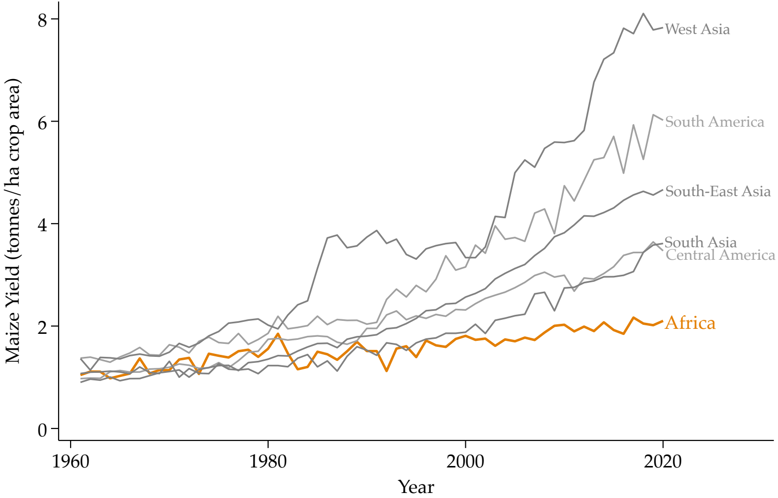
Liquidity Constraints and Capital Allocation:

Evidence from a Selective Trial with Ugandan Farmers

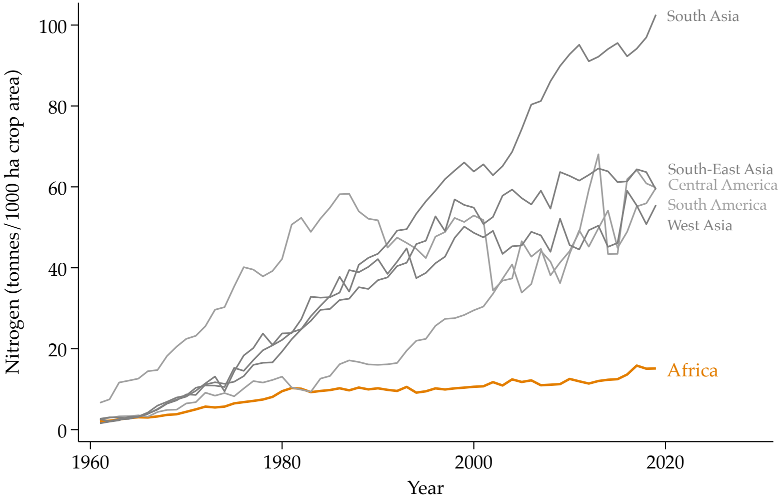
K. Burchardi, J. De Quidt, B. Lerva, S. Tripodi

7th of November 2023

Background: Agricultural Output in Africa



Background: Fertilizer Adoption in Africa



See also: Gollin, Lagakos and Waugh (2014) → Phosphate → Potash

Today

Our Paper

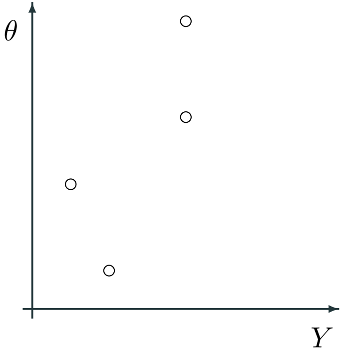
- Functioning of market mechanism in allocating fertilizer.
  - What is the average return of non-adopters?
    $\rightarrow$ Too low to hand out free.
  - What is the optimal subsidy?
    $\rightarrow$ About 30% of the market price.
- Why does market mechanism not work efficiently?
  $\rightarrow$ Partly because of liquidity constraints.

Theory: Selection and Misallocation in Markets

Consider an *investment good* yielding a heterogeneous return.

How to ensure the investment good is allocated to individuals with highest return?

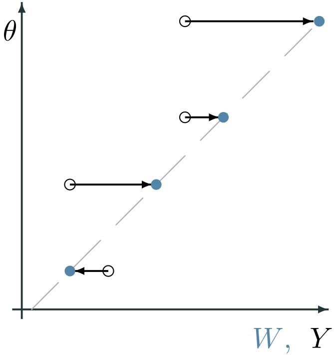
Theory: Allocative Power of Markets



- Consider an *investment good* yielding heterogeneous return θ.
- Access to liquid wealth Y.

What is the investor's willingness-to-pay, W?

Theory: Allocative Power of Markets

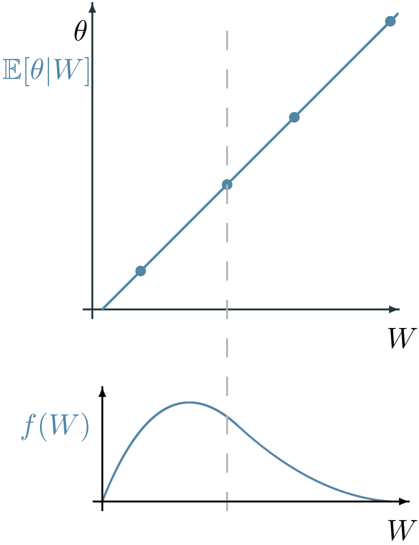


- Consider an *investment good* yielding heterogeneous return θ.
- Access to liquid wealth Y.

What is the investor's willingness-to-pay, W?

$\rightarrow$ Without liquidity constraints: $W = \theta$, independent of Y.

Theory: Allocative Power of Markets



- Consider an *investment good* yielding heterogeneous return θ.
- Access to liquid wealth Y.

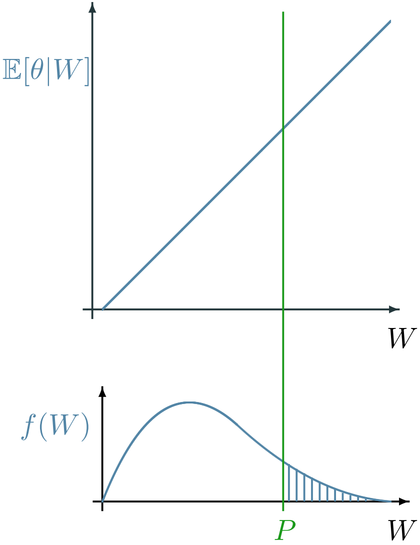
What is the investor's willingness-to-pay, W?

$\rightarrow$ Without liquidity constraints: $W = \theta$, independent of Y.

Implications:

$\rightarrow$ $\mathbb{E}[\theta|W]$ is 45 degree line.

Theory: Allocative Power of Markets



- Consider an *investment good* yielding heterogeneous return θ.
- Access to liquid wealth Y.

What is the investor's willingness-to-pay, W?

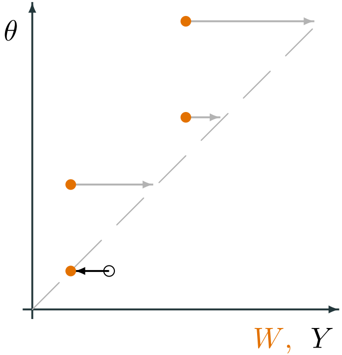
$\rightarrow$ Without liquidity constraints: $W = \theta$, independent of Y.

Implications:

$\rightarrow$ $\mathbb{E}[\theta|W]$ is 45 degree line.

$\rightarrow$ **Selection effect of markets:** charging price P, projects with $W \geq P$ receive investment good: $W \geq P$ if and only if $\theta \geq P$!

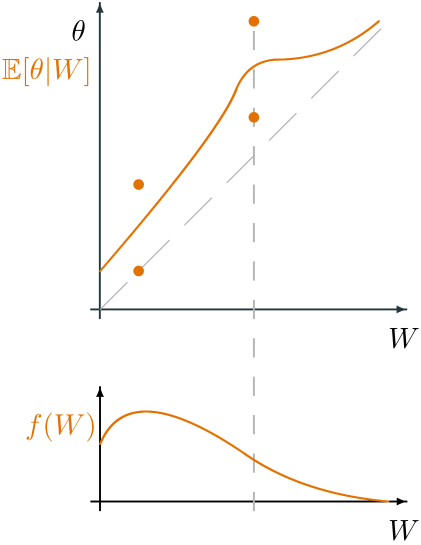
Theory: Misallocation in Markets



With liquidity constraints:

- $W = \min\{Y, \theta\}$.

Theory: Misallocation in Markets



With liquidity constraints:

- $W = \min\{Y, \theta\}$.

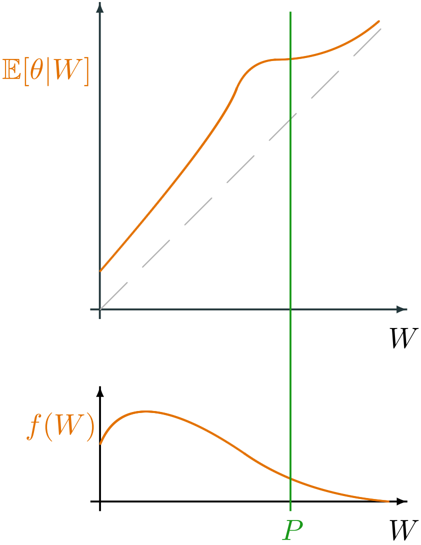
Implications:

$\rightarrow$ $\mathbb{E}[\theta|W] \geq W$.

$\rightarrow$ Demand is lower.

$\rightarrow$ Selection effect breaks down.

Theory: Misallocation in Markets



With liquidity constraints:

- $W = \min\{Y, \theta\}$.

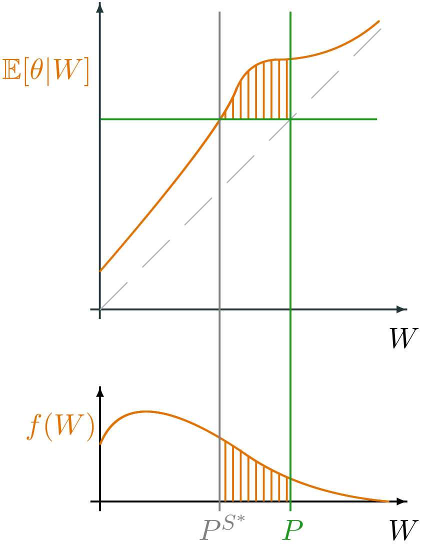
Implications:

→ $\mathbb{E}[\theta|W] \geq W$.

→ Demand is lower.

→ Selection effect breaks down.

Theory: Misallocation in Markets



With liquidity constraints:

- $W = \min\{Y, \theta\}$.

Implications:

→ $\mathbb{E}[\theta|W] \geq W$.

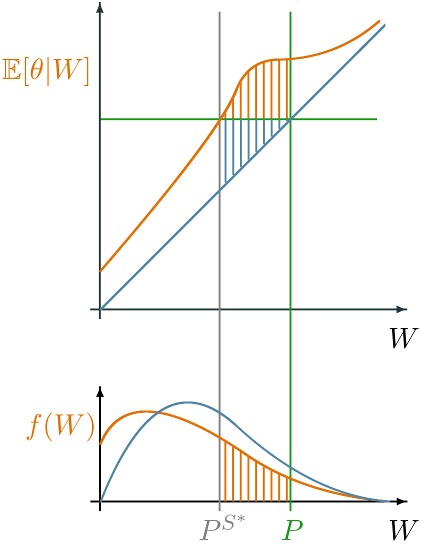
→ Demand is lower.

→ Selection effect breaks down.

Subsidy might be welfare-increasing.

- Optimal subsidy $P - P^{S^*}$.
- P^{S^*} might be 0.

Theory: Misallocation in Markets



With liquidity constraints:

- $W = \min\{Y, \theta\}$.

Implications:

→ $\mathbb{E}[\theta|W] \geq W$.

→ Demand is lower.

→ Selection effect breaks down.

Subsidy might be welfare-increasing.

- Optimal subsidy $P - P^{S^*}$.
- P^{S^*} might be 0.

Without frictions: no point in subsidy.

Related Work

- Fertiliser adoption rates are low.
- Measures of labour productivity and capital intensity are low (Gollin, Lagakos, Waugh, 2014).
- Duflo, Kremer, Robinson (2008, 2011): maize farmers' returns to fertilizer usage around 36%, but takeup very low.
- Methodologically related to Berkower and Dean (2022).

Theory to Empirics

Experimental Design: Measure W

Multiple Price List

Possible R	**Purchase**
200k UGX	
.	.
.	.
140k UGX	
120k UGX	
100k UGX	
.	.
.	.
0k UGX	

Experimental Design: Measure W

Multiple Price List

Possible R	**Purchase**
200k UGX	NO
.	.
.	.
140k UGX	NO
120k UGX	YES
100k UGX	YES
.	.
.	.
0k UGX	YES

Experimental Design: Measure W

Multiple Price List

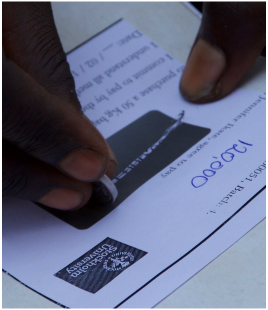
Possible R		**Purchase**
200k UGX		NO
.		.
.		.
140k UGX		NO
$\boxed{\text{120k UGX}}$	= W	YES
100k UGX		YES
.		.
.		.
0k UGX		YES

Experimental Design: Measure W

Multiple Price List

Possible R		Purchase
200k UGX		NO
.		.
.		.
140k UGX		NO
$\boxed{\text{120k UGX}}$	$= W$	YES
100k UGX		YES
.		.
.		.
0k UGX		YES

Scratchcard with R



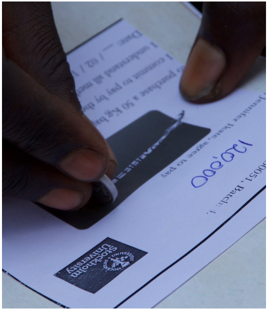
Iff $R \leq W$, purchase fertilizer at R.

Experimental Design: Measure W

Multiple Price List

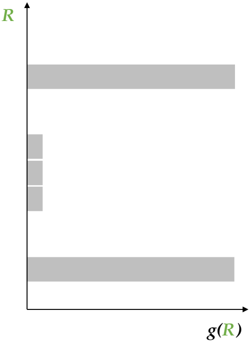
Possible R	**Purchase**
200k UGX	NO
.	.
.	.
140k UGX	NO
$\boxed{\text{120k UGX}}$ = W	YES
100k UGX	YES
.	.
.	.
0k UGX	YES

Scratchcard with R



Iff **R** ≤ W, purchase fertilizer at **R**.

Distribution of R



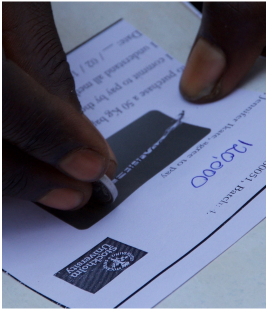
- Version of Becker-DeGroot-Marschak (BDM, 1964).
- Incentive-compatible mechanism to elicit W.

Experimental Design: Identify $\mathbb{E}[\theta|W]$

Multiple Price List

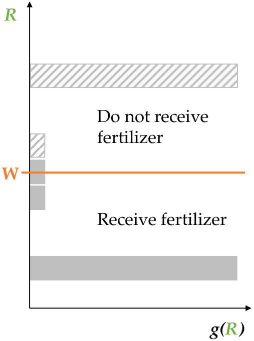
Possible *R*	**Purchase**
200k UGX	NO
.	.
.	.
140k UGX	NO
$\boxed{\text{120k UGX}}$ = W	YES
100k UGX	YES
.	.
.	.
0k UGX	YES

Scratchcard with *R*



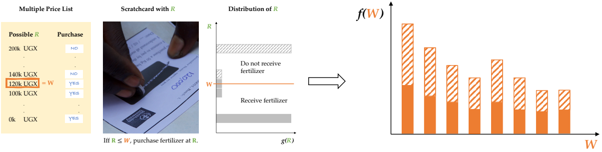
Iff $R \leq W$, purchase fertilizer at R.

Distribution of *R*



- Randomization of treatment conditional on W: identify $\mathbb{E}[\theta|W]$.
- BDM is a *selective trial* (Chassang et al., 2012). [Details]
- Related work: Berry et al. (2019), Berkouwer and Dean (2022).

Experimental Design: Relieving Liquidity Constraints



Experimental Design: Relieving Liquidity Constraints

- Lottery prize money equal to market price of the investment good.
- For $W \leq P$ we know that $Y \geq W$, i.e. liquidity constraints relieved!
- Also incentivise attendance.

Setting and Sampling

Setting: Adoption of **chemical fertiliser** (DAP/CAN) by maize farmers in Eastern Ugandan districts of Tororo, Manafa, and Mbale (Duflo et al., 2008).

Product: 50kg DAP + 50kg CAN; 1 acre; P = 200k UGX.

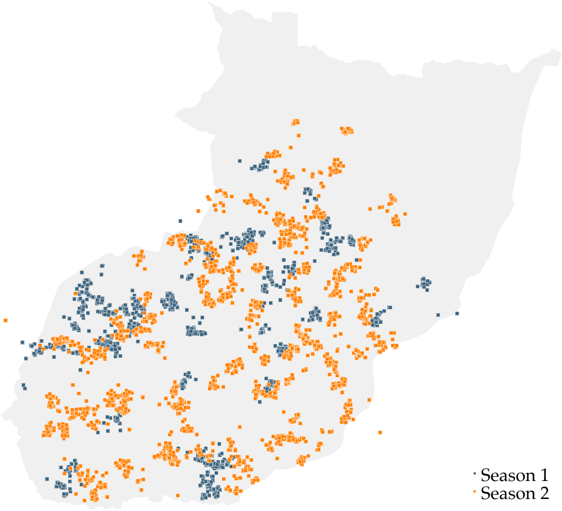
Timing: 1st season of 2017 (S1) and 2018 (S2).

Setting and Sampling

Sampling:

1. Sample of 51/102 rural villages (less than 100 households/km^2), stratified by pop. density.
2. Conducted *census* of households.
3. Eligible households (i) engaged in commercial farming, (ii) cultivate maize, (iii) self-report cultivating 2 to 6 acres and (iv) have mobile money account.
4. Select 408/816 eligible households.

Study Sample



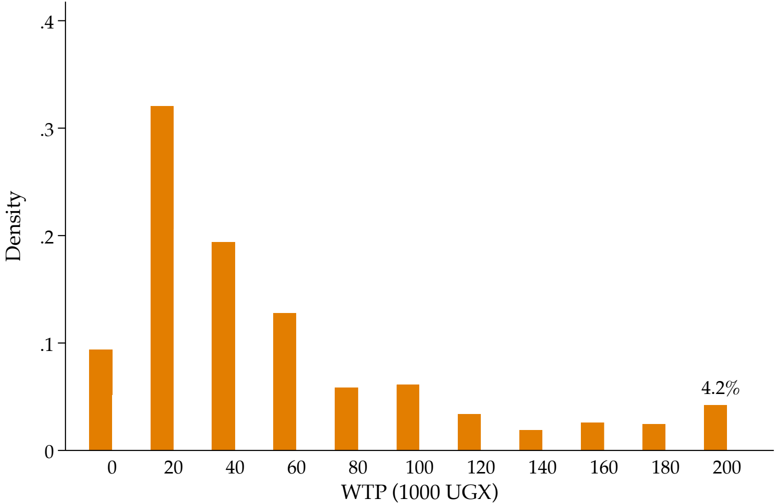
Survey Work: Season 1

1. **Census:** 1/12-16/12/2016.
2. **Baseline:** 24/1-4/2/2017.
3. **Lottery Payout:** 5 days before WTP. [WTP & Planting Time]
4. **WTP Elicitation:** 10/2-13/2/2017. [Implementation]
5. **Phone Survey 1:** 3/4-12/4/2017.
6. **Crop Survey:** 5/6-28/6/2017. [Plot Size Measurement] [Quadrants]
7. **Cob Survey:** 2/7-21/7/2017. [Optimal Survey Design]
8. **Follow-Up Survey:** 21/8-30/8/2017.
9. **Phone Survey 2:** 27/11-30/11/2017.

[Timeline: Season 2]

Results:
Lottery Effect on WTP

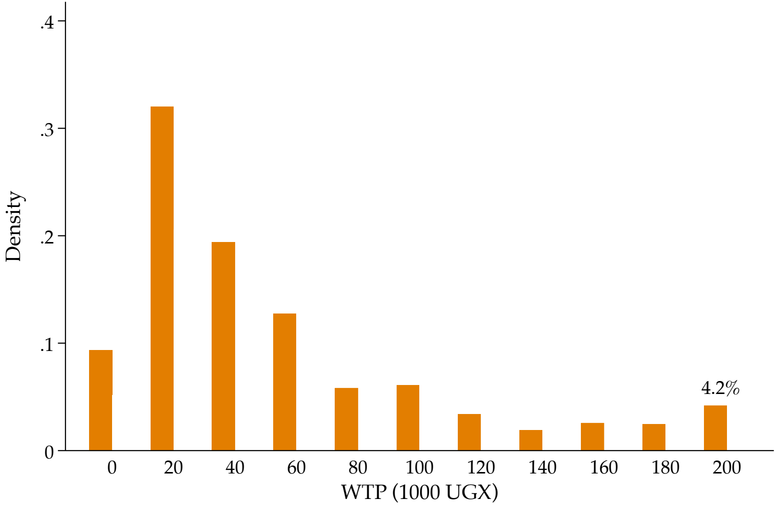
Willingness To Pay (Lottery: 5k)



$\mathbb{E}[\text{WTP}|\text{Lottery Lost}] = 56k\,\text{UGX}$

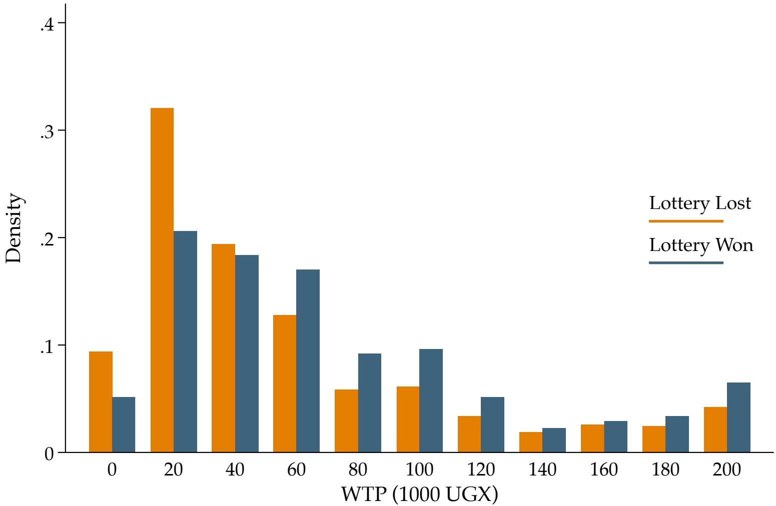
Willingness To Pay (Lottery: 5k)

Baseline: 9% use chemical fertilizer, 4% DAP and 2% CAN.



$\mathbb{E}[\text{WTP}|\text{Lottery Lost}] = 56k\,\text{UGX}$

Willingness To Pay



$\mathbb{E}[\text{WTP}|\text{Lottery Lost}] = 56k\,\text{UGX}$; $\mathbb{E}[\text{WTP}|\text{Lottery Won}] = 72k\,\text{UGX}$

Results:
ATE of Fertilizer

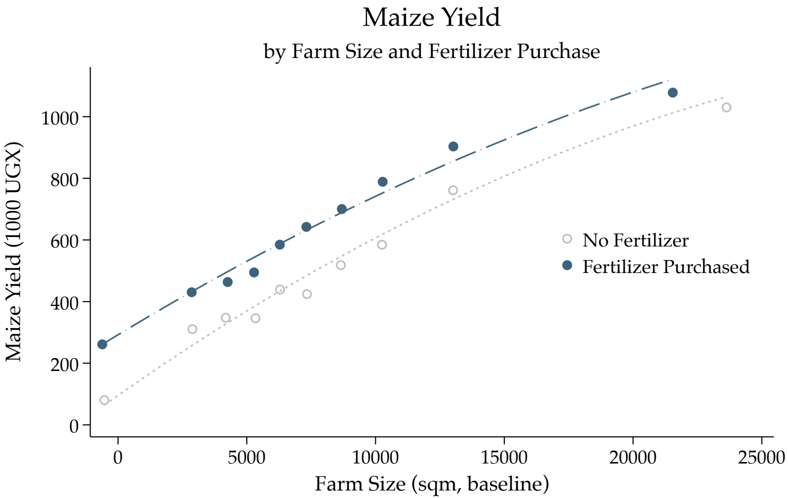
Estimation Approach

Estimating Equation:

$$
\begin{aligned}
y_{hv} = \quad & \sum_w \alpha_w \mathbf{1}[W_h = w, L_h = 0] + \sum_w \beta_w \mathbf{1}[P_h \leq W_h,\, W_h = w, L_h = 0] \\
+ \quad & \sum_w \gamma_w \mathbf{1}[W_h = w, L_h = 1] + \sum_w \delta_w \mathbf{1}[P_h \leq W_h,\, W_h = w, L_h = 1] \\
+ \quad & \zeta X_{hv} + \epsilon_{vh},
\end{aligned}
$$

- where β_w, γ_w, and δ_w are coefficients of interest,
- X_{hv} always includes village fixed effects, baseline farm size and baseline maize output,
- the sample excludes observations with $0 < P_h \leq W_h$.

Maize Yield and Farm Size



Estimation Approach

Estimating Equation:

$$\begin{aligned} y_{hv} = & \sum_w \alpha_w \mathbf{1}[W_h = w, L_h = 0] + \sum_w \beta_w \mathbf{1}[P_h \leq W_h, W_h = w, L_h = 0] \\ + & \sum_w \gamma_w \mathbf{1}[W_h = w, L_h = 1] + \sum_w \delta_w \mathbf{1}[P_h \leq W_h, W_h = w, L_h = 1] \\ + & \zeta X_{hv} + \epsilon_{vh} \end{aligned}$$

ATE of fertilizer on y_{hv} in the $L_h = 0$ group is calculated as:

$$\text{ATE}_F(L_h = 0) = \sum_w \beta_w f_{L=0}(w),$$

where $f_{L=0}(w)$ is the distribution of WTP in the $L_h = 0$ group.

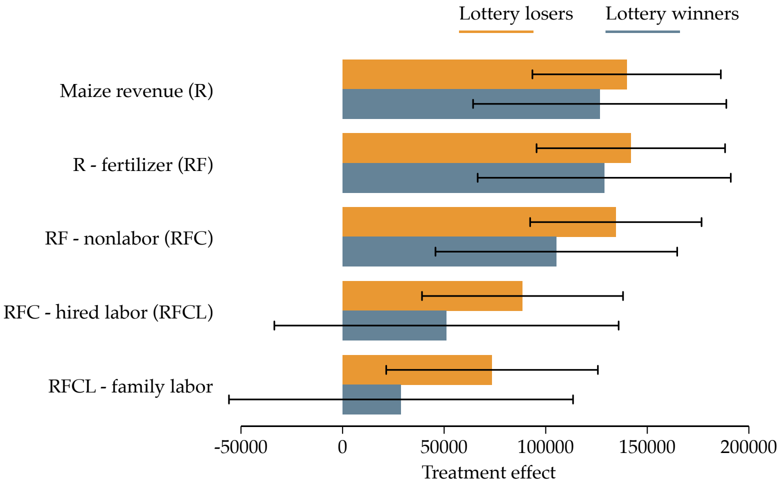
Similarly for $\text{ATE}_F(L_h = 1)$ and ATE of lottery treatment.

Effects on Plant Growth

TABLE 11: QUADRANT LEVEL EFFECTS

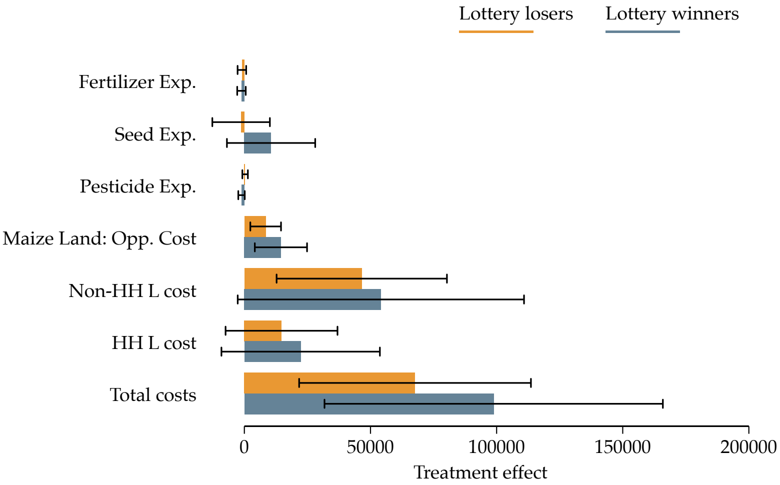
	Maize: # (1)	Expectations (2)	Maize: Exp. # (3)	D: Intercropped (4)	D: Line Sowing (5)
Fertilizer (Lottery Lost)	0.477	0.074	0.815	-0.107	0.041
	(0.134)	(0.018)	(0.149)	(0.027)	(0.018)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.023]
Fertilizer (Lottery Won)	0.412	0.085	0.713	-0.184	0.035
	(0.222)	(0.023)	(0.235)	(0.037)	(0.025)
	[0.064]	[0.000]	[0.003]	[0.000]	[0.159]
Lottery Won	0.321	0.037	0.362	0.007	0.062
	(0.171)	(0.020)	(0.169)	(0.031)	(0.022)
	[0.060]	[0.066]	[0.032]	[0.819]	[0.005]
N (Lottery Lost/Won)	4880/3080	4880/3080	4880/3080	4880/3080	4880/3080
Mean Y in Control	5.206	0.615	3.415	0.440	0.898

Effects on Profit Measures



→ Weather → 'First Stage' → Table

Effects on Costs



Discussion

- **Duflo, Kremer, Robinson (2008)** find – at best – 136% gross return over the season (CAN only).
- With comparable output measure, we find 60%-70% return.
- With appropriate profit measure, we find 15%-30% return.
- All calculations do not include fertiliser cost!

Results:
ATE of Lottery

Lottery Effects

Table 1: Profit measures

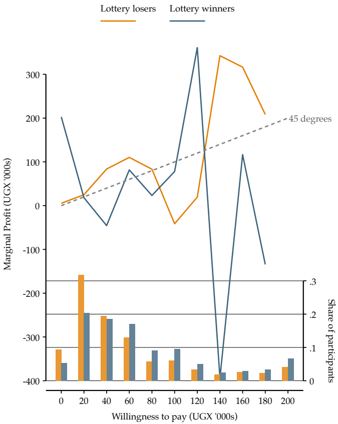
	Maize revenue (R)	R - fertilizer (RF)	RF - nonlabor (RFC)	RFC - hired labor (RFCL)	RFCL - family labor
Fertilizer (Lottery Lost)	138528 (28474) [0.000]	139905 (28482) [0.000]	136983 (26513) [0.000]	92349 (31159) [0.003]	51865 (39690) [0.192]
Fertilizer (Lottery Won)	120132 (38567) [0.002]	122227 (38576) [0.002]	102971 (37947) [0.007]	54140 (54327) [0.319]	30212 (62465) [0.629]
Lottery Won	**76731** (33198) [0.021]	**75457** (33217) [0.023]	**62743** (32040) [0.051]	**64352** (37824) [0.089]	**79023** (48729) [0.105]
N (Lottery Lost/Won)	684/412	684/412	680/408	680/408	680/408
Mean Y in Control	336805	336726	235620	103185	-236670

Karlan, Osei, Osei-Akoto and Udry (QJE, 2014):
420$ grant in Ghana increases costs by 2$ and profits by 65$.

Beaman, Karlan, Thuysbaert, Udry (ECMA, 2023):
140$ grant in Mali increases output by 66$ and “gross profits” by 39$.
(Does not subtract the value of own, family and unpaid labor.)

Results:
Marginal Treatment Effects

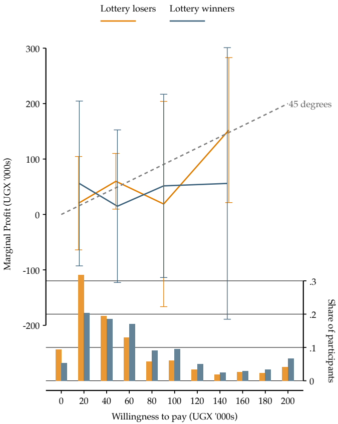
Marginal Treatment Effects on Profit



Market Allocation:

- For $W \geq 140k$ UGX find large returns in constrained group.

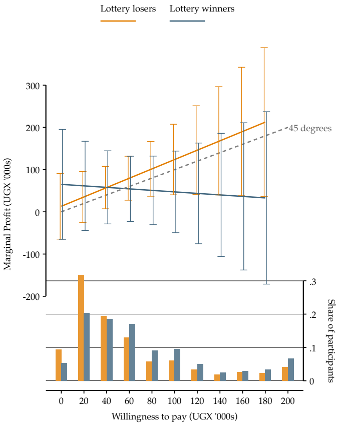
Marginal Treatment Effects on Profit



Market Allocation:

- For $W \geq 140k$ UGX find large returns in constrained group.

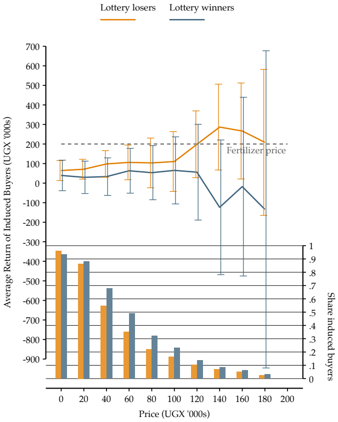
Marginal Treatment Effects on Profit



Market Allocation:

- For $W \geq 140k$ UGX find large returns in constrained group.

Average Treatment Effects of Induced Buyers



Market Allocation:

- For $W \geq 140k$ UGX find large returns in constrained group.

Subsidy:

- Large positive returns realised with optimal subsidy of 60k UGX.
- 7% induced to buy.
- But: optimal subsidy is 0 in unconstrained group!

Results:
Long-Run Adoption

Long-Run Adoption

Table 4: Long run adoption

	D: FER (P) (1)	D: DAP (P) (2)	D: CAN (P) (3)	DAP left (F) (4)	CAN left (F) (5)	Exp. FER (P) (6)	Exp. DAP (P) (7)	Exp. CAN (P) (8)
Fertilizer (Lottery Lost)	0.300 (0.035) [0.000]	0.306 (0.032) [0.000]	0.290 (0.030) [0.000]	0.233 (0.026) [0.000]	0.239 (0.026) [0.000]	3332 (2511) [0.185]	2479 (1215) [0.042]	1806 (1087) [0.097]
Fertilizer (Lottery Won)	0.409 (0.049) [0.000]	0.405 (0.046) [0.000]	0.375 (0.042) [0.000]	0.192 (0.030) [0.000]	0.242 (0.037) [0.000]	8760 (3848) [0.023]	5007 (2466) [0.043]	2484 (1344) [0.065]
Lottery Won	0.036 (0.036) [0.308]	0.016 (0.030) [0.600]	0.006 (0.027) [0.818]	0.005 (0.015) [0.752]	-0.006 (0.018) [0.759]	-2185 (2880) [0.448]	-1910 (1600) [0.233]	-421 (1088) [0.699]
N (Lottery Lost/Won)	642/388	642/388	642/388	642/388	642/388	642/388	642/388	642/388
Mean Y in Control	0.105	0.058	0.035	-0.009	-0.005	6160	2479	834

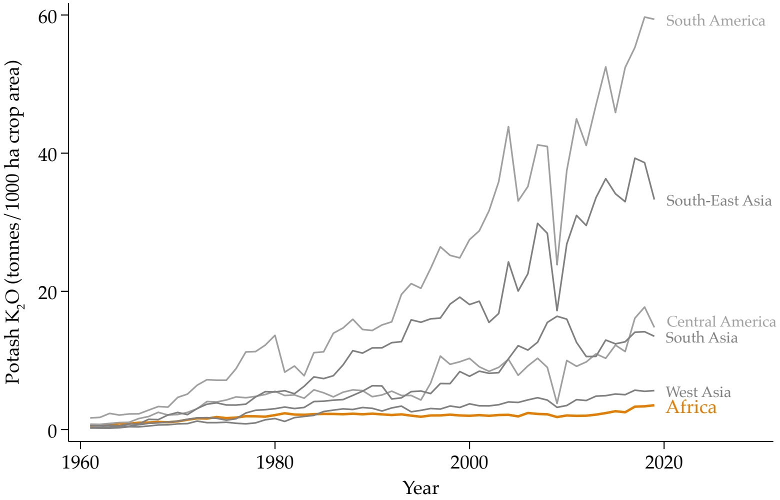
- Very large effects on long-run adoption.
- Not driven by income effect of fertilizer; suggests learning.

Our Paper

- Functioning of market mechanism in allocating fertilizer.
  - What is the average return of non-adopters?
    $\rightarrow$ Too low to hand out free.
  - What is the optimal subsidy?
    $\rightarrow$ About 30% of the market price.
- Why does market mechanism not work efficiently?
  $\rightarrow$ Partly because of liquidity constraints.

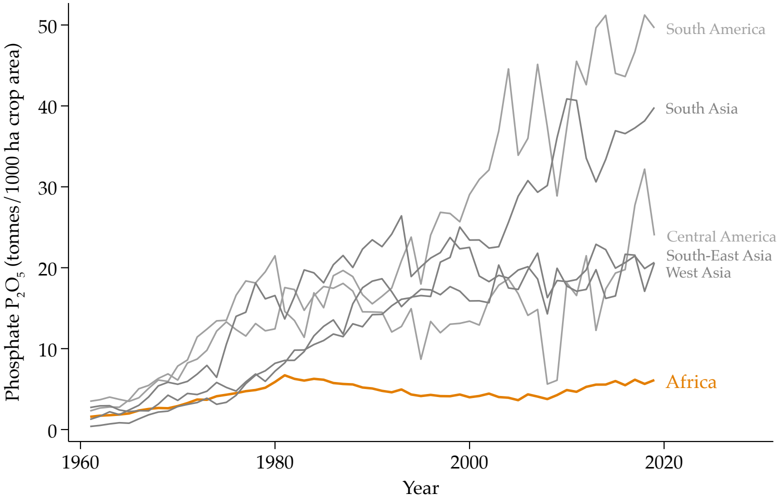
Appendix

Background: Fertilizer Adoption in Africa



←

Background: Fertilizer Adoption in Africa



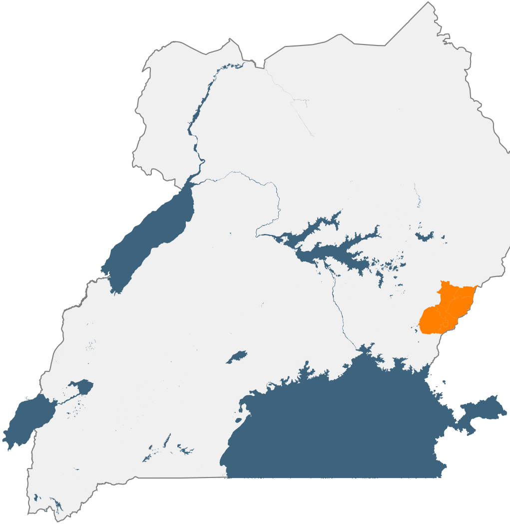
Lottery Effects

Table 1: Profit measures

	Maize revenue (R)	R - fertilizer (RF)	RF - nonlabor (RFC)	RFC - hired labor (RFCL)	RFCL - family labor
Fertilizer (Lottery Lost)	138528 (28474) [0.000]	139905 (28482) [0.000]	136983 (26513) [0.000]	92349 (31159) [0.003]	51865 (39690) [0.192]
Fertilizer (Lottery Won)	120132 (38567) [0.002]	122227 (38576) [0.002]	102971 (37947) [0.007]	54140 (54327) [0.319]	30212 (62465) [0.629]
Lottery Won	**76731** (33198) [0.021]	**75457** (33217) [0.023]	**62743** (32040) [0.051]	**64352** (37824) [0.089]	**79023** (48729) [0.105]
N (Lottery Lost/Won)	684/412	684/412	680/408	680/408	680/408
Mean Y in Control	336805	336726	235620	103185	-236670

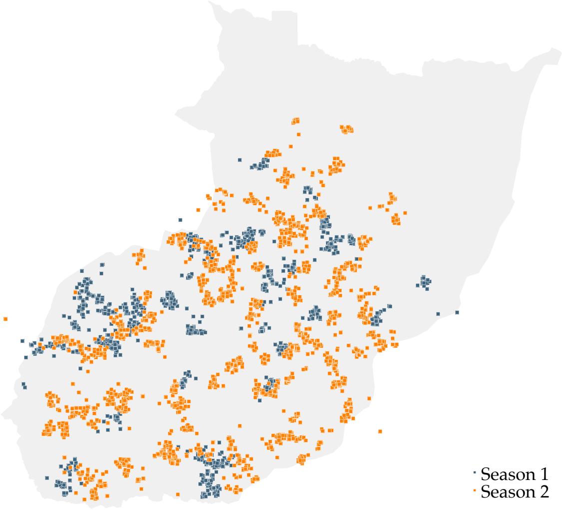
←

Tororo, Mbale, Manafa



←

Study Sample



Marginal Treatment Effects

1. We can identify, for each w

$$\Delta^{MTE}(w) = E[y|\tau = 1, W = w] - E[y|\tau = 0, W = w]$$

Marginal Treatment Effects

1. We can identify, for each w

$$\Delta^{MTE}(w) = E[y|\tau = 1, W = w] - E[y|\tau = 0, W = w]$$

2. Can recover conventional average treatment effect:

$$\int_{W_{min}}^{W_{max}} \Delta^{MTE}(w) dF(w) = E[y|\tau = 1] - E[y|\tau = 0] = ATE$$

Marginal Treatment Effects

1. We can identify, for each w

$$\Delta^{MTE}(w) = E[y|\tau = 1, W = w] - E[y|\tau = 0, W = w]$$

2. Can recover conventional average treatment effect:

$$\int_{W_{min}}^{W_{max}} \Delta^{MTE}(w) dF(w) = E[y|\tau = 1] - E[y|\tau = 0] = ATE$$

3. Can recover other quantities of interest, e.g.
   - $\int_{P}^{W_{max}} \Delta^{MTE}(w) dF(w)$: ATE for buyers at price P.
   - $\int_{P}^{P'} \Delta^{MTE}(w) dF(w)$: LATE of price change from P' to P.

[← Experimental Design]

Implementation Challenges: Understanding

Berry et al. (2015) use approach with water filters in Ghana. BDM demand curve similar to take-it-or-leave-it offers.

Implementation Issues:

- People do not always bid truthfully in strategy-proof mechanisms like BDM: more dominant strategy play in ascending than second-price sealed-bid auctions.

Implementation Challenges: Understanding

Berry et al. (2015) use approach with water filters in Ghana. BDM demand curve similar to take-it-or-leave-it offers.

Implementation Issues:

- People do not always bid truthfully in strategy-proof mechanisms like BDM: more dominant strategy play in ascending than second-price sealed-bid auctions.
- Worry about auctioneer misbehaviour. Participants will not bid truthfully if they suspect price manipulation.

Implementation Challenges: Understanding

Berry et al. (2015) use approach with water filters in Ghana. BDM demand curve similar to take-it-or-leave-it offers.

Implementation Issues:

- People do not always bid truthfully in strategy-proof mechanisms like BDM: more dominant strategy play in ascending than second-price sealed-bid auctions.
- Worry about auctioneer misbehaviour. Participants will not bid truthfully if they suspect price manipulation.
- With uniform price, low power for ‘extreme’ WTP.

Implementation Challenges: Understanding

Berry et al. (2015) use approach with water filters in Ghana. BDM demand curve similar to take-it-or-leave-it offers.

Implementation Issues:

- People do not always bid truthfully in strategy-proof mechanisms like BDM: more dominant strategy play in ascending than second-price sealed-bid auctions.
- Worry about auctioneer misbehaviour. Participants will not bid truthfully if they suspect price manipulation.
- With uniform price, low power for ‘extreme’ WTP.
- With actual randomisation of R, no stratification.

Pilot Experiment

Pilot Experiment:

1. WTP elicitation in setting where optimal bid is known: instantly-redeemable cash voucher worth 1400 UGX (\$0.40).
2. Test different mechanisms to compare performance.

Test population: rural village residents in Eastern Uganda.

Pilot Experiment

Test 3 design features:

- **BDM** vs **ascending MPL**.

Pilot Experiment

Test 3 design features:

- **BDM** vs **ascending MPL**.
- **Public randomization** vs **preassigned prices**.

Pilot Experiment

Test 3 design features:

- **BDM** vs **ascending MPL**.
- **Public randomization** vs **preassigned prices**.
  - Public: scratchcard with 11 boxes, subject scratches one. We show examples, can scratch all boxes after, enumerators "do not know" the price locations.
  - Preassigned: scratchcard with single box. "The price has been randomly selected by the computer and enumerator does not know it."

Pilot Experiment

Test 3 design features:

- **BDM** vs **ascending MPL**.
- **Public randomization** vs **preassigned prices**.
  - Public: scratchcard with 11 boxes, subject scratches one. We show examples, can scratch all boxes after, enumerators "do not know" the price locations.
  - Preassigned: scratchcard with single box. "The price has been randomly selected by the computer and enumerator does not know it."
- **Unstated** vs **stated** price distribution.

Pilot Experiment

Test 3 design features:

- **BDM** vs **ascending MPL**.
- **Public randomization** vs **preassigned prices**.
  - Public: scratchcard with 11 boxes, subject scratches one. We show examples, can scratch all boxes after, enumerators "do not know" the price locations.
  - Preassigned: scratchcard with single box. "The price has been randomly selected by the computer and enumerator does not know it."
- **Unstated** vs **stated** price distribution.
  - Unstated: "The price might be 0, 200, ..., up to 2,000 UGX."
  - Stated: "... Each of these values appears exactly once/is equally likely to be on your scratchcard."

Pilot Experiment

Treatments

1. BDM, public randomization, stated uniform prices
2. MPL, public randomization, stated uniform prices
3. MPL, preassigned price, stated uniform prices
4. MPL, preassigned price, unstated price distribution

Main observables:

- Participant's cash on hand
- Bid for voucher
- Optimal bid = 1 if $bid \in \{1200, 1400\}$ or all cash if constrained.

Pilot Experiment: Soap Auction

After the voucher, we also sell a bar of soap worth 2000 UGX (\$0.55) using the same mechanism ($P \in \{0, 400, ..., 4000\}$).

Data

- Bid for bar of soap
- Participant belief about market price of soap (at end of survey)
- "Optimal bid" = 1 if bid $\in \{marketprice(rounded), marketprice - 400\}$, or (cash on hand + voucher profit) if constrained.

Note: only "optimal". True optimal bid equals market price only if a) participant would buy soap in the market and b) no transaction costs.

Pilot Experiment: Implementation Details

Multiple design features to aid trust and understanding

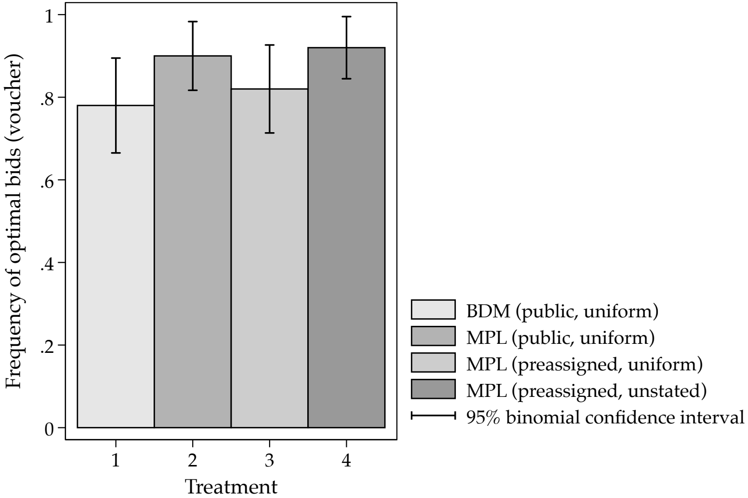
- Give participants an instantly redeemable 100 UGX voucher at start, to demonstrate voucher mechanism.
- Explain mechanism in concrete and intuitive terms.
- For a hypothetical 1600 UGX voucher, ask participant to compute win/lose outcomes and profits for different bid-price combinations.
- Check participant comprehension of price range (& distribution in stated treatments).
- After WTP elicited, check participant knows what would happen for different price draws.
- Allow to change bid if desired (nobody does).
- Final confirmation.
- Price revelation.

Pilot Experiment: Implementation Details

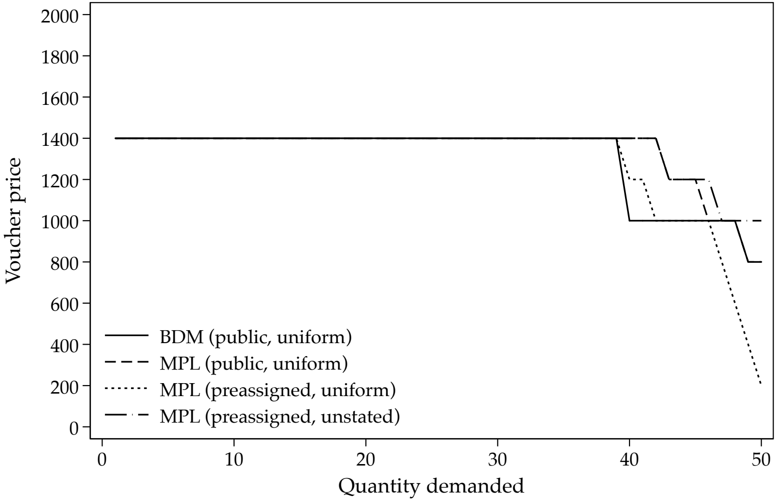
Participants:

- Recruit from 7 rural villages in Mbale District (Eastern Region) Uganda.
- Mobilizers visit village 1 day before, ask chairperson to to suggest 35 participants.
- Interviews continue until N = 200.
- Simple randomization (50 per treatment).

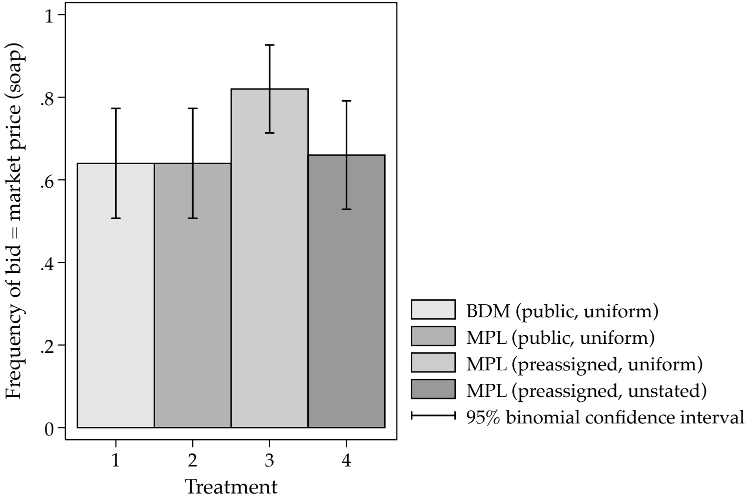
Pilot Experiment: Results



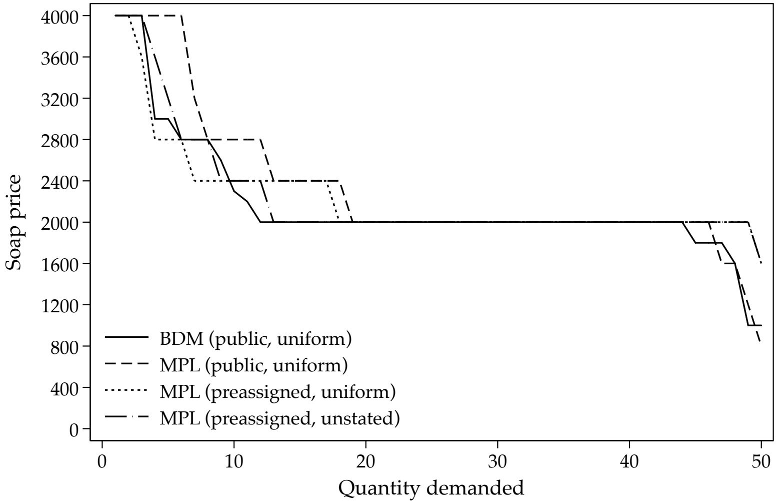
Pilot Experiment: Results



Pilot Experiment: Results



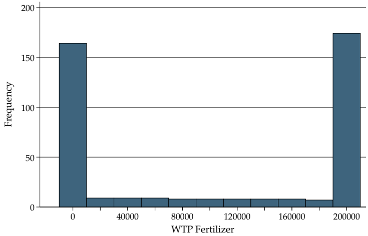
Pilot Experiment: Results

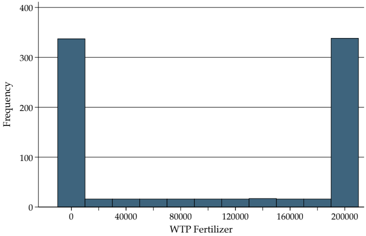


Pilot Experiment: Summary of Findings

- All mechanisms performed surprisingly well, no clear differences.
  - Contrast with earlier results.
- Gives confidence in ability to measure WTP.
- MPL (preassigned, unstated) seems safe to use.

Price Distribution





Price Revelation

[← Timeline S1]

Aim: Measure outcome y_i in population of interest $\mathcal{P}$.

Problem: Prohibitively costly. Could measure y_i in $\mathcal{P}_y \subset \mathcal{P}$.

Standard: Measure $\mathbf{x}_i$ that maybe proxies for y_i, $y_i = f(\beta \mathbf{x}_i, \epsilon_i)$, and use it instead of y_i.

Optimal Survey Design

How to use $\mathbf{x_i}$ optimally? How to choose $\mathbf{x}_i$?

(Carneiro, Lee, Wilhelm, 2017)

Optimal Survey Design: Part 1

How to use $\mathbf{x_i}$ optimally?

1. Measure y_i for $\mathcal{P}_y$ and $\mathbf{x}_i$ for $\mathcal{P}$.

2. Find $\hat{f}(\mathbf{x}_i) = \arg\min_{g(\mathbf{x}) \in \mathcal{C}(\mathbf{x})} L(y_i, g(\mathbf{x}_i))$.

3. Predict $\hat{y}_i = \hat{f}(\mathbf{x}_i)$ for all $i \notin \mathcal{P}_y$.

Optimal Survey Design: Part 1

How to use $\mathbf{x_i}$ optimally?

1. Measure y_i for $\mathcal{P}_y$ and $\mathbf{x}_i$ for $\mathcal{P}$.

2. Find $\hat{f}(\mathbf{x}_i) = \arg\min_{g(\mathbf{x}) \in \mathcal{C}(\mathbf{x})} L(y_i, g(\mathbf{x}_i))$.

- $L(y_i, g(\mathbf{x}_i))$ usually MSE.

3. Predict $\hat{y}_i = \hat{f}(\mathbf{x}_i)$ for all $i \notin \mathcal{P}_y$.

Optimal Survey Design: Part 1

How to use $\mathbf{x_i}$ optimally?

1. Measure y_i for $\mathcal{P}_y$ and $\mathbf{x}_i$ for $\mathcal{P}$.

2. Find $\hat{f}(\mathbf{x}_i) = \arg\min_{g(\mathbf{x}) \in \mathcal{C}(\mathbf{x})} L(y_i, g(\mathbf{x}_i))$.

- $L(y_i, g(\mathbf{x}_i))$ usually MSE.
- Machine learning good at this!

3. Predict $\hat{y}_i = \hat{f}(\mathbf{x}_i)$ for all $i \notin \mathcal{P}_y$.

Optimal Survey Design: Part 1

How to use $\mathbf{x_i}$ optimally?

1. Measure y_i for $\mathcal{P}_y$ and $\mathbf{x}_i$ for $\mathcal{P}$.

2. Find $\hat{f}(\mathbf{x}_i) = \arg\min_{g(\mathbf{x}) \in \mathcal{C}(\mathbf{x})} L(y_i, g(\mathbf{x}_i))$.

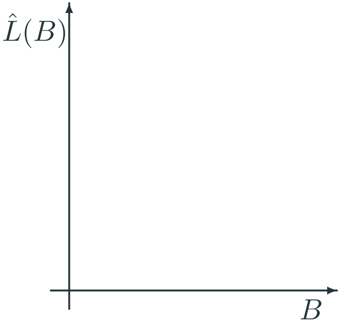
- $L(y_i, g(\mathbf{x}_i))$ usually MSE.
- Machine learning good at this!
- Evaluate $L(y_i, \hat{f}(\mathbf{x}_i))$ in $\mathcal{P}_{\text{Test}} \subset \mathcal{P}_y$.

3. Predict $\hat{y}_i = \hat{f}(\mathbf{x}_i)$ for all $i \notin \mathcal{P}_y$.

How to choose $\mathbf{x}_i$?

1. Prices p_j for each x_{ij}.

$\rightarrow B(\mathbf{x}_i; \mathbf{p}) := \sum_{i \in \mathcal{P}} \mathbf{p}'\mathbf{x}_i$.

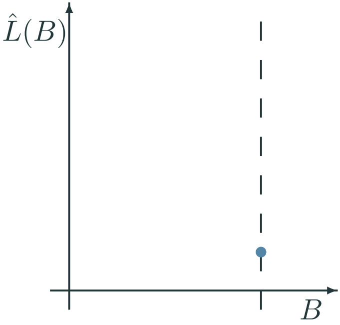


Optimal Survey Design: Part 2

How to choose $\mathbf{x}_i$?

1. Prices p_j for each x_{ij}.

$\rightarrow B(\mathbf{x}_i; \mathbf{p}) := \sum_{i \in \mathcal{P}} \mathbf{p}'\mathbf{x}_i$.

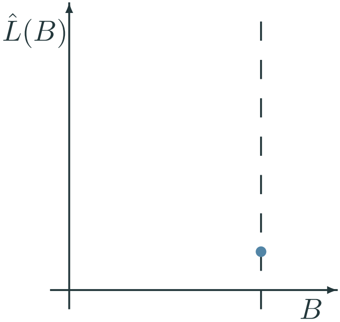


How to choose $\mathbf{x}_i$?

1. Prices p_j for each x_{ij}.

$\rightarrow B(\mathbf{x}_i; \mathbf{p}) := \sum_{i \in \mathcal{P}} \mathbf{p}'\mathbf{x}_i$.

2. For all $B < B(\mathbf{x}_i; \mathbf{p})$:
find subset of variables $\mathbf{z}$ that minimises $L(y_i; \hat{f}(\mathbf{z}_i))$, s.t. $B(\mathbf{z}; \mathbf{p}) \leq B$; evaluate $L(y_i; \hat{f}(\mathbf{z}_i))$ in $\mathcal{P}_{\text{Test}^*} \subset \mathcal{P}_y$.



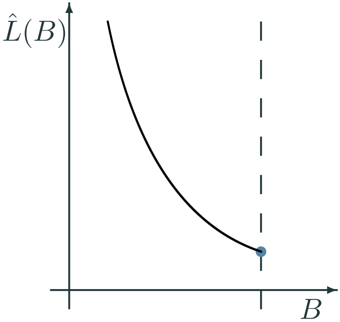
Optimal Survey Design: Part 2

How to choose $\mathbf{x}_i$?

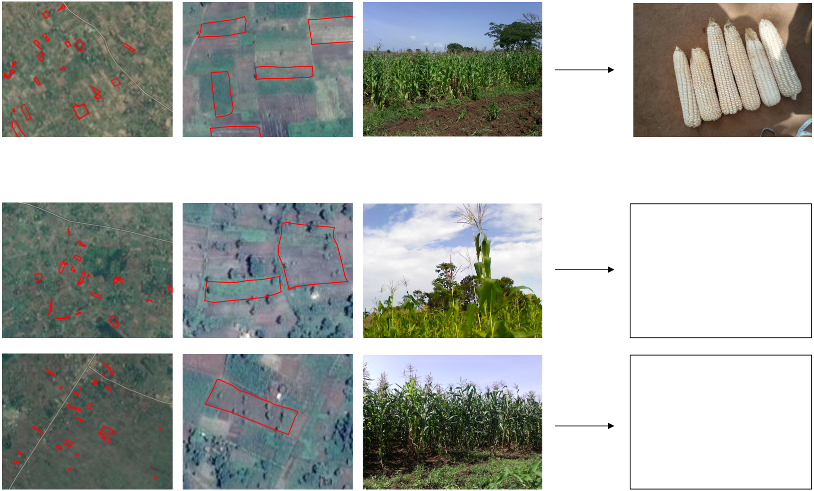
1. Prices p_j for each x_{ij}.

$\rightarrow B(\mathbf{x}_i; \mathbf{p}) := \sum_{i \in \mathcal{P}} \mathbf{p}'\mathbf{x}_i$.

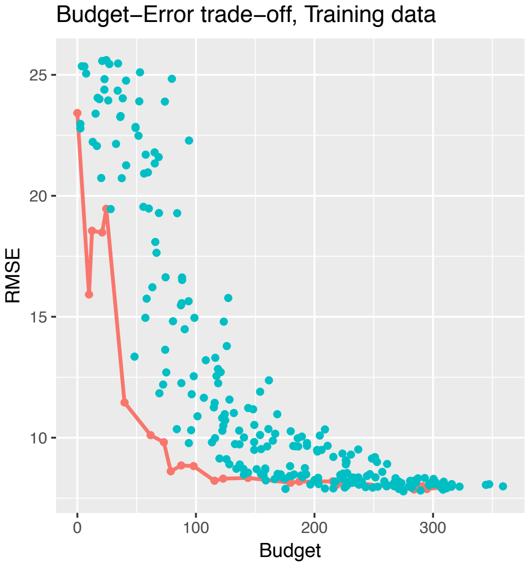
2. For all $B < B(\mathbf{x}_i; \mathbf{p})$:
find subset of variables $\mathbf{z}$ that minimises $L(y_i; \hat{f}(\mathbf{z}_i))$, s.t. $B(\mathbf{z}; \mathbf{p}) \leq B$; evaluate $L(y_i; \hat{f}(\mathbf{z}_i))$ in $\mathcal{P}_{\text{Test}^*} \subset \mathcal{P}_y$.



Optimal Survey Design: Our Example

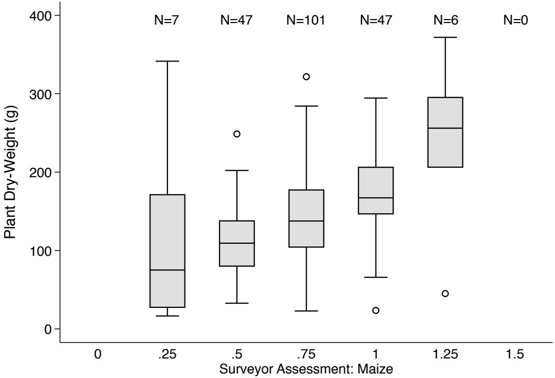


Optimal Survey Design: Our Example



Optimal Survey Design: Our Example

Figure 1: Enumerator Expectations and Realized Maize Yield.



Notes: The figure depicts box plots of realized dry yield per plant in a quadrant, at each discrete level of enumerator expectations. *Plant Dry-Weight (g)* is measured in grams and calculated as the average weight (at 12% moisture) of sampled maize ears in a quadrant multiplied by the observed number of ears per plant. The enumerators expectations are the enumerators' assessment of the expected yield per plant of maize plants in a quadrant relative to the yield of healthy plant. Answers were given as shares from a discrete set of options, $s \in \{0\%, 25\%, 50\%, 75\%, 100\%, 125\%, 150\%\}$. The grey area of each box plot depicts the 25th to 75th percentile range, and the inside black line depicts the median. Above each box plot the number of observations at the is stated. All data is at the quadrant level.

Optimal Survey Design: Our Example

Table 2: Correlation of Yield Measures

	Realized Yield	Enumerator Assessment	Farmer-Reported
Realized Yields	1.000		
Enumerator Assessment	0.816	1.000	
Farmer-Reported	0.395	0.463	1.000

Notes: The table reports the pair-wise Pearson's correlation coefficient between any two measures of yields, using information from the 69 plots for which all information is available. *Realized Yield* is based on a combination of Crop-Cutting and Pre-Harvest Crop Assessment data with the key ingredient being the measured dry-weight of ears; *Enumerator Assessment* is solely based on Pre-Harvest Crop Assessment with the key ingredient being the enumerator assessment of yield per plants; and *Farmer Reported* is based on the farmers' self-reported yield in the Follow-Up Survey.

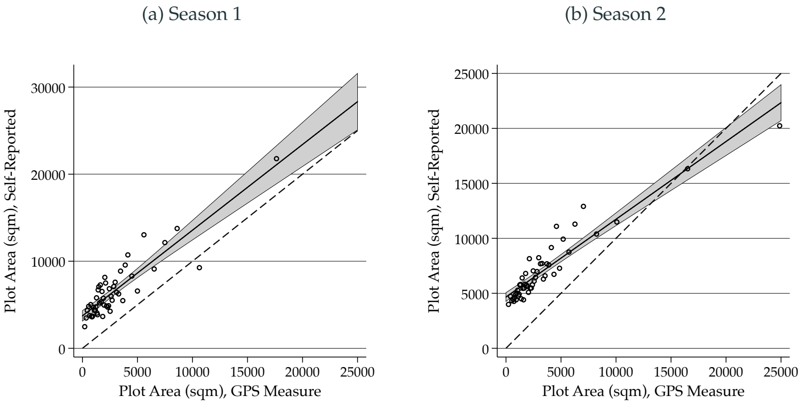
[← Timeline S1]

Measuring Plot Sizes

Figure 3: Crop Assessment Field Tracks

Measuring Plot Sizes

Figure 4: Land Size Measurement



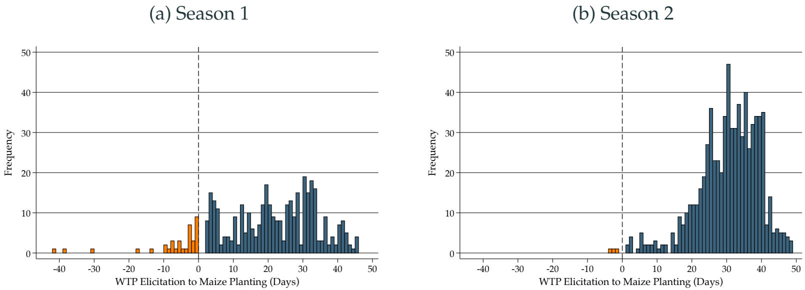
Implications for measuring farm size – productivity relation, wealth inequality, and gains from reallocation. [← Timeline S1]

Quadrant Placement

[← Timeline S1]

Planting Time

Figure 5: WTP Elicitation to Maize Planting



[← Timeline S1]

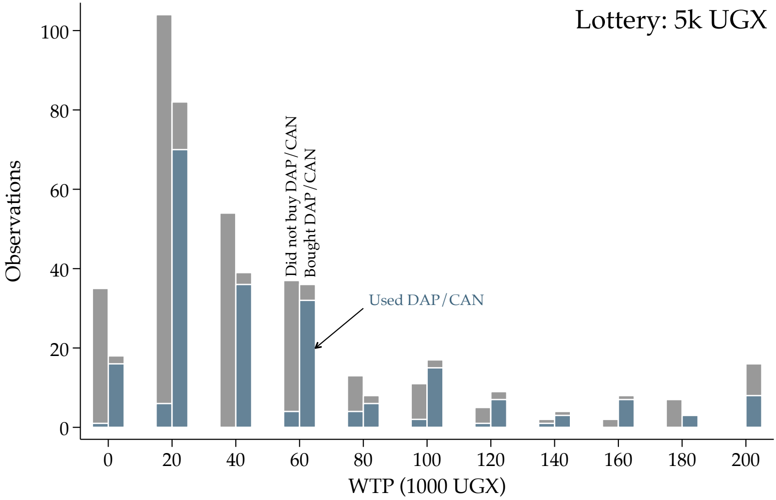
Timeline: Season 2

Survey Activity

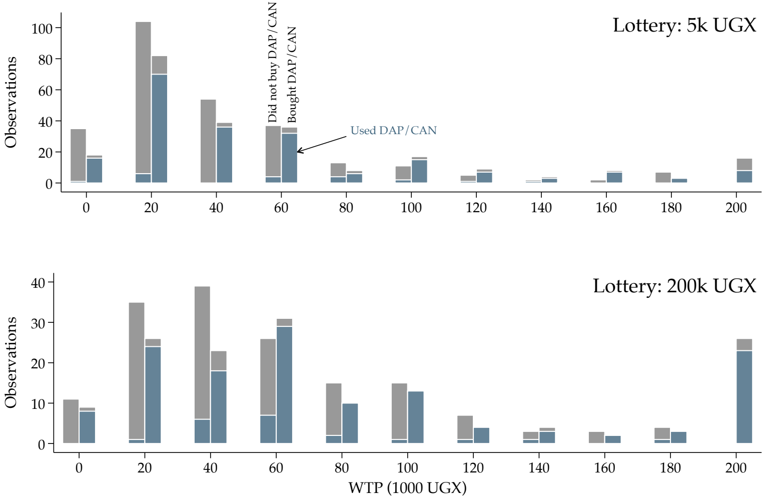
1. **Census:** 18/9-20/11/2017.
2. **Baseline:** 14/1-7/2/2018.
3. **WTP Elicitation:** 3/2-14/2/2018.
4. **Phone Survey 1:** 21/3-28/3/2018.
5. **Crop Survey:** 11/6-3/7/2018.
6. **Follow-Up Survey:** 27/8-13/9/2018.
7. **Phone Survey 2:** 14/12-23/12/2018.

[← Survey Activities: Season 1]

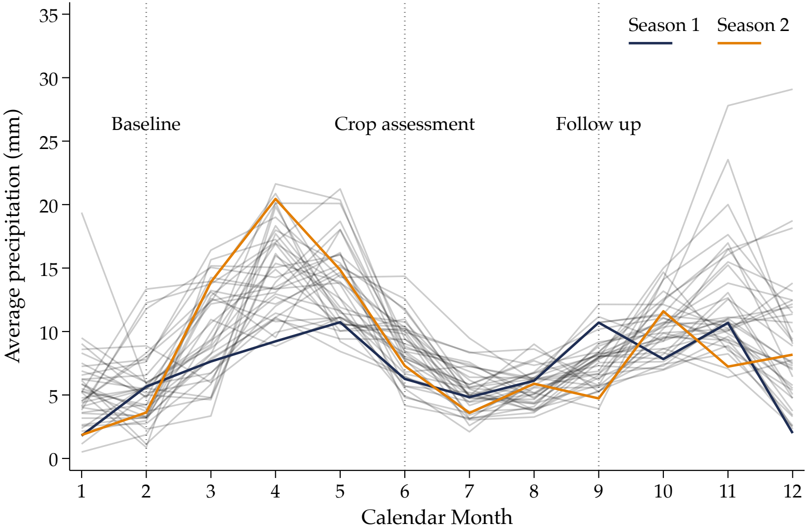
First Stage (Lottery: 5k)



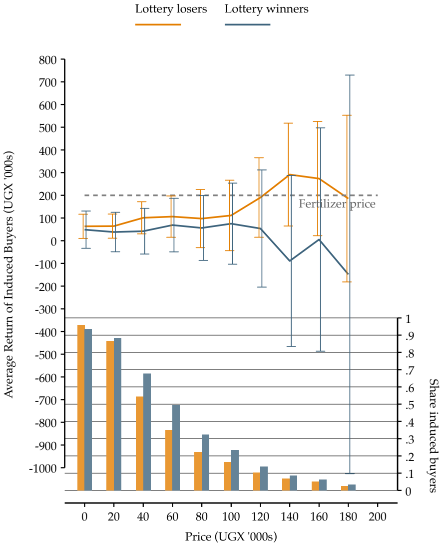
First Stage


Lottery: 5k UGX
Observations
Did not buy DAP/CAN
Bought DAP/CAN
Used DAP/CAN
0
20
40
60
80
100
0
20
40
60
80
100
120
140
160
180
200
Lottery: 200k UGX
Observations
0
10
20
30
40
WTP (1000 UGX)

Weather Patterns 1981-2018



Average Treatment Effects of Induced Buyers: 50% wage



←