

# Steps In Implementing an Impact Evaluation

Christopher Udry

Yale University

# Outline

1. Identify Priorities
2. Understand roll-out of intervention and opportunities for impact evaluation
3. Appropriate design
4. Assignment to treatment and control
5. Collect baseline data
6. Check for balance
7. Monitor and gather information on implementation
8. Follow-up data
9. Estimate impact
10. Analyze impact
11. Disseminate
12. Iterate

# Step 1: Identify Priorities, and focus on priorities for learning

- Identify the highest priorities of the intervention/program
- The overall goal of the MCC-Ghana programme is to reduce poverty through agriculture transformation
- This is being done by addressing two sub-goals
  - Increasing production and productivity of high-value cash and staple food crops
  - enhance the competitiveness of horticulture and other traditional crops of Ghana

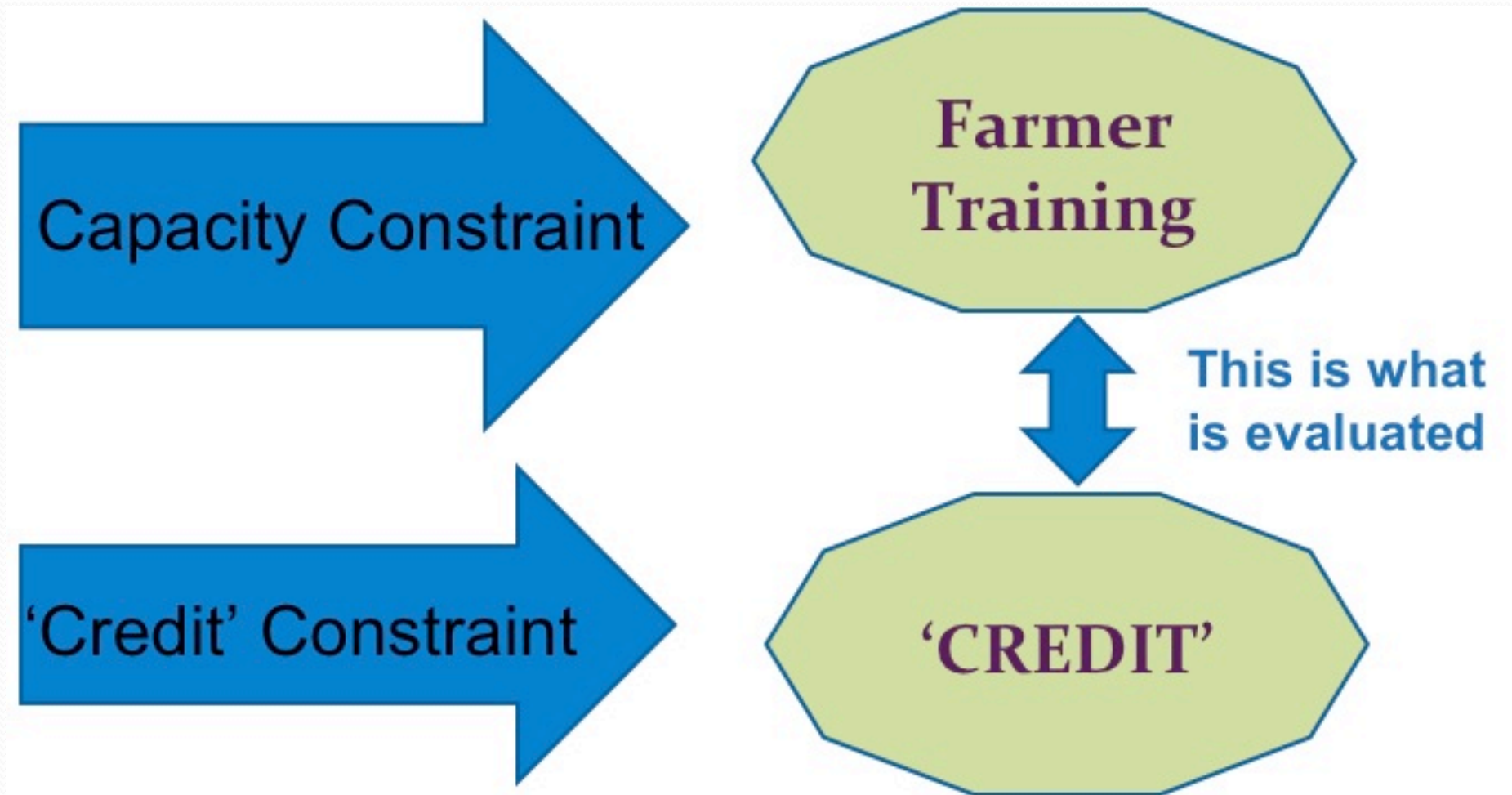
## *Identify Priorities for Learning*

- The Agriculture Project has the following components
  - **Farmer and Enterprise Training in Commercial Agriculture**
  - Irrigation development
  - Land tenure facilitation
  - Improving post harvest handling
  - **Improve credit services**
  - Feeder roads improvements
- The MCA-Ghana programme covers 30 districts (initially 23 but changed with a re-demarcation exercise)

## *Identify Priorities for Learning*

- Decide on the Outcomes of Interest
  - End-goals (poverty, welfare, education, consumption ....)
  - Intermediate, or mechanism-based
    - Yield, profit, adaption rates, knowledge
- Scaling-up? Test alternatives, test cost-benefit, examine sustainability
- Is this feasible?

# What do we seek to evaluate?



## 2. Understand rollout of intervention and opportunities for impact evaluation

- How will the program be rolled out? Different interventions?
  - Piloted in a random sample of communities / population / regions?
  - Rolled out nationwide?
  - **Rolled out in communities/population/regions satisfying a certain criteria?**
  - **Rolled out to a targeted high-potential or high-poverty population/areas?**

→ Understand **TARGETING** and **PROGRAM PARTICIPATION**

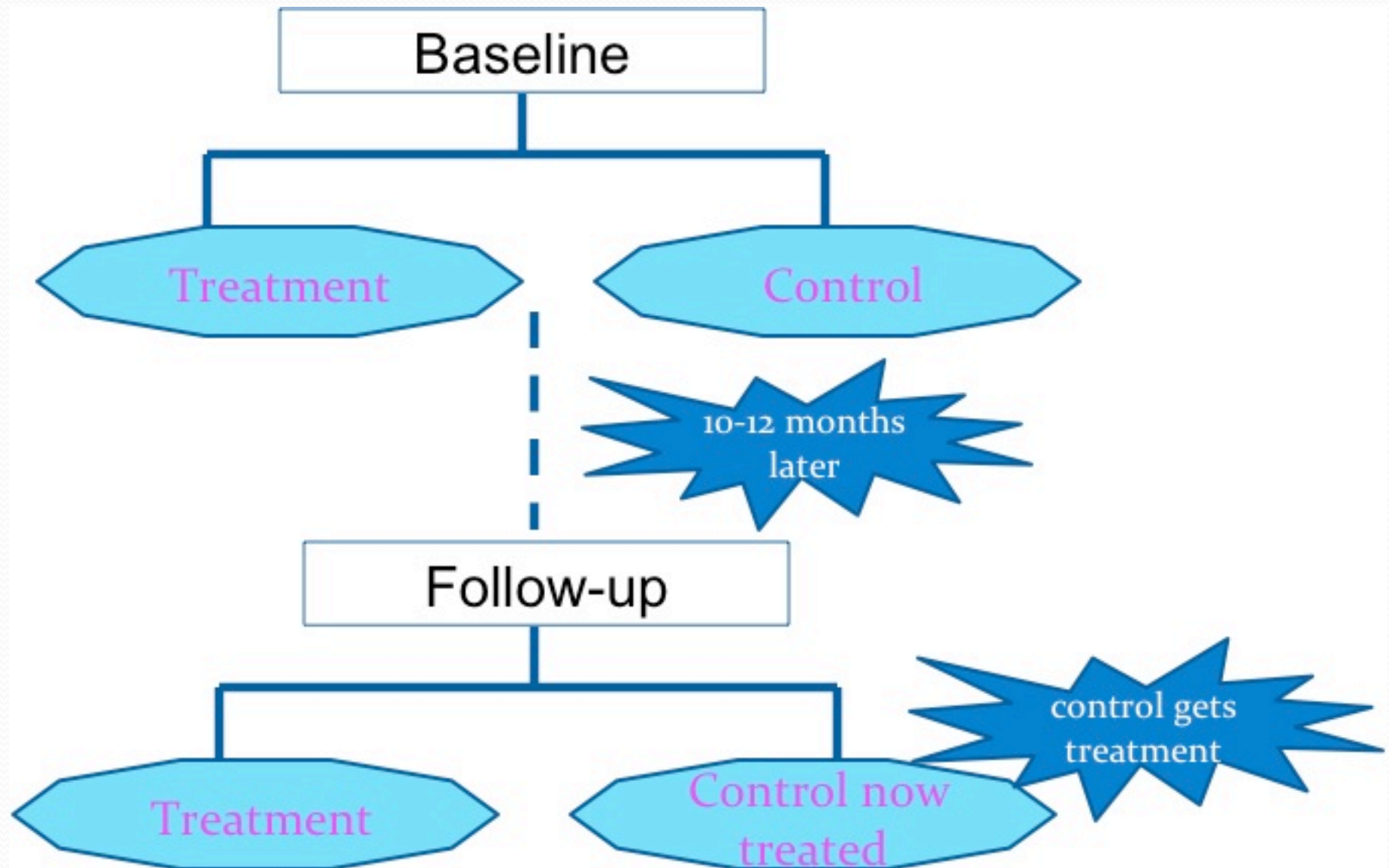
- Each roll-out strategy yields distinct opportunities for impact evaluation

# Back to MCC-Ghana Program

- Limited possibilities for randomized evaluation of roads, irrigation, land tenure, post-harvest handling
- But training, “starter pack” and credit has possibilities
  - Investment
  - Yields, Crop incomes (profits)
  - Welfare
- To answer these questions ISSER used a randomised phasing-in of beneficiaries
  - Unfortunately, cannot distinguish components of program



## *Rollout, instead of C/T*



### *3. Appropriate Design*

- Determine Scale: Large scale or small pilot?
  - Large scale – easier to make representative
  - But costlier
  - May permit rich analysis of effectiveness by region, agroclimate
  - Small pilot cheaper, easier to implement
  - Not as informative
    - For agriculture, aggregate shocks matter, small pilot can be utterly uninformative on impact

### *3. Appropriate Design*

- Rollout across MIDA districts in Ghana, simultaneous
- Training of farmers is being done at the FBO level, so
  - Of 1200 FBOs we select 5 farmers from each to get a total of 3000 farmers
  - The FBOs are randomly selected into treatment and control groups (this was done in a participatory way to deal with initial ethical issues that arose)
  - For practical purposes the data was collected for two batches (of 600 each)

### 3. Design: Rollout timing

		Year 1		Year 2		Year 3	
		survey	treat	survey	treat	survey	Treat t
Batch 1	Early Treatment	B	X	F			
	Late Treatment	B		F	X		
Batch 2	Early Treatment			B	X	F	
	Late Treatment			B		F	X

- Plan for analysis:

$$y_{ict} = \alpha_c + \alpha_t + T_{ct}\beta(+\lambda_i) + \varepsilon_{ict}$$

- Or in a simpler set-up

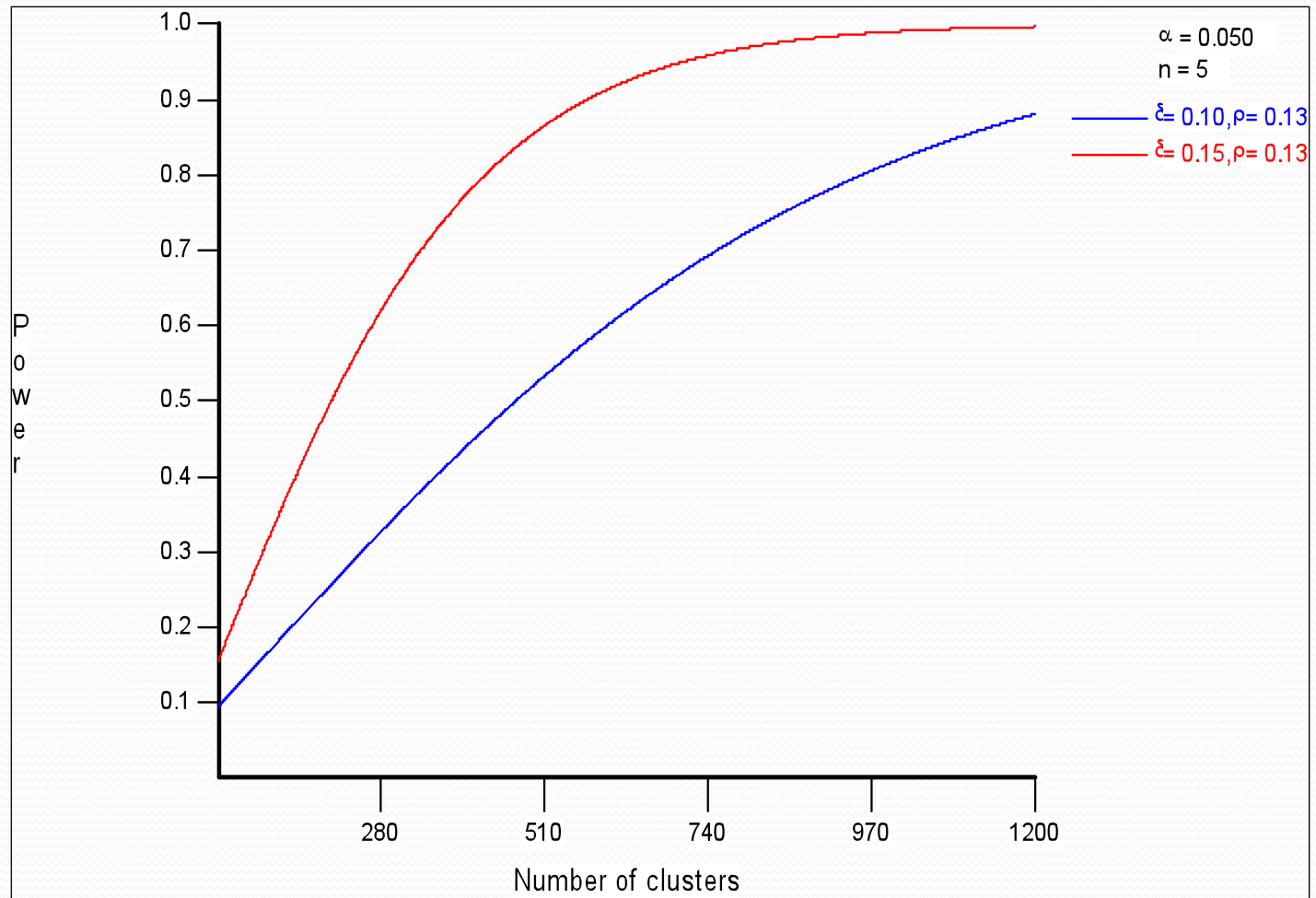
$$\bar{Y}_{T,post} - \bar{Y}_{C,post}$$

or

$$\bar{Y}_{T,post} - \bar{Y}_{C,post} - (\bar{Y}_{T,pre} - \bar{Y}_{C,pre})$$

- This is the moment for power analysis – will your strategy be able to detect the kinds of impacts you want to detect?

# Power Calculation: MCC FBO Evaluation



## *4. Assignment to treatment and control*

- Is random assignment feasible?

Large-scale irrigation project : Not feasible to assign farm households or communities randomly to the intervention, determined by location of canals

Input vouchers: Can randomly assign at region, community, or household level- contamination?

- At what level? The unit of intervention often simplest
- Trade-off: higher level means bigger sample
- Public, transparent often best



## 5. Collect baseline data

- ‘optional’ step
- Random assignment: Implies control and treatment are identical, but we must check for balance
- If not balanced, reassign/correct?
- Regression Discontinuity Design and other quasi-experimental methods: Baseline essential, matching on observables and checking balance
- Informs project design and implementation: Who was targeted? Did the program mostly benefit patients who were poor or at high risk at baseline? How well were they targeted?

## 5. MCC Ghana

- Used a Household survey instrument covering
  - Demographic Characteristics, Household membership and FBO activities
  - Educational characteristics of households
  - Household Health
  - Activity status of household members
  - Migration
  - Transfers in and out of the households
  - Information seeking behaviour of households

- Household assets as well as their borrowing, savings and lending behaviour
- Housing characteristics of households
- Household agriculture activities including land ownership and transactions and agriculture processing
- Non-farm enterprises of households

## 6. Check for balance, describe sample

- If imbalanced, serious difficulty but all is not lost

# BASLINE SURVEY: SUMMARY OF KEY FINDINGS

## Distribution of sample by Batch and Treatment Group

	Early Treatment (Treatment)	Late treatment (Control)	Total
Batch 1	2,808	2,508	5,316
Batch 2	2,829	2,875	5,704
Total	5,637	5,383	11,020

# Distribution Farmers in Sample by Zone

MiDA Zone	Average household size	Number of persons	Number of households
<u>Batch I Round II Sample</u>			
Southern Horticulture	5.7	3,541	625
Afram Basin	5.2	5,173	992
Northern	7.6	5,882	776
Total	6.1	14,596	2,393
<u>Batch II Round II Sample</u>			
Southern Horticulture	5.1	3,927	771
Afram Basin	4.9	5,238	1,070
Northern	7.2	7,254	1,001
Total	5.8	16,419	2,842

# Crop Yields Batch 1

	Treatment	Control	
VARIABLES	Mean	Mean	P-value
Cassava	4.7	3.6	0.2158
Groundnut/ Peanut	0.7	0.8	0.3954
Sorghum	0.5	1.3	0.0861
Maize	1.5	1.5	0.9415
Millet	0.9	1.2	0.1646
Pepper	1.7	1.4	0.6487
Pineapple	22.6	23.9	0.7973
Rice	0.8	0.9	0.4745
Soybean	0.8	0.8	0.7699
Yam	2.6	3.2	0.0858

- Yields are highest for pineapples at about 23tons/ha
- Of the grains rice and sorghum seem to have the lowest yields



# Crop Yields Batch 2

	Treatment	Control	
VARIABLES	Mean	Mean	p-value
Cassava	4.4	3.3	0.067
Groundnut/Peanut	1	1.1	0.7899
Sorghum	1.1	0.7	0.4025
Maize	1.6	1.5	0.2237
Millet	0.7	0.9	0.5301
Pepper	1.5	0.9	0.4735
Pineapple	15.1	4.1	0.1863
Rice	1.6	1.3	0.0438
Soybean	0.7	0.6	0.6277
Yam	6.9	6.2	0.4465
Observations	2,479	2,718	

- Baseline Yields are similar across the batches

# Crop Incomes Batch 1

	Treatment	Control	
VARIABLES	Mean	Mean	P-value
Maize	543	344.9	0.1278
Cassava	473	196.8	0.2885
Soya	153.8	92.8	0.1675
Yams	497	360.3	0.4029
Rice	278.6	403.1	0.2372
Millet	153	49.6	0.0097
Groundnuts	299	244.3	0.5365
Pineapples	1,318.40	769.1	0.2952
Pepper	988.4	321.7	0.2251

- Crop incomes are highest for the export oriented crops
- For this batch one observes that crop incomes increase over the two periods

# Crop Incomes Batch 2

	Treatment	Control	
VARIABLES	Mean	Mean	P-value
Maize	625.1	636.9	0.869
Cassava	843.3	1,790.80	0.2422
Soya	187.7	118.7	0.0837
Yams	1,453.20	1,504.50	0.9132
Sorghum	208.3	1,118.70	0.04
Rice	604.1	502.8	0.2297
Millet	295.2	300.8	0.9569
Groundnuts	264.7	283.5	0.7111
Pineapples	1,733.80	3,901.30	0.4366
Pepper	571.9	607.5	0.9087
	692.4	802.6	0.2037

## 7. Roll out intervention

- Monitor to ensure that evaluation is not compromised
  - Contamination: treatment provided to control group
  - Can still use ITT, but immediately weakens power, ultimately can make evaluation impossible
- Is treatment group being provided with additional treatments?
  - e.g., NGO targets treatment communities with another intervention
  - Changes evaluation
- Compensating policy changes, can lead to finding of no impact

# MiDA Training and Starter Pack

- Training
  - Business capacity
  - Technical
  - Marketing
- Starter Pack (GH¢400=US\$230)
  - Fertilizer
  - Seeds (for one acre)
  - Protective clothing
  - GH¢30 (US\$22)

## 8. Follow-up Data

- Collect follow-up data for both the treatment and control groups
- Appropriate intervals
  - Consider how long it should take for outcomes to change
  - One year or at next harvest
    - Provide initial outcomes
    - Adjust program if needed
  - Two years: Changes in longer term outcomes?
  - After end of program: Do effects endure?
    - What happens once the input voucher program has phased out?

- Key issue for MCC evaluation: only 1 year gap between treatment of early FBOs and treatment of late (control FBOs).
  - Can only measure short run impact
  - Unlikely to see any persistent changes...

## 9. Estimate program impact

- Randomization: Simply compare average outcomes for treatment and comparison. So easy!
- Other methods: Make statistical assumptions to estimate impact of program
- Combination of methods: Matching with difference-in-difference



# **IMPACT EVALUATION: SOME PRELIMINARY RESULTS**

# Yields (tonnes/ha)

	(1) Yield	(2) Income	(3) Revenue	(4) Cost
VARIABLES	lyield	lincome	lrev	lcost2
Period 2	0.1***	0.5***	-0.0	-0.9***
Period 3	0.1	1.0***	0.3**	-1.0***
Treatdum	0.0	0.1	0.0	-0.0
Treattime	-0.1	-0.2	-0.1*	-0.0
Batchdum	-0.2***	-0.2**	-0.2**	0.1
Constant	-0.0	4.9***	6.3***	6.1***
Observations	10,706	4,006	4,006	4,006
R-squared	0.0	0.0	0.0	0.1

- Training and starter pack does not impact on crop yield and income

# Inputs use

	Size	Chemical	Seeds	Lab_Total
VARIABLES	lsize	lchem	lseed	TLab
Period 2	-0.2***	0.5***	0.3***	-422.2***
Period 3	-0.2***	0.5***	0.2***	-846.1***
Treatdum	-0.1***	-0.1	-0.2***	-141.0*
Treattime	-0.0	0.3***	0.1**	163.2
Batchdum	0.0	-0.1**	-0.0	213.4**
Constant	0.6***	4.5***	4.3***	1,138.4***
Observations	3,612	7,008	8,066	8,974

- Programme increases the value of chemical and seeds costs by 30 and 10 % respectively
- There is however no effect on cultivated area

# Labour Use

VARIABLES	TLabLP	TLabFM	TLabH	TLabPH
Period 2	-491.3***	35.5	-27.0	60.6***
Period 3	-575.0***	-104.1	-165.8***	-1.2
Treatdum	-18.9	-110.7***	-9.3	-2.1
Treattime	31.0	95.7*	19.7	16.8
Batchdum	27.7	140.8***	15.7	29.3*
Constant	627.8***	196.8***	270.8***	43.0***
Observations	8,974	8,974	8,974	8,974

- Labour use for field management increased by about 96 hours as a result of the programme

# Main findings

- This study evaluates the impact of the training component on crop yields and income and finds that
  - The training has had no impact on crops yields and crop incomes
  - However, expenditures by farmers on seed and chemical use were positively impacted
  - We conclude by noting that although no impact is found for crop yields and income, farmer behaviour seems to have been impacted positively. Part of this however, may have been due to the starter pack

## 10. Analyze impacts

- Are the effects statistically significant?
  - Basic statistical test tells whether differences are due to the program or to noisy data
- Are they policy significant?
  - If the input voucher scheme costs a million dollars and has positive effect but it's tiny, may not be worthwhile
- Are they sustainable?
  - If input use falls to pre-program levels when the intervention ends, the program is not financially sustainable in its current form

# 10. Disseminate!

- If no one knows about it, it won't make a difference to policy!
- Make sure the information gets into the right policy discussions
- Ownership by government, capacity building
- Forums
  - Workshop
  - Report
  - Policy brief

## 12. Iterate

- Re-examine sector **priorities**: Identify **next learning opportunity**
- Or suppose the effects aren't as large as you hoped
  - Test **variations**
    - Alternate subsidy amounts
    - Alternate packages of inputs
    - Alternate implementation and targeting mechanisms: Government extension workers or input dealers? Beneficiary selection?
  - Test other **interventions** to affect same outcomes
    - Matching grants for technology adoption
    - Training in use of improved technologies
    - Improving access to markets and providing complementary infrastructure to increase the share of marketed output