Selection and Comparative Advantage in Technology Adoption

Tavneet Suri

MITSloan MANAGEMENT

The Empirical Puzzle

Average gross returns (experimental) to agricultural technologies such as hybrid maize and fertilizer are extremely high

Puzzles:

i. Despite high average returns, a significant number of households do not use these technologies: why not?
ii. Adoption rates show no accelerating increases

Explanations in the literature:

i. Lack of good information and slow learning
ii. Credit constraints
iii. Consumption tastes

MIT Sloan MANAGEMENT

Motivation

TRENDS IN YIELDS OF STAPLES: AVERAGE ANNUAL % CHANGES IN YIELDS (HG/HA) BY DECADE[a]

	1961–1970	1971–1980	1981–1990	1991–2004
Kenya				
Maize	0.362	2.373	1.169	–1.198
Wheat	5.646	2.333	–3.078	0.984
India				
Maize	1.502	0.842	1.900	2.572
Wheat	4.876	2.514	3.343	1.235
Rice	0.954	1.714	3.310	0.838
Mexico				
Maize	2.057	4.267	–0.548	1.447
Wheat	4.586	3.204	–0.255	1.664
Zambia				
Maize	–0.267	10.403	1.571	–1.707

MIT Sloan MANAGEMENT

What About the Farmers?

Priorities to Improve Family Well Being	HHs for whom it is the Top Priority	HHs that Place it in the Top 3 Priorities
Increase Yields on Existing Land	39.4%	72.7%
Obtain More Land	16.4%	29.2%
Obtain More Animals	14.5%	55.4%
Start a Business/ Earn more from Business	23.2%	48.4%
Education	2.3%	5.9%
Credit	0.3%	0.5%

Explanations in the Literature

i. Lack of good information about the technologies

ii. Slow learning about the technologies (several trials needed before you understand the true benefits)

iii. Credit constraints – households can't afford the technologies

iv. Consumption tastes are different (e.g. Latin America)

v. Farmers are hyperbolic when it comes to decisions about technologies (they can not commit to the expenditures at the time of harvest)

Hypothesis in This Paper

- The hypothesis is quite simple
- Are there are differences in returns to the technology across farmers?
- If so, is it simply the case that the farmers who do not use it simply do not benefit from it?
- Why might there be differences in these returns?

In Particular...

What is the spatial distribution of returns? How do these returns correlate with adoption decisions & observables? Are average returns high, but marginal returns low?

I model two forms of heterogeneity:

i. Absolute advantage (individual specific intercepts/average yields, irrespective of technology)
ii. Comparative advantage (individual specific gains to hybrid, i.e. individual specific slopes)

Bottom line: how important is the role of comparative advantage?

MITSloan MANAGEMENT

Summary of Findings

I find evidence of cross-sectional heterogeneity in returns to hybrid with three interesting sub-groups of farmers in the population:

i. A small group has potentially high returns from adopting, yet they do not adopt. These farmers seem to have high unobservable costs (they have poor access to input distribution/infrastructure)
ii. A larger group has smaller positive returns yet adopt every period
iii. Farmers that switch in and out of adoption have approximately zero returns

Implies adoption decisions are well explained by variation (observable and unobservable) in heterogeneous net benefits to the technology

MITSloan MANAGEMENT

The Policy Questions

Policy implications are very different if returns are not homogeneous

Given the limited resources of policy makers, how do you target policy to be the most cost effective?

Should policy makers expand extension or develop new varieties?

If households have zero returns, policy makers should not encourage adoption of existing varieties via say extension services

Households that have lower returns and use hybrid don't seem to be constrained and would benefit from development of new strains

Talk Outline

- **Maize Cultivation, Data, and Adoption Trends**
- Modeling Adoption Decisions
- Estimating a Model with Heterogeneous Returns
- Results
  - Summary Statistics
  - Baseline OLS and Fixed Effects Results
  - Motivation for Heterogeneity
  - Correlated Random Coefficients Estimates
- Are Adoption Decisions Unconstrained? Implications for Policy
  - What Costs?
- Conclusions

MITSloan MANAGEMENT

Maize in Kenya

Maize is the main staple, little or no irrigation

Technologies available since 1960's (>20 seeds released since 1955)

Initial adoption rates were high, but no sustained yield increases

Gerhart (1975) about the diffusion of H611 in Western Kenya:

"at rates as fast as or faster than among farmers in the US corn belt"

Pan territorial seed pricing for most of this period, most of the seed is distributed by Kenya Seed Company (this is key)

MITSloan MANAGEMENT

Data and Field Work

Tegemeo Agricultural Monitoring & Policy Analysis (TAMPA) Project: a panel household survey across a lot of rural Kenya, 1997 to 2004

The hybrid decision is a binary choice (only 1% of HHs plant both)

I use data on yields, hybrid decisions and inputs (land, seed use, fertilizer use, labor (family and hired), land preparation, rainfall)

I have complete input/output data for only 1997, 2000 and 2004

There is also a large qualitative component to these surveys

Sample Villages

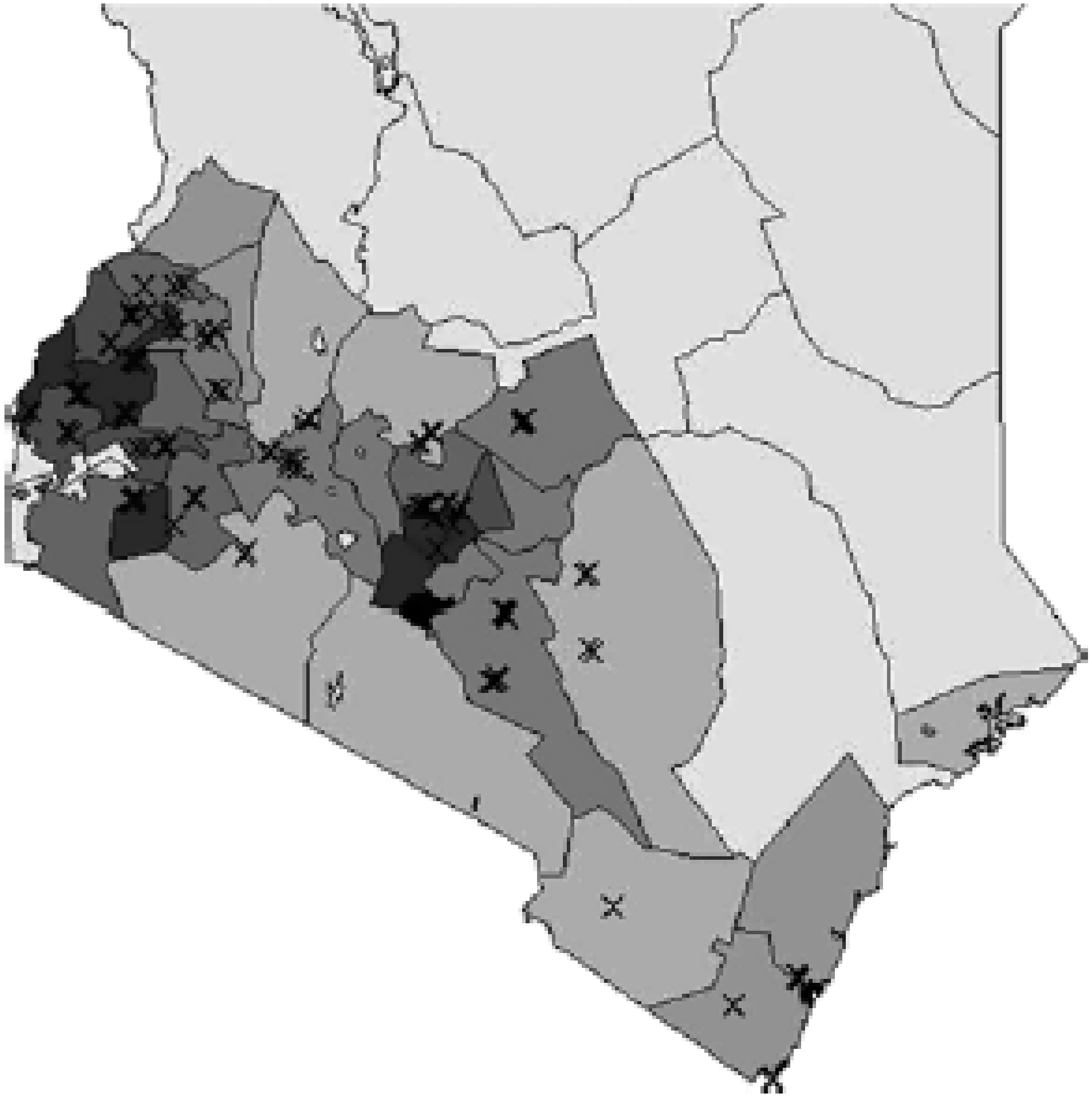


FIGURE 1.—Population density and location of sample villages.

Hybrid Adoption (Province)

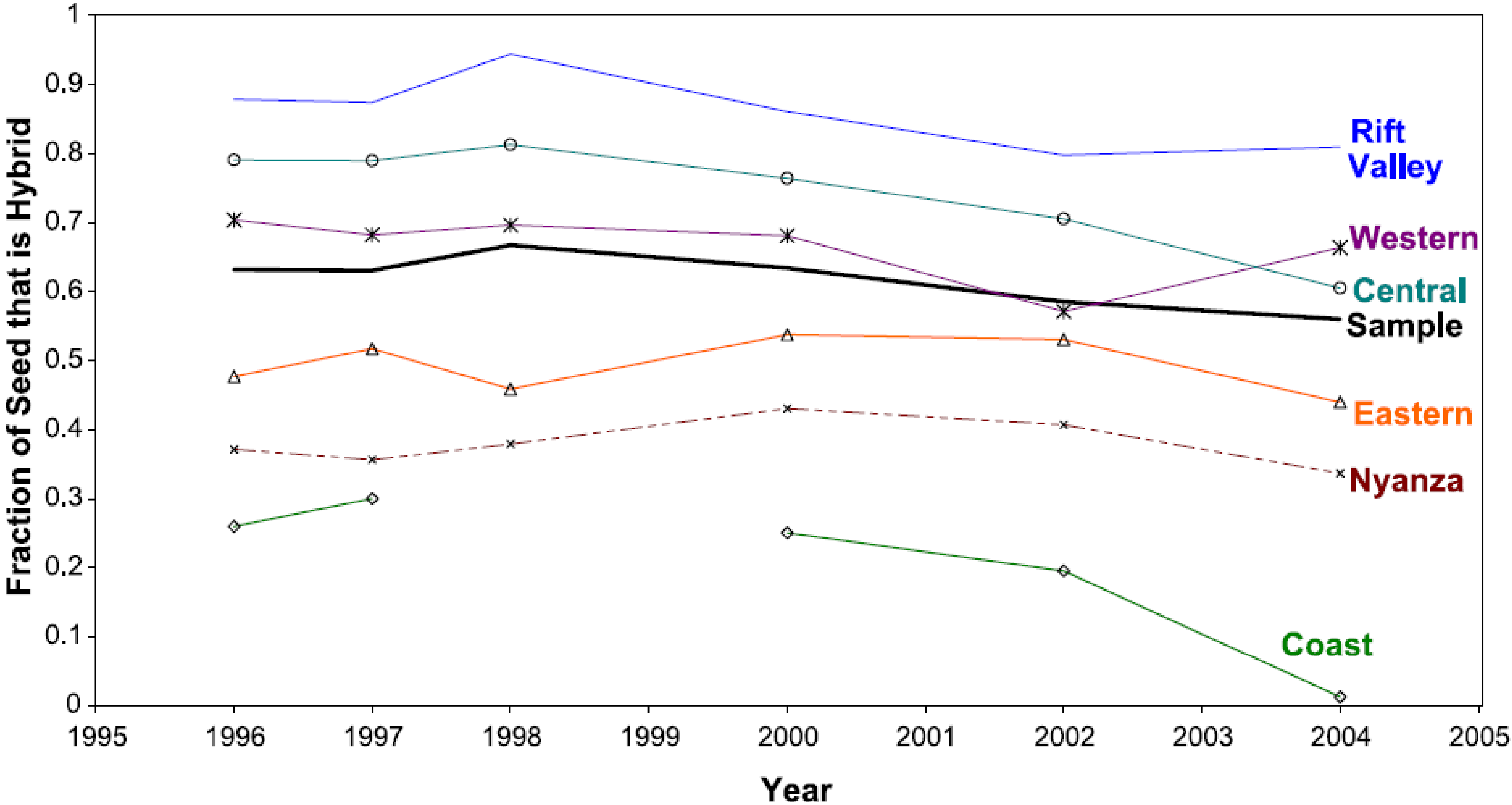


FIGURE 2.—Hybrid maize adoption patterns by province.

Fertilizer Adoption (Province)

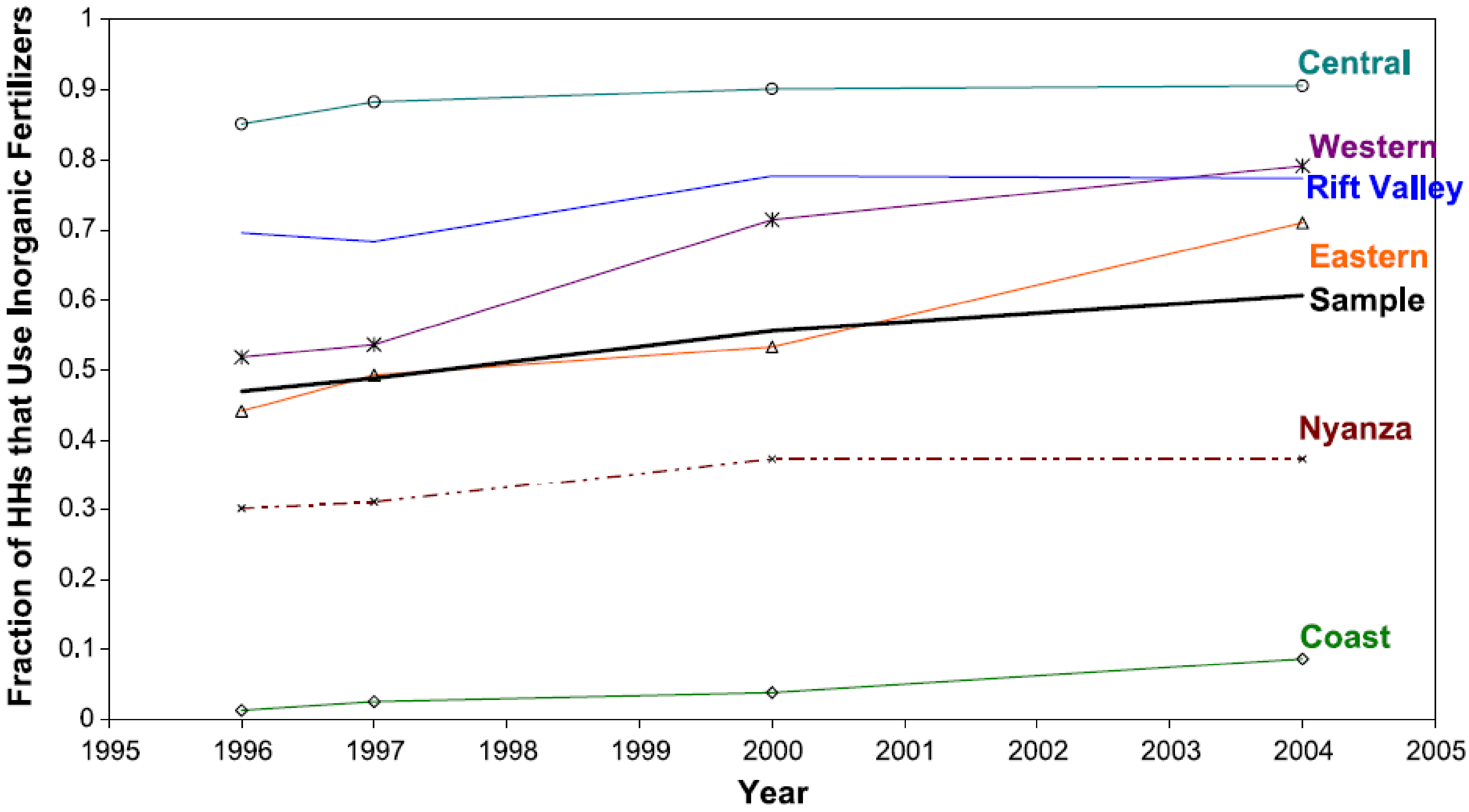


FIGURE 3A.—Fraction of households using inorganic fertilizer by province.

Fertilizer Expenditure

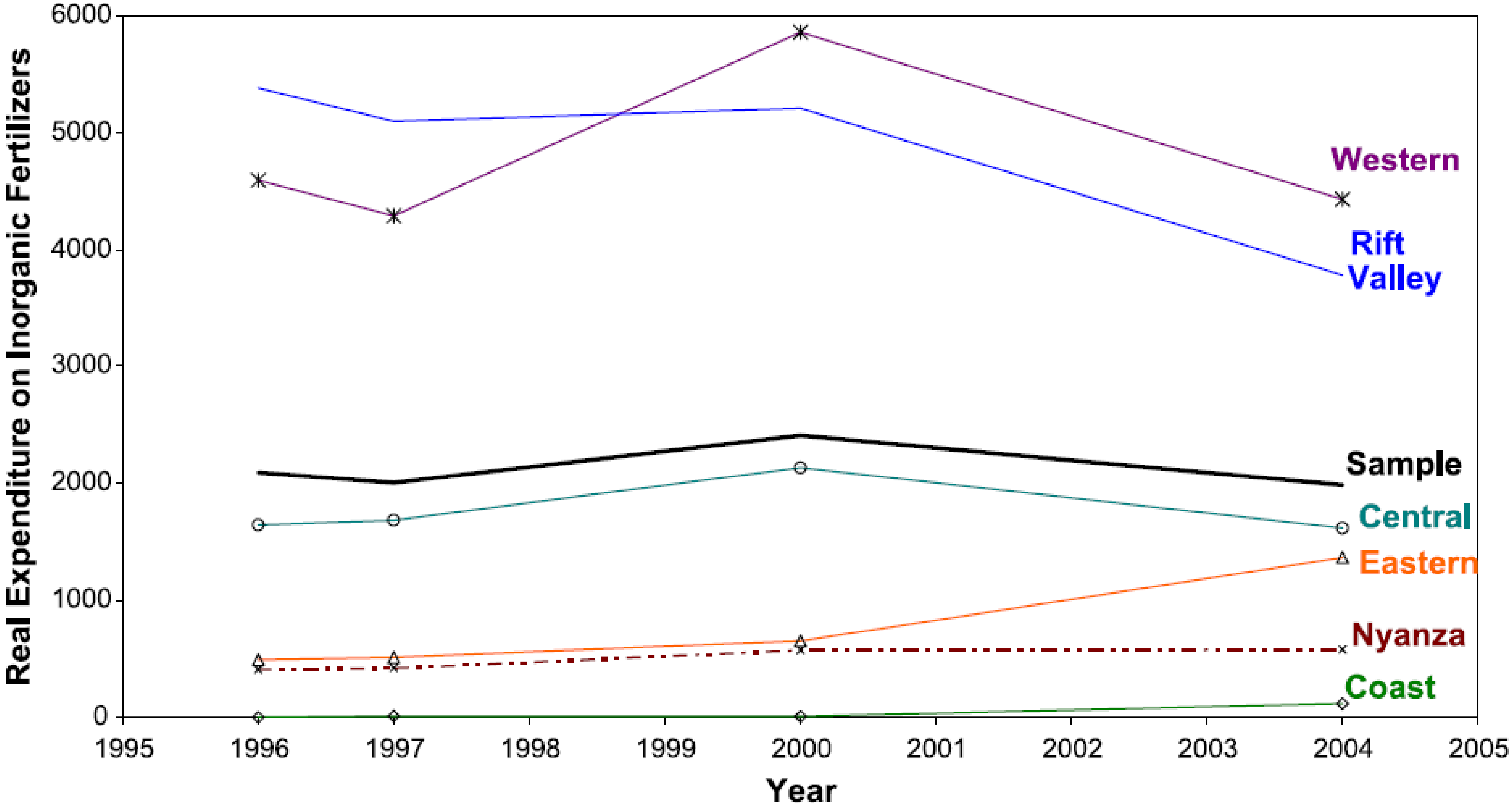


FIGURE 3B.—Real expenditure on inorganic fertilizer by province.

MIT Sloan MANAGEMENT

Switching

TRANSITIONS ACROSS HYBRID/NONHYBRID SECTORS FOR THE SAMPLE PERIODS 1997, 2000, AND 2004[a]

Transition in Terms of Technology Used (1997 2000 2004)	Fraction of Sample (%) ($N = 1202$ Households)
N N N	20.38
N N H	2.83
H N H	6.07
N H H	4.91
H N N	5.99
H N H	3.16
H H N	7.15
H H H	49.50

Yields by Sector, 2004

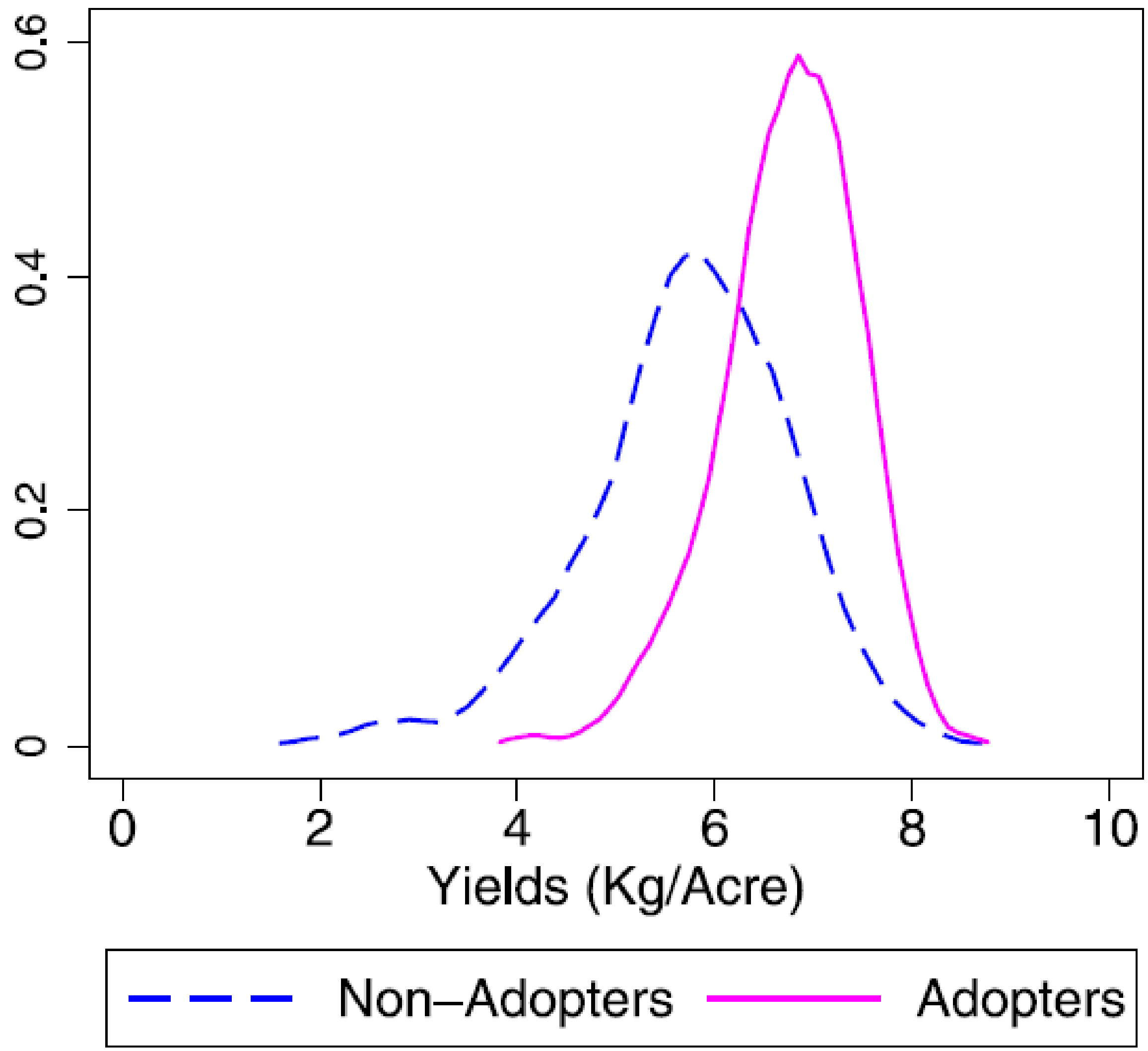


FIGURE 4B.—Marginal distribution of yields by sector, 2004.

Talk Outline

- Data and Adoption Trends
- **Modeling Adoption Decisions**
- Estimating a Model with Heterogeneous Returns
- Results
  - Summary Statistics
  - Baseline OLS and Fixed Effects Results
  - Motivation for Heterogeneity
  - Correlated Random Coefficients Estimates
- Are Adoption Decisions Unconstrained? Implications for Policy
  - What Costs?
- Conclusions

MITSloan MANAGEMENT

Modeling Adoption Decisions

Say the hybrid and non-hybrid profit functions for a farmer are of the following Cobb-Douglas form:

$$(1) \qquad \pi_{it}^{H} = p_{it} Y_{it}^{H} - (b_t s_{it} + a_{it}) - \sum_{j=1}^{J} w_{jit} X_{jit}^{H}$$

$$(2) \qquad \pi_{it}^{N} = p_{it} Y_{it}^{N} - c_{it} s_{it} - \sum_{j=1}^{J} w_{jit} X_{jit}^{N}$$

MITSloan MANAGEMENT

What Drives Adoption I

The simplest case is the Roy model where the farmer adopts if

$$\pi_{it}^{*H} > \pi_{it}^{*N}$$

This implies sorting based on comparative advantage

Happens when (equation 4):

$$\left(Y_{it}^{*H} - \sum_{j=1}^{J} \frac{w_{jit}}{p_{it}} X_{jit}^{*H}\right) - \left(Y_{it}^{*N} - \sum_{j=1}^{J} \frac{w_{jit}}{p_{it}} X_{jit}^{*N}\right) > \frac{a_{it}}{p_{it}} + \frac{\delta_{it}^{s}}{p_{it}} \equiv A_{it} + \Delta_{it}^{s}$$

MITSloan MANAGEMENT

What Drives Adoption II

With some assumptions this boils down to:

$$(5) \qquad ({Y_{it}^{*}}^{H} - {Y_{it}^{*}}^{N}) > A_{it} + \Delta_{it}^{s}$$

Note that I never actually estimate the adoption decision directly

Instead I focus on estimating comparisons of underlying yield functions

- One advantage of this is household labor (large fraction of total labor) is hard to value

Yields

Instead, I estimate the underlying yield functions:

$$(6) \qquad Y_{it}^{H} = e^{\beta_t^H} \left(\prod_{j=1}^{k} X_{ijt}^{\gamma_j^H} \right) e^{u_{it}^H},$$

$$(7) \qquad Y_{it}^{N} = e^{\beta_t^N} \left(\prod_{j=1}^{k} X_{ijt}^{\gamma_j^N} \right) e^{u_{it}^N}.$$

Or in logs,

$$(8) \qquad y_{it}^{H} = \beta_t^H + x'_{it}\gamma^H + u_{it}^H$$

$$(9) \qquad y_{it}^{N} = \beta_t^N + x'_{it}\gamma^N + u_{it}^N$$

Structure of the Errors I

Now put some structure on the errors to allow both absolute and comparative advantage:

$$u_{it}^{H} = \theta_i^{H} + \xi_{it}^{H} \tag{10}$$

$$u_{it}^{N} = \theta_i^{N} + \xi_{it}^{N} \tag{11}$$

Decompose as follows:

$$\theta_i^{H} = b_H(\theta_i^{H} - \theta_i^{N}) + \tau_i \tag{12}$$

$$\theta_i^{N} = b_N(\theta_i^{H} - \theta_i^{N}) + \tau_i \tag{13}$$

MITSloan MANAGEMENT

Structure of the Errors II

Following Heckman & Honore (1990) and Lemieux (1998), I use linear projections of θ_i^H and θ_i^N on (θ_i^H-θ_i^N):

$$\theta_i^H = b_H(\theta_i^H - \theta_i^N) + \tau_i$$

$$\theta_i^N = \underbrace{b_N(\theta_i^H - \theta_i^N)}_{\text{Redefine to be } \theta_i} + \tau_i$$

⇨

$$\theta_i^H = (\phi + 1)\theta_i + \tau_i$$

$$\theta_i^N = \theta_i + \tau_i$$

where

$$b_H = (\sigma_H^2 - \sigma_{HN})/(\sigma_H^2 + \sigma_N^2 - 2\sigma_{HN})$$

$$\sigma_{HN} \equiv cov(\theta_i^H, \theta_i^N)$$

$$\sigma_H^2 \equiv Var(\theta_i^H)$$

$$\phi \equiv 1 - \frac{b_H}{b_N}$$

Yields Once Again

Substituting back into the yield equations:

$$(17) \qquad y_{it}^{H} = \beta_t^{H} + \tau_i + (\phi + 1)\theta_i + X_{it}\gamma^{H} + \xi_{it}^{H}$$

$$(18) \qquad y_{it}^{N} = \beta_t^{N} + \tau_i + \theta_i + X_{it}\gamma^{N} + \xi_{it}^{N}.$$

These yield functions look like they can be estimated by fixed effects

But, they cannot!! The thetas cannot be differenced away (they are technology specific)

MITSloan MANAGEMENT

Yields Once Again

Using a generalized yield equation:

$$(19) \qquad y_{it} = h_{it} y_{it}^{H} + (1 - h_{it}) y_{it}^{N}$$

Substituting in for yields

$$(20) \qquad y_{it} = \beta_t^N + \theta_i + (\beta_t^H - \beta_t^N) h_{it} + X_{it}' \gamma^N + \phi \theta_i h_{it} + X_{it}' (\gamma^H - \gamma^N) h_{it} + \tau_i + \varepsilon_{it}$$

where θ_i (comparative advantage) and τ_i (absolute advantage) are uncorrelated by construction

MITSloan MANAGEMENT

Now Back to Adoption…

- Relating this back to adoption, the farmer adopts if

$$E(u_{it}^{H} - u_{it}^{N}) > A_{it} + \Delta_{it}^{s} - (\beta_{t}^{H} - \beta_{t}^{N}) + \sum_{j=1}^{J} (\gamma_{j}^{N} x_{jit}^{*N} - \gamma_{j}^{H} x_{jit}^{*H})$$

- With some assumptions, this boils down to

$$(25) \qquad \phi\theta_i - \alpha_i > \Delta_{it}^{s} - (\beta_{t}^{N} - \beta_{t}^{H}) + \vartheta_{it}$$

i.e. when benefits are greater than costs (fixed and variable)

Talk Outline

- Data and Adoption Trends
- Modeling Adoption Decisions
- **Estimating this Model with Heterogeneous Returns**
- Results
  - Summary Statistics
  - Baseline OLS and Fixed Effects Results
  - Motivation for Heterogeneity
  - Correlated Random Coefficients Estimates
- Are Adoption Decisions Unconstrained? Implications for Policy
  - What Costs?
- Conclusions

MITSloan MANAGEMENT

Correlated Random Coefficients

The basic yield function I want to estimate is:

$$y_{it} = \beta_t^N + \theta_i + (\beta_t^H - \beta_t^N)h_{it} + X_{it}'\gamma^N + \phi\theta_i h_{it} + X_{it}'(\gamma^H - \gamma^N)h_{it} + \tau_i + \varepsilon_{it}.$$

Identification assumption:

$$(26) \qquad E(\tau_i + \varepsilon_{it} | \theta_i, h_{i1}, \ldots, h_{iT}, X_{i1}, \ldots, X_{iT}) = 0$$

Identification I

Remember

$$(10) \qquad u_{it}^H = \theta_i^H + \xi_{it}^H$$

$$(11) \qquad u_{it}^N = \theta_i^N + \xi_{it}^N$$

Key Assumption: transitory errors ξ_{it}^H and ξ_{it}^N do not affect farmer's decision to use hybrid [θ_i^H and θ_i^N will as farmers know these]

Implies: ξ_{it}^H and ξ_{it}^N are known **after** the planting decision is made

MITSloan MANAGEMENT

Identification II

The timing of maize production and rainfall is crucial here

An important component of the transitory shock ξ_{it} is rainfall which is observed by the farmer after planting

The seed type is fixed before this shock is (fully) realized

But other inputs may be correlated with the shock

I condition on the covariates affecting yields in my data and most importantly I condition on seasonal rainfall (which I observe)

Estimating CRC I

Start with explaining the simpler version of the yield equation:

$$y_{it} = \delta + \beta h_{it} + \theta_i + \phi \theta_i h_{it} + u_{it}$$

Project θ_i onto the history and the interactions of the hybrid histories:

$$(29) \qquad \theta_i = \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i1} h_{i2} + v_i$$

Estimation and identification similar to Chamberlain CRE model

In the estimation, I impose directly $\sum \theta_i = 0$

Estimating CRC II

Basic strategy:

- Substitute the project back in to the yield functions period by period
- Gives you the structural equations for each period
- Can estimate reduced forms for each period (using SUR) and then the structural parameters using minimum distance
- Six reduced form parameters ($\gamma_1 - \gamma_6$) – the reduced forms include all the interactions of the hybrid histories
- These map onto five structural parameters ($\varphi\ \beta\ \lambda_1\ \lambda_2\ \lambda_3$)
- The structural parameters over-identified

Estimating CRC III

Minimum distance restrictions:

$$
\begin{aligned}
(34) \qquad \gamma_1 &= (1+\phi)\lambda_1 + \beta + \phi\lambda_0, \\
\gamma_2 &= \lambda_2, \\
\gamma_3 &= (1+\phi)\lambda_3 + \phi\lambda_2, \\
\gamma_4 &= \lambda_1, \\
\gamma_5 &= (1+\phi)\lambda_2 + \beta + \phi\lambda_0, \\
\gamma_6 &= (1+\phi)\lambda_3 + \phi\lambda_1.
\end{aligned}
$$

Talk Outline

- Data and Adoption Trends
- Modeling Adoption Decisions
- Estimating a Model with Heterogeneous Returns
- **Results**
  - **Summary Statistics**
  - Baseline OLS and Fixed Effects Results
  - Motivation for Heterogeneity
  - Correlated Random Coefficients Estimates
- Are Adoption Decisions Unconstrained? Implications for Policy
  - What Costs?
- Conclusions

Summary Statistics I

	1997 Sample	2004 Sample
Yield (log maize harvest per acre)	5.907 (1.153)	6.350 (0.977)
Acres planted	1.903 (3.217)	1.957 (2.685)
Total seed planted (kg per acre)	9.575 (7.801)	9.072 (6.863)
Total purchased hybrid planted (kg per acre)	6.273 (6.926)	5.080 (5.260)
Hybrid (dummy)	0.658 (0.475)	0.604 (0.489)
Fertilizer (kg DAP (diammonium phosphate) per acre)	20.300 (38.444)	24.610 (34.001)
Fertilizer (kg MAP (monoammonium phosphate) per acre)	1.566 (10.165)	0.308 (4.538)
Fertilizer (kg CAN (calcium ammonium nitrate) per acre)	6.473 (24.727)	8.957 (21.702)
Total fertilizer expenditure (KShs per acre)	1361.7 (2246.3)	1354.6 (1831.2)
Land preparation costs (KShs per acre)	960.88 (1237.1)	541.43 (1022.8)
Family labor (hours per acre)	293.25 (347.49)	354.27 (352.68)
Hired labor (KShs per acre)	1766.0 (3346.4)	1427.4 (2130.3)
Main season rainfall (mm)	620.83 (256.43)	728.11 (293.29)
Distance to closest fertilizer seller (km)	6.288 (9.774)	3.469 (5.964)
Household size	7.109 (2.671)	8.409 (3.521)

[a] Standard deviations are given in parentheses. KShs is Kenyan shillings (exchange rate over this period was KShs 75 = $ 1). All monetary variables are in real terms.

Summary Statistics II

	1997 Sample		2004 Sample	
	Hybrid	Nonhybrid	Hybrid	Nonhybrid
No. of households	791	411	726	476
Yield (log maize harvest per acre)	6.296 (0.934)	5.158 (1.167)	6.751 (0.692)	5.738 (1.030)
Total maize acres cultivated	1.982 (3.557)	1.753 (2.428)	2.087 (3.029)	1.758 (2.042)
Total seed planted (kg per acre)	9.669 (6.569)	9.394 (9.750)	8.746 (4.156)	9.569 (9.608)
Fertilizer (kg DAP per acre)	28.755 (44.115)	4.028 (13.266)	37.148 (37.294)	5.488 (13.909)
Fertilizer (kg CAN per acre)	9.087 (29.715)	1.442 (7.152)	12.708 (24.961)	3.235 (13.622)
Land preparation costs (KShs/acre)	1043.9 (1242.7)	801.08 (1211.7)	659.83 (1079.7)	360.83 (901.0)
Expenditure on fertilizer (KShs/acre)	1922.3 (2542.9)	282.64 (740.53)	1893.3 (1964.7)	533.09 (1211.4)
Inorganic fertilizer use (dummy)	0.7421 (0.4378)	0.2311 (0.4221)	0.8994 (0.3009)	0.4055 (0.4915)
Main season rainfall (mm)	651.70 (228.82)	561.44 (293.88)	825.41 (215.20)	579.69 (332.05)
Hired labor (KShs/acre)	1864.3 (2680.6)	1576.7 (4347.8)	1616.5 (2197.4)	1139.0 (1991.6)
Family labor (hours/acre)	260.35 (264.13)	356.57 (461.71)	343.6 (336.1)	370.58 (376.33)
Distance to closest fertilizer seller (km)	4.684 (7.993)	9.374 (11.93)	2.419 (2.420)	5.069 (8.760)
Household size	7.162 (2.616)	7.007 (2.773)	8.457 (3.340)	8.336 (3.783)

Talk Outline

- Data and Adoption Trends
- Modeling Adoption Decisions
- Estimating a Model with Heterogeneous Returns
- **Results**
  - Summary Statistics
  - **Baseline OLS and Fixed Effects Results**
  - Motivation for Heterogeneity
  - Correlated Random Coefficients Estimates
- Are Adoption Decisions Unconstrained? Implications for Policy
  - What Costs?
- Conclusions

MITSloan MANAGEMENT

OLS and FE Results

	OLS, Pooled	OLS, Pooled	OLS, Pooled	FE	FE
Hybrid	1.074 (0.040)	0.695 (0.039)	0.541 (0.041)	0.017 (0.070)	0.090 (0.065)
Acres (× 1000)	—	–	0.035 (5.749)	—	−0.509 (0.140)
Seed kg per acre (× 10)	—	–	0.184 (0.024)	—	0.179 (0.032)
Land preparation costs per acre (× 1000)	—	–	0.066 (0.016)	—	0.075 (0.023)
Fertilizer per acre (× 1000)	—	–	0.075 (0.009)	—	0.054 (0.012)
Hired labor per acre (× 1000)	—	–	0.037 (0.006)	—	0.027 (0.008)
Family labor per acre (× 1000)	—	–	0.374 (0.050)	—	0.467 (0.072)
Year = 2004	0.501 (0.038)	0.480 (0.035)	0.566 (0.041)	0.444 (0.032)	0.587 (0.044)
Constant	5.200 (0.038)	4.636 (0.080)	3.954 (0.113)	5.896 (0.051)	−2.383 (5.582)
Province dummies	No	Yes	Yes	—	—
R-squared	0.266	0.400	0.502	0.049	0.089

Talk Outline

- Data and Adoption Trends
- Modeling Adoption Decisions
- Estimating a Model with Heterogeneous Returns
- **Results**
  - Summary Statistics
  - Baseline OLS and Fixed Effects Results
  - **Motivation for Heterogeneity**
  - Correlated Random Coefficients Estimates
- Are Adoption Decisions Unconstrained? Implications for Policy
  - What Costs?
- Conclusions

Motivation for Heterogeneity

i. Heckit Selection Equations

ii. ATE, TT, MTE (selection corrected)

iii. IV/LATE Estimates

Excluded Regressor: Distance to closest fertilizer distributor

MIT Sloan Management

ATE, TT, MTE Estimates

Exclusion restriction: use distance to closest fertilizer seller (km)

	Heckman Two-Step Estimates: Selection Correction λ	
Year	Hybrid Sector	Nonhybrid Sector
1997	−0.854 (0.170)	1.659 (0.864)
2004	−0.957 (0.181)	0.028 (0.152)

	Implied Treatment Effects		
Year	ATE	TT	MTE Slope
1997	2.391	0.917	−2.512 (0.880)
2004	1.279	0.921	−0.985 (0.237)

MITSloan MANAGEMENT

LATE Estimates

Exclusion restriction: use distance to closest fertilizer seller (km)

First stage: Effect of distance	–0.288 (0.108)	—
First stage: Effect of distance interacted with wealth quintile (× 100)		
Second wealth quintile (coefficient on interaction)	—	–0.221 (0.302)
Third wealth quintile (coefficient on interaction)	—	–0.057 (0.032)
Fourth wealth quintile (coefficient on interaction)	—	0.329 (0.288)
Fifth wealth quintile (coefficient on interaction)	—	0.507 (0.273)
F test p-value on excluded instruments	0.008	0.108
Second stage: Effect of predicted hybrid on yields	2.768 (1.123)	1.536 (0.816)

Selection and Heterogeneity

Variable	Without Covariates: 1997 Yield	Without Covariates: 2004 Yield	With Covariates: 1997 Yield	With Covariates: 2004 Yield
Hybrid stayers	1.505 (0.066)	1.280 (0.056)	0.869 (0.073)	0.683 (0.063)
Leavers	0.809 (0.094)	0.648 (0.079)	0.537 (0.084)	0.370 (0.069)
Joiners	1.007 (0.114)	0.883 (0.096)	0.469 (0.101)	0.498 (0.084)
Acres (× 100)	—	—	0.561 (0.782)	−0.744 (0.802)
Seed kg per acre (× 10)	—	—	0.218 (0.035)	0.197 (0.032)
Land preparation costs per acre (× 1000)	—	—	0.066 (0.023)	0.058 (0.021)
Fertilizer per acre (× 1000)	—	—	0.063 (0.012)	0.061 (0.012)
Hired labor per acre (× 1000)	—	—	0.028 (0.008)	0.057 (0.010)
Family labor per acre (× 1000)	—	—	0.415 (0.075)	0.318 (0.064)

Heterogeneity by Observables

Variable	OLS With Covariates: Hybrid	OLS With Covariates: Nonhybrid	FE With Covariates: Hybrid	FE With Covariates: Nonhybrid
Acres ($\times$ 10)	0.144 (0.056)	−0.053 (0.149)	−0.381 (0.153)	−0.941 (0.379)
Seed kg per acre ($\times$ 10)	0.281 (0.035)	0.129 (0.036)	0.219 (0.047)	0.147 (0.063)
Land preparation costs per acre ($\times$ 1000)	0.056 (0.018)	0.135 (0.031)	0.033 (0.024)	0.097 (0.060)
Fertilizer per acre ($\times$ 1000)	0.064 (0.008)	0.143 (0.032)	0.040 (0.011)	0.081 (0.086)
Hired labor per acre ($\times$ 1000)	0.047 (0.007)	0.026 (0.010)	0.054 (0.011)	0.035 (0.029)
Family labor per acre ($\times$ 1000)	0.297 (0.064)	0.435 (0.081)	0.497 (0.094)	0.581 (0.177)
Year = 2004	0.568 (0.050)	0.595 (0.068)	0.467 (0.058)	0.689 (0.096)
Average return (β) when returns vary by observables (evaluated at mean X's)	0.480 (0.048)		0.091 (0.076)	
Number of observations	1517	887	1517	887

MITSloan MANAGEMENT

Talk Outline

- Data and Adoption Trends
- Modeling Adoption Decisions
- Estimating a Model with Heterogeneous Returns
- Results
  - Summary Statistics
  - Baseline OLS and Fixed Effects Results
  - Motivation for Heterogeneity
  - **Correlated Random Coefficients**
- Are Adoption Decisions Unconstrained? Implications for Policy
  - What Costs?
- Conclusions

MITSloan MANAGEMENT

CRC Generalizations

In CRC I can account for both exogenous and endogenous covariates:

Exogenous covariates are uncorrelated with the θ_i and are easily added to the model (enter the RFs, but not the projection)

Endogenous covariates are correlated with the θ_i and are accounted for by extending the projection to include them – I only consider fertilizer to be endogenous (clearly a joint decision for most HHs)

Interpretation

$$y_{it} = \pi + \beta h_{it} + \underbrace{\theta_i + \tau_i}_{} + \underbrace{\phi\theta_i}_{} h_{it} + \varepsilon_{it}$$

= α_i, HH Specific Intercept (HH specific average yield)

HH Specific Slope (HH specific gain to hybrid)

The covariance between the household specific slopes and intercepts:

$$cov(\alpha_i, \varphi\theta_i) = \varphi\sigma_\theta^2$$

Structural coefficient $\boldsymbol{\varphi}$ tells us if high intercept HHs are high slope HHs

If $0<\boldsymbol{\varphi}<1$, high θ_i farmers also gain the most

If $-1<\boldsymbol{\varphi}<0$, the gains are largest in the left tail of the distribution

MITSloan MANAGEMENT

OMD Structural Estimates I

	Full Sample		
	Without Covariates	With Covariates	With Interactions With Hybrid
λ_1	0.648 (0.093)	0.565 (0.087)	0.456 (0.090)
λ_2	1.007 (0.112)	0.665 (0.104)	0.473 (0.116)
λ_3	1.636 (4.854)	−1.690 (4.316)	−0.485 (0.199)
β	−0.543 (1.874)	1.023 (1.480)	3.534 (24.05)
ϕ	−0.794 (0.411)	−1.317 (1.262)	−17.82 (137.4)
χ^2_1	40.089	11.25	139.5

	Without HIV Districts		
	Without Covariates	With Covariates	With Interactions With Hybrid
λ_1	0.471 (0.099)	0.305 (0.089)	0.139 (0.092)
λ_2	1.139 (0.122)	0.710 (0.112)	0.466 (0.123)
λ_3	−4.800 (9.173)	−0.936 (0.308)	−0.497 (0.257)
β	2.287 (4.222)	0.623 (0.100)	0.790 (0.169)
ϕ	−1.010 (0.228)	−1.518 (0.310)	−2.196 (1.142)
χ^2_1	175.5	114.1	305.2

MITSloan MANAGEMENT

Comparative Advantage

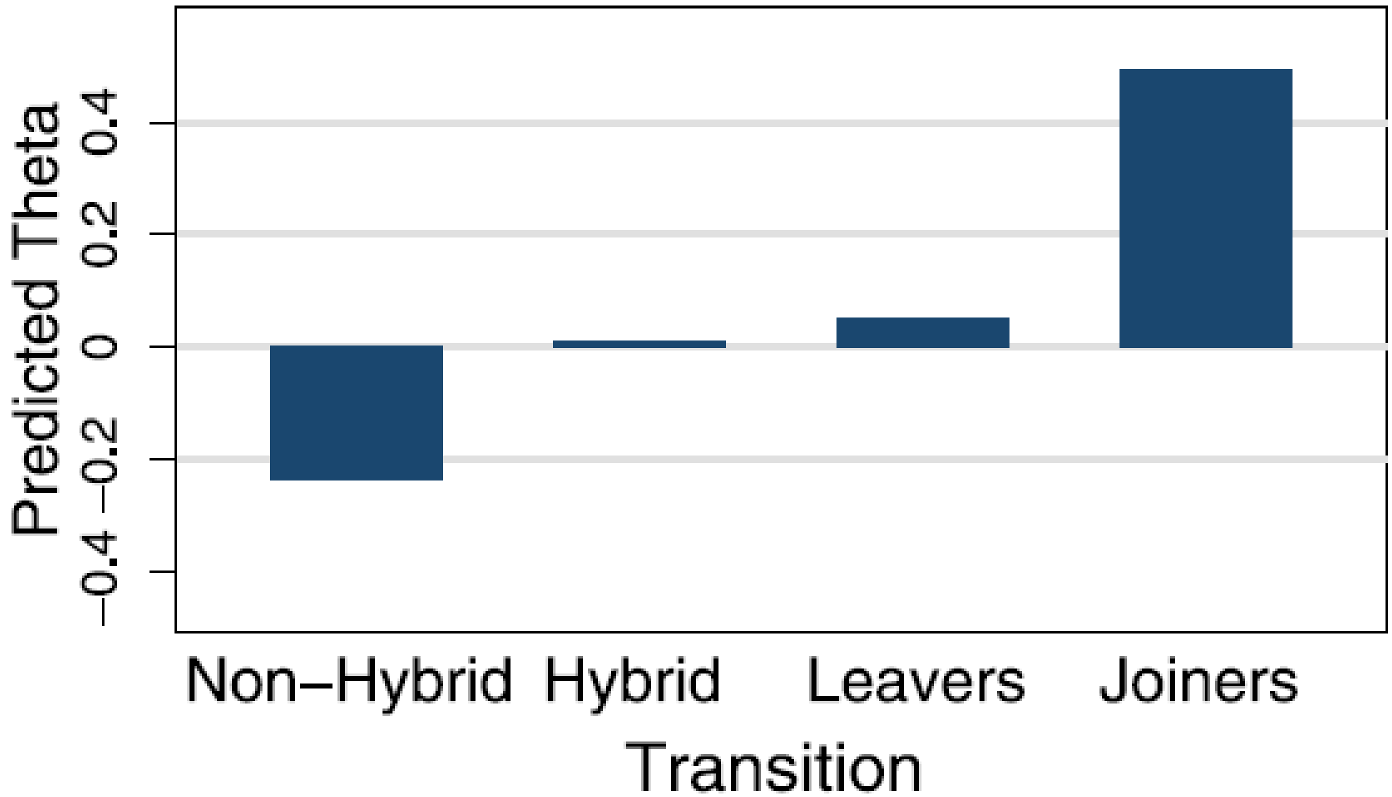


FIGURE 5A.—Distribution of comparative advantage.

Implied Returns $(\beta+\varphi\theta_i)$

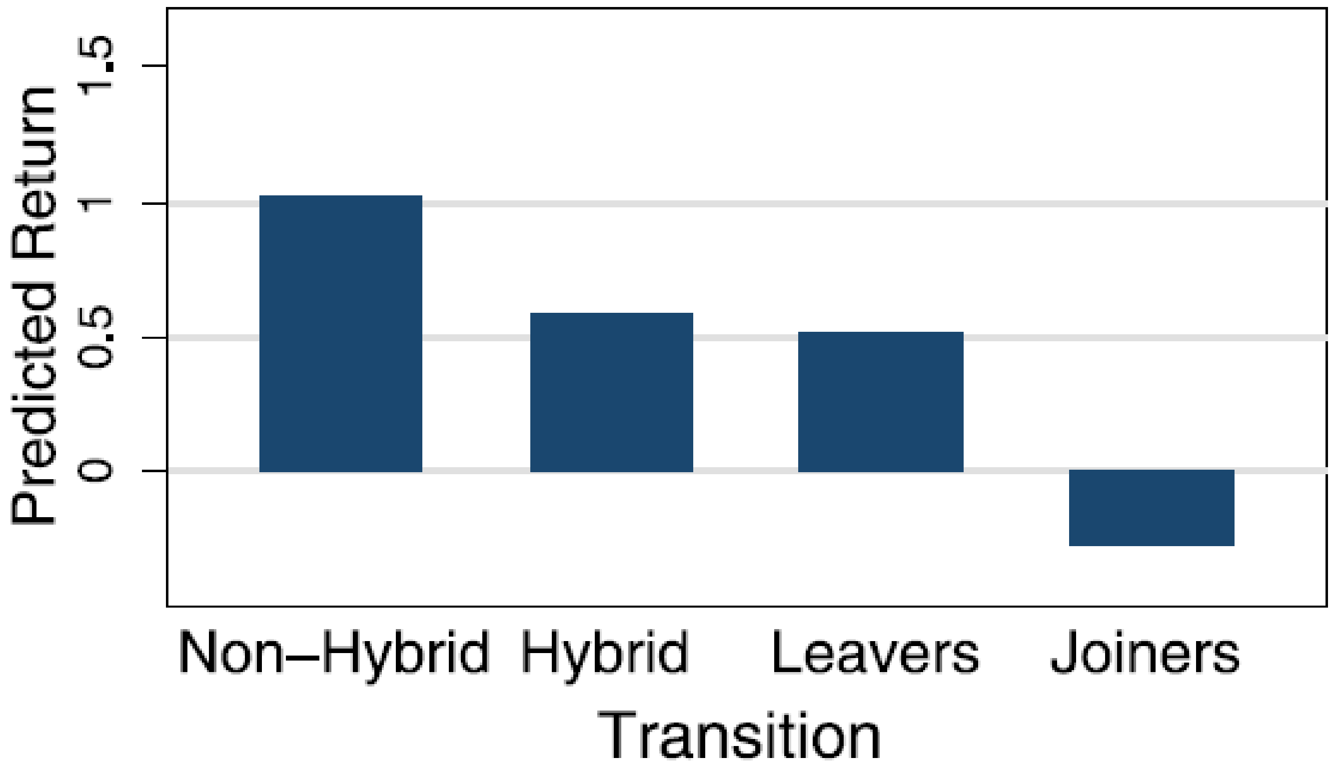


FIGURE 5B.—Distribution of returns.

MITSloan MANAGEMENT

OMD Structural Estimates II

With Both Fertilizer and Hybrid as Endogenous

Projection: $\theta_i = \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i1} h_{i2} + \lambda_4 h_{i1} f_{i1} + \lambda_5 h_{i2} f_{i1} + \lambda_6 h_{i1} h_{i2} f_{i1} + \lambda_7 h_{i1} f_{i2}$
$+ \lambda_8 h_{i2} f_{i2} + \lambda_9 h_{i1} h_{i2} f_{i2} + \lambda_{10} f_{i1} + \lambda_{11} f_{i2} + v_i$

	Full Sample		Without HIV Districts	
	With Covariates	With Interactions With Hybrid	With Covariates	With Interactions With Hybrid
β	0.088 (0.096)	0.915 (0.417)	0.603 (0.060)	0.686 (0.174)
ϕ	−0.449 (0.176)	−3.772 (2.707)	−1.788 (0.277)	−2.118 (0.641)

Comparative Advantage II

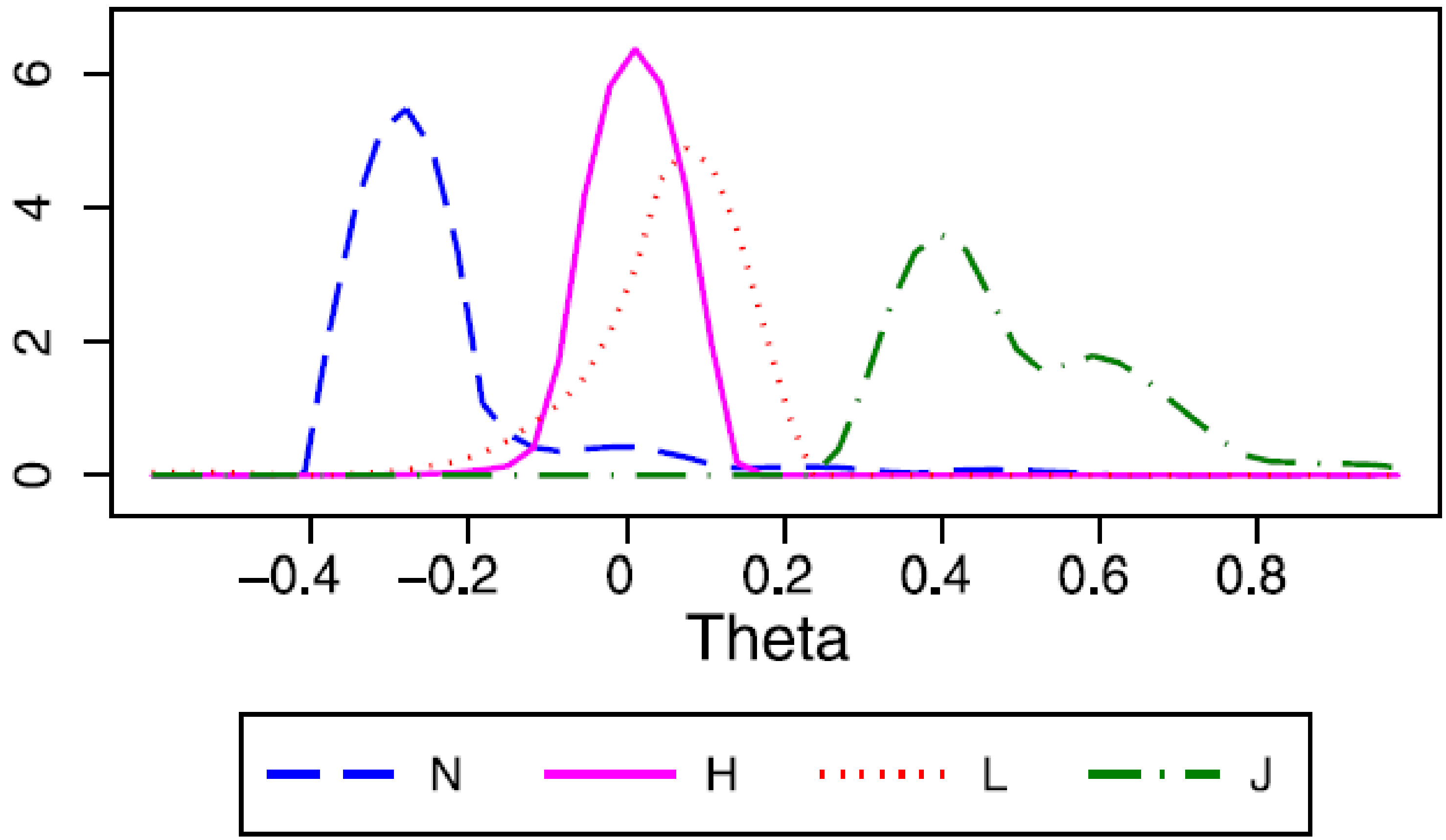


FIGURE 5C.—Endogenous hybrid and fertilizer use.

Absolute Advantage

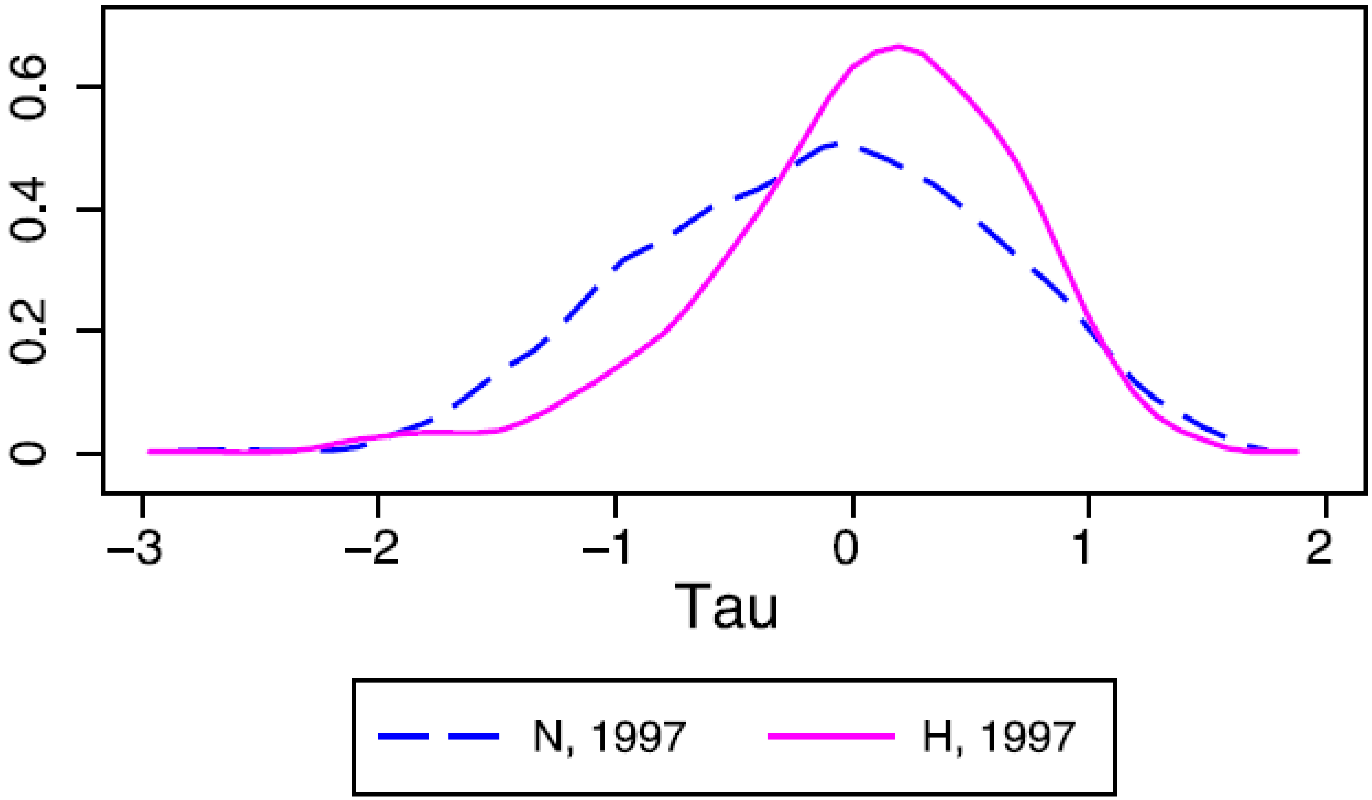


FIGURE 5D.—Distribution of tau, by adoption decision in 1997.

Talk Outline

- Data and Adoption Trends
- Modeling Adoption Decisions
- Estimating a Model with Heterogeneous Returns
- Results
  - Summary Statistics
  - Baseline OLS and Fixed Effects Results
  - Motivation for Heterogeneity
  - Correlated Random Coefficients Estimates
- Robustness Checks
- **Are Adoption Decisions Unconstrained? Policy Implications**
  - **What Costs?**
- Conclusions

MITSloan MANAGEMENT

Is Adoption Constrained?

There is a group of farmers who would have high returns but do not use hybrid at all

LATE estimates point to these farmers facing larger costs/constraints, in particular infrastructure/input distribution

LATE estimates close to the gains for these farmers from hybrid

A larger fraction of the sample has small positive returns and adopts universally; these farmers seem to be unconstrained

For these farmers, the risk properties of hybrid may be important

MITSloan MANAGEMENT

What Costs or Constraints?

	(1)	(2)	(3)	(4)
Distance to closest fertilizer seller (× 100)	−0.301 (0.122)	−0.289 (0.121)	−0.285 (0.121)	−0.285 (0.121)
Distance to motorable road (× 100)	−0.904 (0.503)	−0.887 (0.501)	−0.901 (0.502)	−0.898 (0.501)
Distance to matatu stop (× 100)	0.032 (0.298)	−0.034 (0.298)	−0.016 (0.298)	−0.028 (0.299)
Distance to extension services (× 100)	−0.130 (0.155)	−0.063 (0.155)	−0.063 (0.155)	−0.061 (0.155)
Tried to get credit (× 10)	—	−0.138 (0.153)	—	—
Tried but did not receive credit (× 10)	—	—	0.027 (0.347)	—
Received credit (× 10)	—	—	—	−0.047 (0.154)
Dummies for household head education	No	Yes	Yes	Yes
(*p*-value on joint significance)		(0.002)	(0.002)	(0.000)
Province dummies	Yes	Yes	Yes	Yes

Alternative Models: Learning

No S-shaped curves of adoption (aggregate adoption is constant over the sample period)

Switching between different types of hybrid? 60-70% plots use 614 in both 2000 and 2004, about 5% of each of 625, 627 and 511

99% of households have used hybrid before (83% fertilizer)

Earlier table on switching didn't show systematic persistence in adoption

Only 0.3% of households using traditional varieties cited things like "experimenting" or "on trial" when asked why they were not using hybrid

Switching

TRANSITIONS ACROSS HYBRID/NONHYBRID SECTORS FOR THE SAMPLE PERIODS 1997, 2000, AND 2004[a]

Transition in Terms of Technology Used (1997 2000 2004)			Fraction of Sample (%) ($N = 1202$ Households)
N	N	N	20.38
N	N	H	2.83
H	N	H	6.07
N	H	H	4.91
H	N	N	5.99
H	N	H	3.16
H	H	N	7.15
H	H	H	49.50

MITSloan MANAGEMENT

Alternative Models: Credit

How About Credit Constraints?

In 2004, 41% of households tried to get credit (83% of which did)

In 1997 34% of households tried to get agricultural credit (90% of which did)

Only 2% of households plant both hybrid and non-hybrid, there are no differential fixed costs to planting hybrid

Earlier patterns did not fit a pure liquidity constraints story

None of the credit variables correlated with comparative advantage

Selection and Heterogeneity

	Without Covariates		With Covariates	
Variable	1997 Yield	2004 Yield	1997 Yield	2004 Yield
Hybrid stayers	1.505 (0.066)	1.280 (0.056)	0.869 (0.073)	0.683 (0.063)
Leavers	0.809 (0.094)	0.648 (0.079)	0.537 (0.084)	0.370 (0.069)
Joiners	1.007 (0.114)	0.883 (0.096)	0.469 (0.101)	0.498 (0.084)
Acres (× 100)	—	—	0.561 (0.782)	−0.744 (0.802)
Seed kg per acre (× 10)	—	—	0.218 (0.035)	0.197 (0.032)
Land preparation costs per acre (× 1000)	—	—	0.066 (0.023)	0.058 (0.021)
Fertilizer per acre (× 1000)	—	—	0.063 (0.012)	0.061 (0.012)
Hired labor per acre (× 1000)	—	—	0.028 (0.008)	0.057 (0.010)
Family labor per acre (× 1000)	—	—	0.415 (0.075)	0.318 (0.064)

MITSloan MANAGEMENT

Conclusion

I lay out a framework to capture heterogeneity in benefits to hybrid

With no systematic pattern in temporal variation in adoption over my sample period, I focus on the spatial variation in adoption to understand the heterogeneity in returns

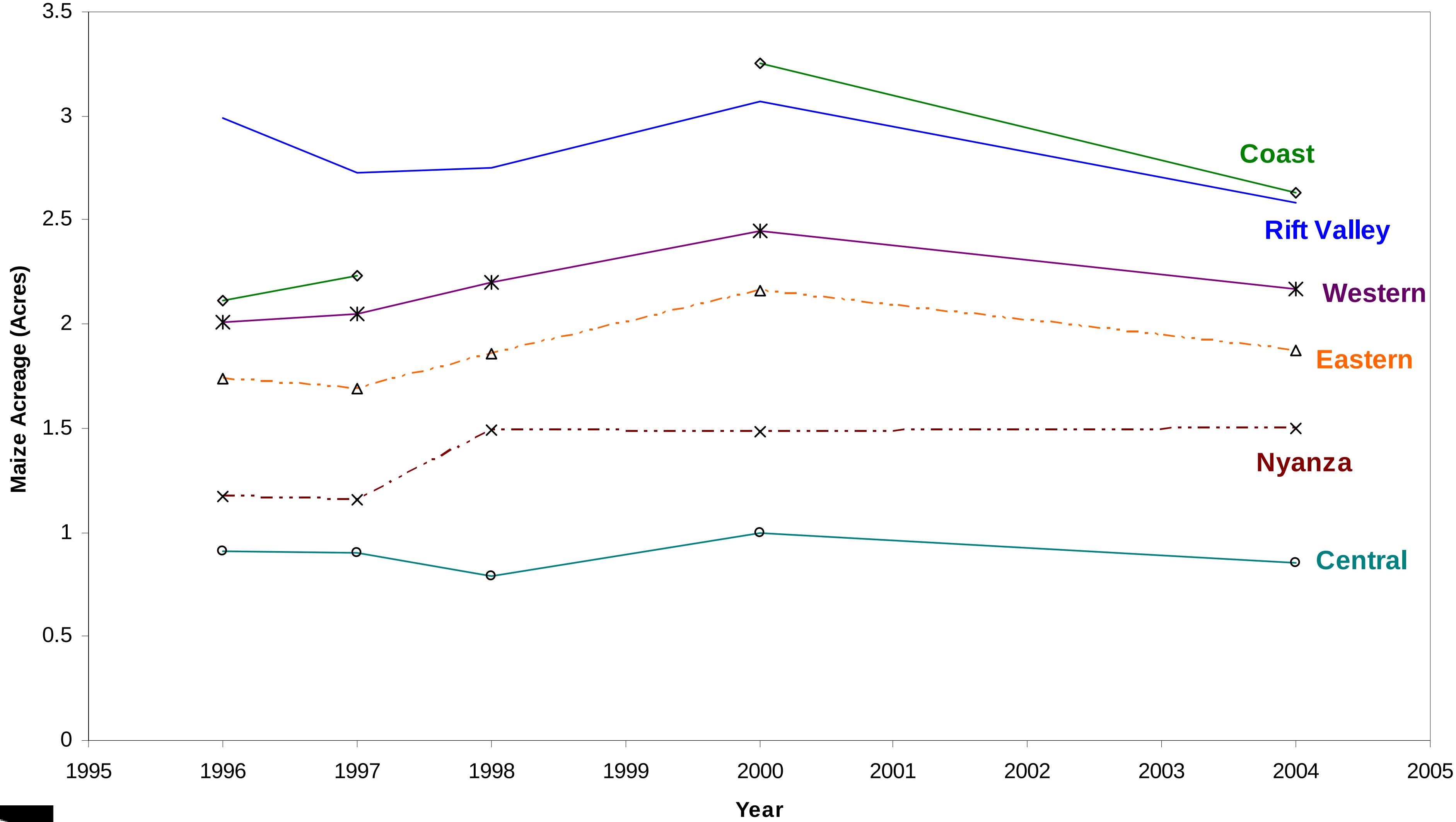
Policy implications are different for the different groups of farmers:

- For those with high returns, they do not use hybrid due to constraints, policy is to alleviate constraints
- For the always adopters, maize supply could be increased via the development of new hybrid strains

MITSloan MANAGEMENT

EXTRA SLIDES

Maize Acreage



Maize Harvests