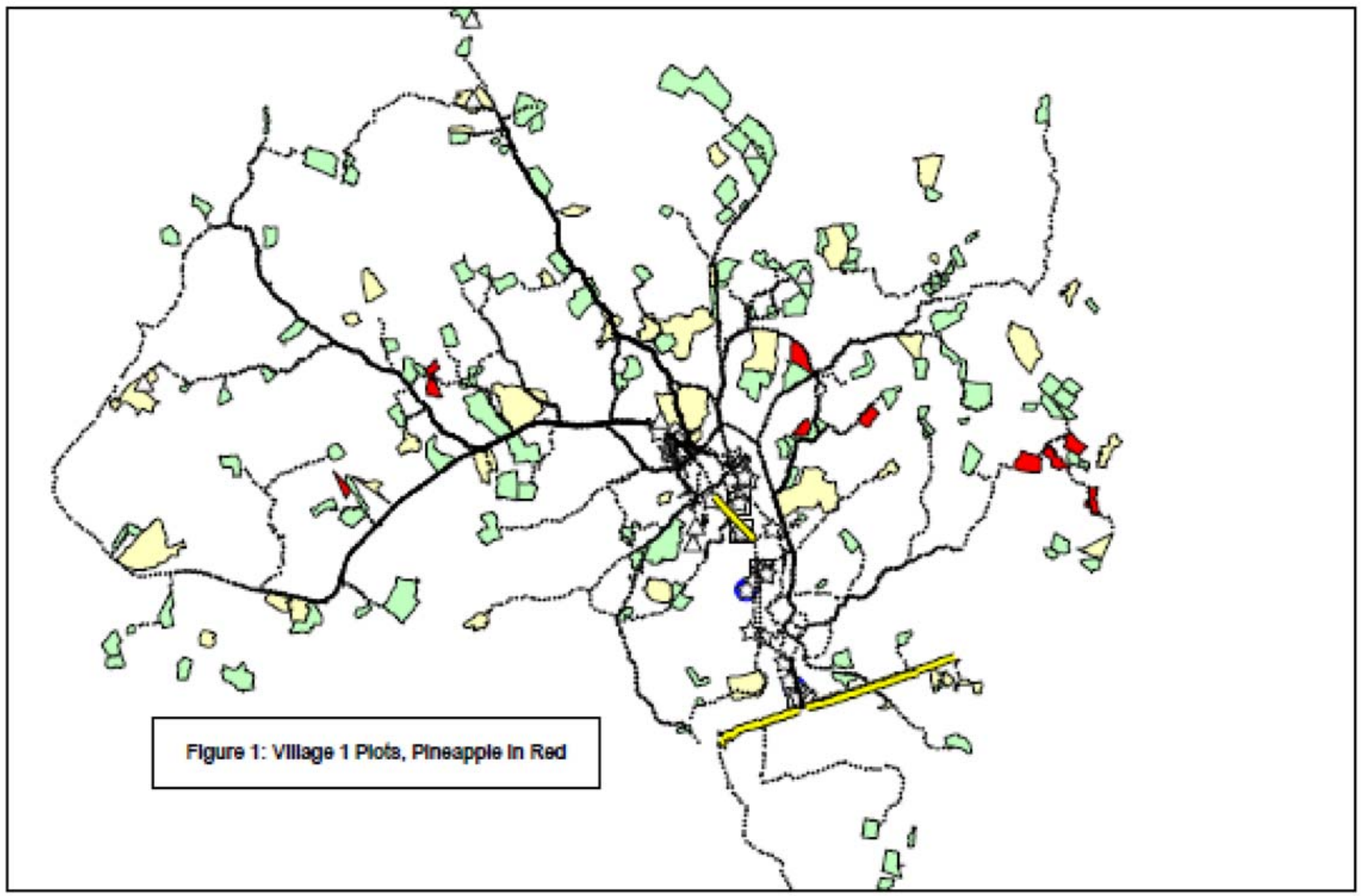


Figure 1: Village 1 Plots, Pineapple in Red

So there is something typically ‘unobserved’ that is correlated with things you care about. Spatial effects exist and matter.

$$\pi_i(Z_{pi}, w_i, \rho_i) = \max_{x_{pi}} E\rho_i f(x_{pi}.Z_{pi}, \omega_{pi}) + \varepsilon_{pi} - w_i x_{pi}$$

where

x_{pi} - vector of all the inputs, w is prices;

$Z_{pi}-$observable characteristics of p and i (human capital? land characteristics?),

ω_{pi} shocks known by i before production choices are made

$f(.)$ (scalar) production function, ρ price of output

E taken over distribution of ε.

First-order approximation:

$$\pi_{pi} = \mathbf{Z}_{pi}\beta + \gamma G_{pi} + \lambda_i + \lambda_{N(pi)} + \varepsilon_{pi}$$

NOTE: Two different fixed effects. This is a simple version, more generally $\lambda_{N(pi)}$ is a smooth function of space.

Estimator:

$$\begin{aligned} \pi_{pi} - \frac{1}{N_{pi}} \sum_{q \in N_p} \pi_q \;&=\; (\mathbf{Z}_{pi} - \frac{1}{N_{pi}} \sum_{q \in N_{pi}} Z_q)\beta + \gamma(G_{pi} - \frac{1}{N_{pi}} \sum_{q \in N_{pi}} G_q) \\ &\quad \lambda_i - \frac{1}{N_{pi}} \sum_{j \in N_{pi}} \lambda_j + \varepsilon_{pi} - \frac{1}{N_{pi}} \sum_{q \in N_{pi}} \varepsilon_{qj}. \end{aligned}$$

(stata program example available)

Table 3: Profits and Gender

	2		3	
	OLS		OLS	
dependent variable	profit x1000 cedis/hect		profit x1000 cedis/hect	
	estimate	std error	estimate	std error
gender: 1=woman	-1043	300	-1667	374
Plot Size Decile = 2	447	179	1002	244
Plot Size Decile = 3	1039	295	475	267
Plot Size Decile = 4	1135	302	788	298
Plot Size Decile = 5	657	134	578	128
Plot Size Decile = 6	811	163	97	210
Plot Size Decile = 7	875	172	220	249
Plot Size Decile = 8	439	302	-374	274
Plot Size Decile = 9	249	284	-120	251
Plot Size Decile = 10	-316	332	-1195	339
Soil Type = Loam	-175	211	-442	160
Soil Type = Clay	-512	294	-525	324
Toposequence: midslope	299	334	-468	389
Toposequence: bottom	663	337	-525	435
Toposequence: steep	3	365	971	577
pH	-260	89	155	43
Organic Matter	-16	52	-347	76
Observations	508		508	
Fixed effects	household x year		spatial (250 meters) and household x year	

Standard errors are consistent for arbitrary heteroskedasticity and spatial correlation.

Profits of a New Technology

$$\pi_i(w_i, \rho_i, \upsilon_i) = \max_{\Upsilon_i, x_i} E\rho_i f(x_i, \omega_i; \Upsilon_i) + \varepsilon_i - w_i x_i - \upsilon_i \Upsilon_i$$

Randomize $\upsilon_i \in \{\upsilon^L, \upsilon^H\}$, $\upsilon^H > \upsilon^L$

- Provides $E\left(\pi(1) - \pi(0)\right)$ over k_i, ω_i, ρ, w
- obvious that $E\pi_i(1) > E\pi_i(0) \not\Leftrightarrow$ yield (1) $>$ yield (2)
- obvious that $E\left(\pi_i(1) - \pi_i(0)\right) > 0$ does not imply that $\pi_i(1) - \pi_i(0) > 0$ for all i.

Selection on ω

suppose $\omega f(x; \Upsilon)$ and $\Upsilon = 1$ increases yield

$\pi^H(\omega; 1) = \max_{x_i} E\rho f(x_i, \omega; 1) + \varepsilon_i - wx_i - \upsilon^H$ and

$\pi^H(\omega; 0)$, $\pi^L(\omega; 1)$ and $\pi^H(\omega; 0)$.

$\exists\ \omega^H$ s.t. for $\omega_i \leq \omega^H$, all $i \in H$ have $\pi^H(\omega_i; 0) > \pi^H(\omega_i; 1)$ and choose $\Upsilon_i = 0$; and $\Upsilon_i = 1$ for all $\omega_i > \omega^H$.

Similarly, there is ω^L $(< \omega^H)$

The obvious approach to estimating the expected profitability of the new technology is to estimate the regression

$$\pi_i = \alpha + \beta \Upsilon_i + e_i$$

using T_i as an instrument for Υ_i.

$$\hat{\beta} = E\left([\pi_i(1) - \pi_i(0)] \,|\omega^L < \omega_i < \omega^H \right)$$

- choosing the level of υ^H, υ^L doesn't just change the power of the experiment. It changes what you are estimating

- So, for example, increasing the subsidy for adoption (lowering υ^L) reduces ω^L and thus reduces $\hat{\beta}$.

Understanding the Production Function

$$x_i(\rho_i, w_i; \omega_i) = \arg\max_x E\rho_i f(x, \omega_i) + \varepsilon_i - w_i x_i$$

randomize factor prices over individuals, and we can estimate the demand functions and hence f.

(obviously, it's not really so simple. Even if all else is easy, we need to be concerned about functional form and dimensionality)

Randomization of factor prices to estimate production functions seems such an obvious thing. But we can't find examples. Why not?

1. Production economists estimate production functions by experimentally varying inputs and observe outputs (Close to Duflo, Kremer and Robinson in their early Busia fertilizer work). Recent examples: Canchi et al. (2010);Tembo et al. (2008)

   (a) Like experiment station trials, but we want these on farmers' plots, integrated into farming systems

   (b) Once integrated into farmers' overall decision-making, we get optimizing behavior, which leads us back to estimating $x()$

2. $x(.)$ depends on farmer knowledge, learning

3. How to induce variation in w? With complete markets, it's not obvious. With a subsidy, just demand an infinite amount and sell the surplus. With a tax, just buy on the market.

   (a) Transaction costs provide a window through which small variation in w may be possible, which in turn permits local estimates of $x(.)$ and $f(.)$.

4. Conditioning on ω is required. This is hard; one possibility is to rely on spatial correlation to take care of most of it.

$x_i \in \{1, 0\}$

$x_i = 1$ iff $\omega_i f(1) - w_i \geq \omega_i f(0)$

IV: input price $w_{i.}$ We set $w_i = w^H$ for randomly-chosen group H, $w_i = w^L$ for group L. For $i \in H$,

$$x_i = 1 \text{ if } \omega_i(f(1) - f(0)) > w^H, 0 \text{ otherwise}$$

for $j \in L$,

$$x_j = 1 \text{ if } \omega_j(f(1) - f(0)) > w^L, 0 \text{ otherwise.}$$

We observe $y_i = \omega_i f(x_i) + \varepsilon_{i.}$ Now estimate

$$y_i = \alpha + \beta x_i + e_i$$

using w_i as an instrument for x_i

$$\hat{\beta} = E\left(\omega(f(1) - f(0))|\frac{f(1) - f(0)}{w^H} < \omega_i < \frac{f(1) - f(0)}{w^L}\right)$$

- Standard LATE issues, as in the discussion above