Lecture 10:
How to Make a Pre-Analysis Plan (PAP)

Edmund Malesky, Ph.D.

July 16, 2018

Duke University

What is a PAP?

- A PAP, sometimes called a research or study protocol, is a document outlining the technical details of a study
  - Written before the study is conducted
  - Includes: type of study to be conducted, sources of data, how the variables will be constructed, model specifications, problems that may arise over the course of the study, and how those problems will be addressed.

Why are PAP's Necessary?

- Raise the credibility and reliability of research results.
- PAPs allow for a higher level of confidence, particularly where researchers are seeking to confirm a hypothesis rather than simply explore potential causal relations.
- By "tying one's 2 hands" through a PAP, reported results will not reflect statistical noise hand-picked from the data.

When Should We Make a PAP?

- A PAP should be written and registered ***before*** the intervention begins, so as to ensure and publicly display the fact that a hypothesis is made prior to seeing any data.

1. Page 1

- **Name/title** of the project
- **Authorship:** all researchers involved and their affiliations.
- **External partner** institutions (if applicable)
- **Conflicts of Interest:** list any conflicts of interest for all study authors
- **Brief summary** of project and objective/rationale for conducting this study.

2. Statement of Policy Problem

- What is the key issue faced that requires policy change or improvement?
  - Outcome variable that we would like to change.
  - Previous research on the issue.
  - Any conflicting views about the problem

3. Background Research

- Delineate different policy suggestions.
- Discuss prevailing theories in literature.
- Where does this research fit-in?
- Discuss previous research on this problem in Vietnam and in other contexts.

4. Theory of Change

- Identify key hypotheses to be tested.
  - Hypotheses are "testable statement about the empirical relationship between cause and effect."
    - Specify units being compared
    - Which variables are expected to be related
    - Tendency of that relationship
- Declare any important sub-group analysis and expected heterogeneous effects

4. Theory of Change

Distribution of Raskin ID cards to Raskin beneficiaries

Sample assumption: Successful delivery to household, no "bypassing" of cards

Reception of Raskin ID cards

Sample assumption: Beneficiaries understand use of cards, use cards, do not confuse with previously existing Raskin cards/coupons

Beneficiaries obtain more of their Raskin subsidies

***Sample assumption*:** Beneficiaries demand reduced prices, village Raskin authorities listen to beneficiary demands and have ability to make change

Increase effectiveness of social protection programs

***Sample assumption*:** Fair(er) practices sustained, poor program effectiveness due to lack of transparency

9

5. Log Frame

Inputs	Activities	Outputs	Outcomes	Final Outcomes
- Budget for new vocational training program. - Staffing from DOET, vocational instructors - Training facilities (rooms, computers)	- Design of new curriculum. - Instructor training. - Development and printing of course materials. - Development of software	- 1000 vocational training instructors trained on new curriculum. - 100,000 textbooks and copies of software delivered to classrooms	- Instructors use new textbooks and software in class. - Students are learning new curriculum. - Improved student performance of end of class exam	- Improved competency of material. - More students hired by firms. - Higher employment - Labor productivity of firms goes up.

Implementation (Supply Side)

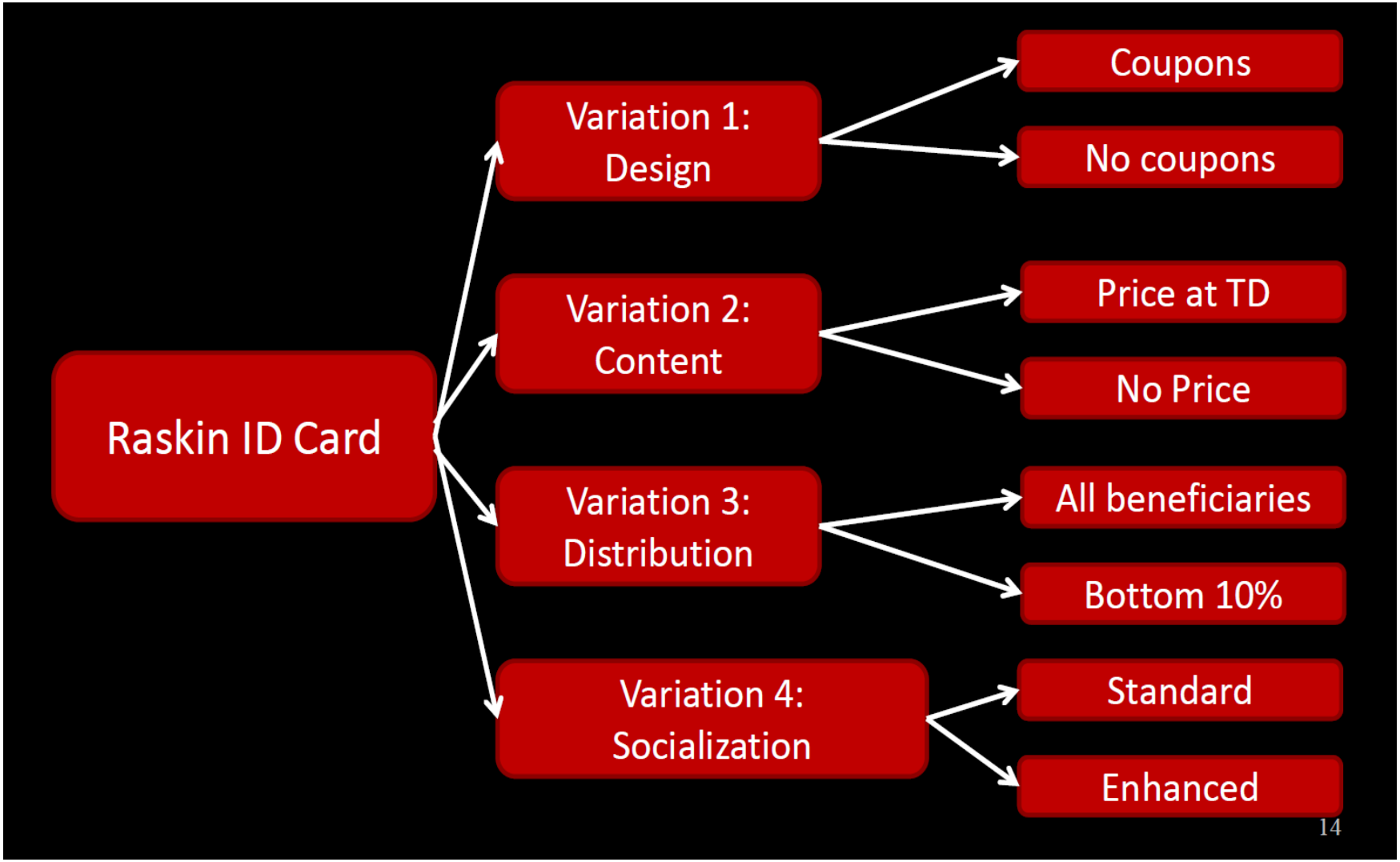
Results (Demand + Supply)

6. RESEARCH DESIGN

6.1. Major Design Choice

1. Randomized Controlled Trial
    - Type
        - Clinical Design
        - Over-Subscription
        - Randomized Oder of Phase-In
        - Encouragement
        - Spillover/Saturation
    - Number of Arms
        - Single arm
        - Multiple arms
        - Factorial design (identify treatment groups)

Tree Approach

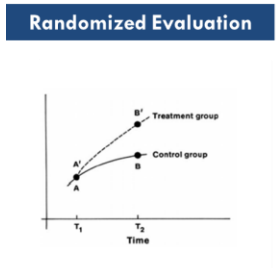


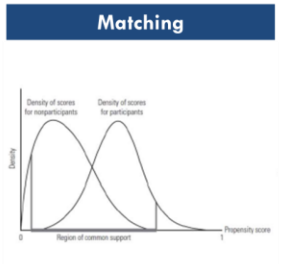
Tabular Approach

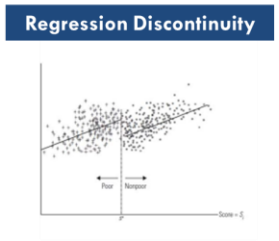
Card Variations			Standard socialization	Enhanced Socialization
All beneficiary	Price	Coupon	Group 1	Group 2
		No Coupon	Group 3	Group 4
	No Price	Coupon	Group 5	Group 6
		No Coupon	Group 7	Group 8
Bottom 10%	Price	Coupon	Group 9	Group 10
		No Coupon	Group 11	Group 12
	No Price	Coupon	Group 13	Group 14
		No Coupon	Group 15	Group 16
			Control (No card, no socialization)	

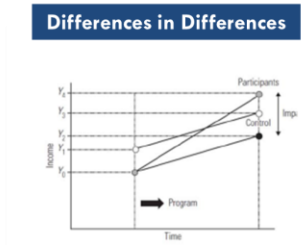
6.1. Major Design Choice

2. Natural Experiment
   - Type
     - Regression Discontinuity
     - Difference-in-Differences
     - Propensity Score Matching
     - Instrumental Variables
     - Combined approach









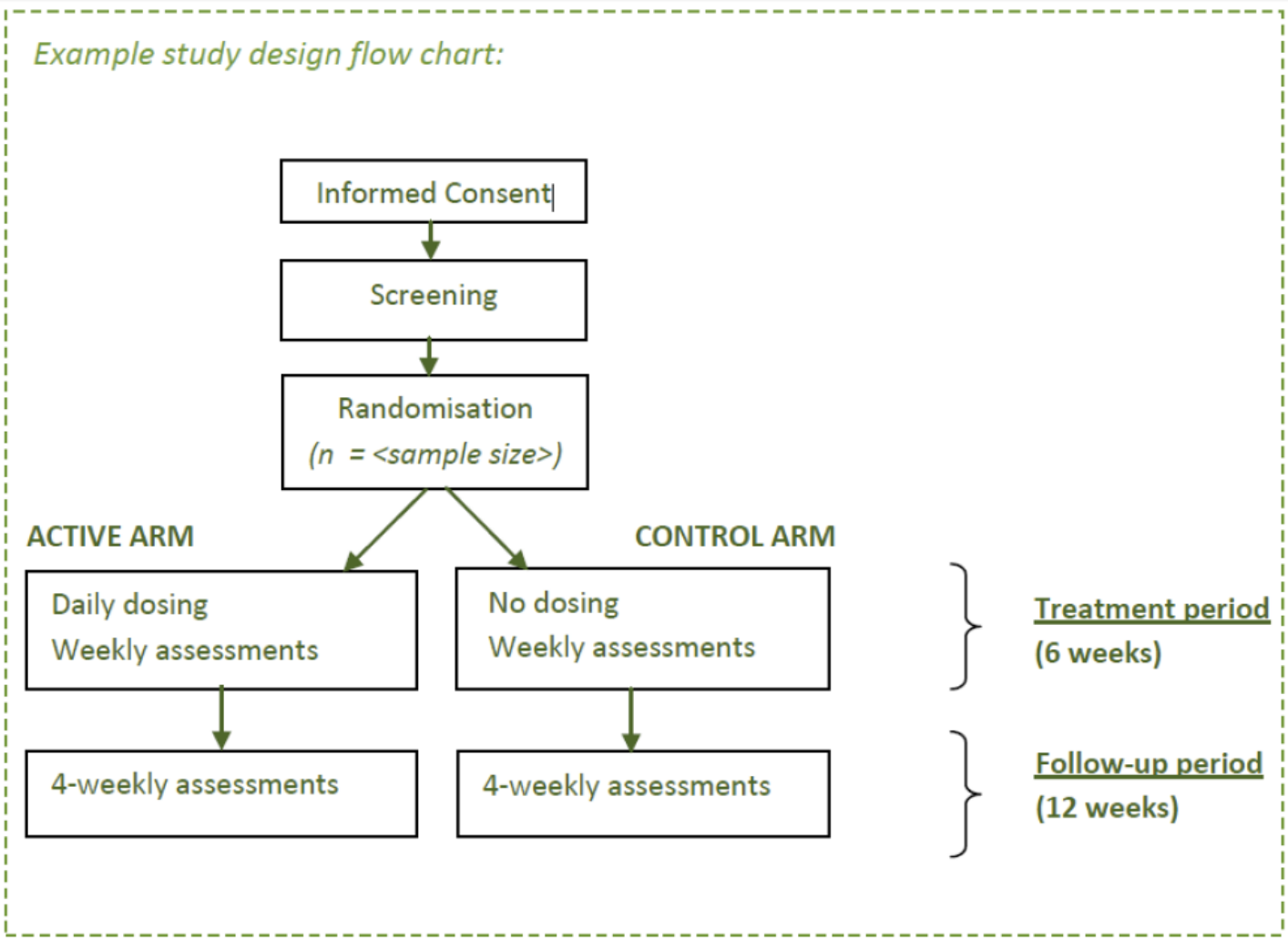
37

Design	When to use	Advantages	Disadvantages
Randomization	▫ Whenever feasible ▫ When there is variation at the individual or community level	▫ Gold standard ▫ Most powerful	▫ Not always feasible ▫ Not always ethical
Randomized Encouragement Design	▫ When an intervention is universally implemented	▫ Provides exogenous variation for a subset of beneficiaries	▫ Only looks at sub-group of sample ▫ Power of encouragement design only known ex post
Regression Discontinuity	▫ If an intervention has a clear, sharp assignment rule	▫ Project beneficiaries often must qualify through established criteria	▫ Only look at sub-group of sample ▫ Assignment rule in practice often not implemented strictly
Difference-in-Differences	▫ If two groups are growing at similar rates ▫ Baseline and follow-up data are available	▫ Eliminates fixed differences not related to treatment	▫ Can be biased if trends change ▫ Ideally have 2 pre-intervention periods of data
Matching	▫ When other methods are not possible	▫ Overcomes observed differences between treatment and comparison	▫ Assumes no unobserved differences (often implausible)

6.2. Study Design and Procedures

- The geographic region
- Research population
- Sampling frame
- Inclusion/exclusion criteria (with clear justification)
- Unit of analysis and definition of cluster (if applicable)
- Attrition criteria as applied to individual participants
- Early termination criteria for the study (e.g. the treatment works so well that the project is halted and the treatment is administered to all study participants)
- Expected timeline of the study, including a detailed description of when the intervention and data collection will take place

Flow Chart of Implementation

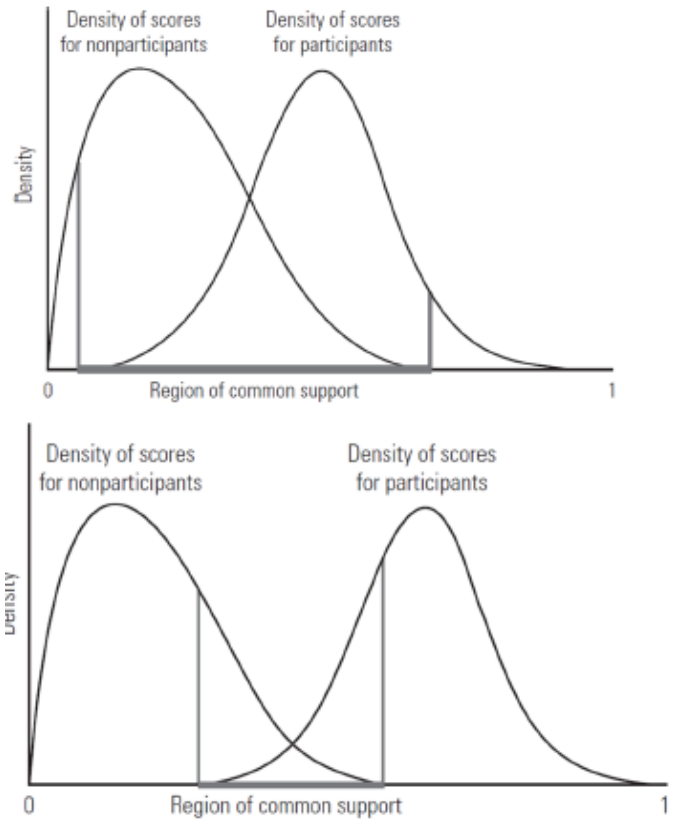


6.3. Describe Intervention in Detail

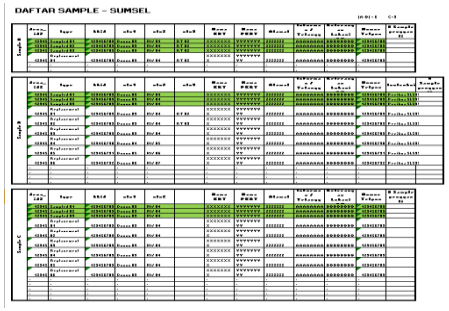
- If the intervention involves a technology or medical treatment, provide technical details for how the technology/treatment will work.
- Provide a description of how the intervention will differ between the control and treatment groups and/or between subgroups (if applicable).
- If there are several treatment arms, detail how each arm will differ from each other as well as the control group.
- Blinding:
  - Single-Blind: Subjects don't know treatment conditions
  - Double-Blind: Data collectors don't know treatment conditions
  - Triple-Blind: Analysts don't know treatment conditions
  -

6.4. Assumptions

- RCT
  - Conditional Independence; No Contamination
- Natural Experiments
  - All: “As if Random”
- Regression Discontinuity
  - No Sorting; Balance; No Polynomial Function; Survives Placebo
- Diff-in-Diff
  - Parallel Trends; Survives Placebo
- Propensity Score Matching
  - Conditional Independence; Common Support
- Instrumental Variables
  - Conditional Independence; Exclusion Criterion, Instrument Strength



6.5 Data



- What will the sources of data be?
- What is your sample frame?
- How will the data be collected (e.g., in-person interviews)?
- Provide rationale for using certain methods of data collection over others.
- If there is data used in the study that is not collected by your team (e.g., third-party administrative data), describe the source of the data.

6.6. Randomization

- If the study is a randomized controlled trial, define the randomization procedure in advance.

- How will it be done (In excel? Physical lottery? Through a statistical program such as STATA?). If there is a process with pieces of paper randomly drawn, how will the process ensure that it is not possible to interfere with randomization?

- What is the unit of randomization (e.g. individual, cluster)? Provide justification for this decision.

- What stratification variables are used (if any)?

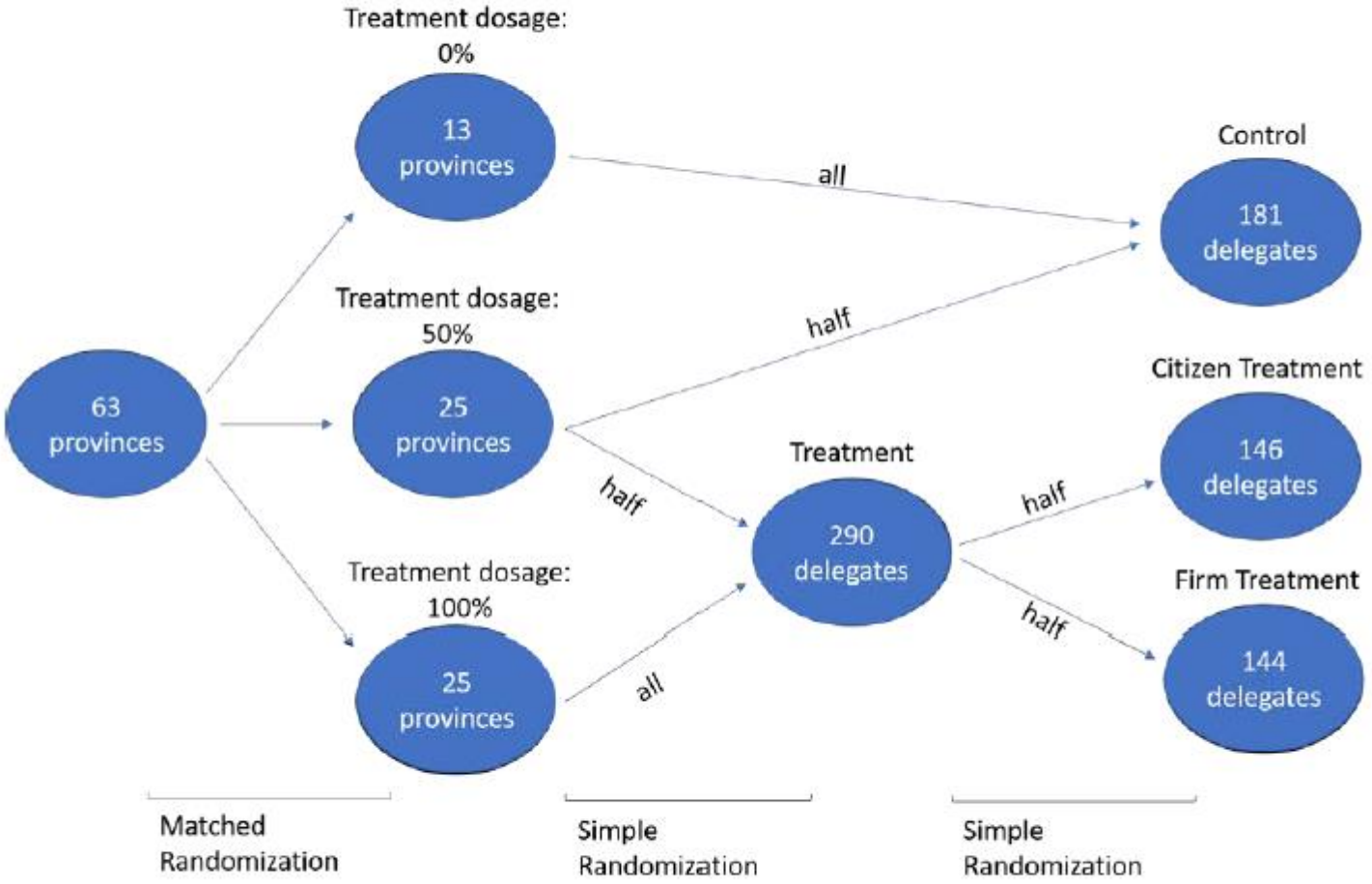


Figure 5: Randomization strategy

6.7 Power Calculations

- If budget is unknown and you are appealing for funding, calculate necessary sample size.

$$N > \frac{\sigma^2}{\left(ETE \Big/ \left(\left(t_{1-\kappa} + t_{\alpha/2} \right) \sqrt{\frac{1}{p(1-p)}} \right) \right)^2}$$

- If budget is known, calculate minimum detectable effect.

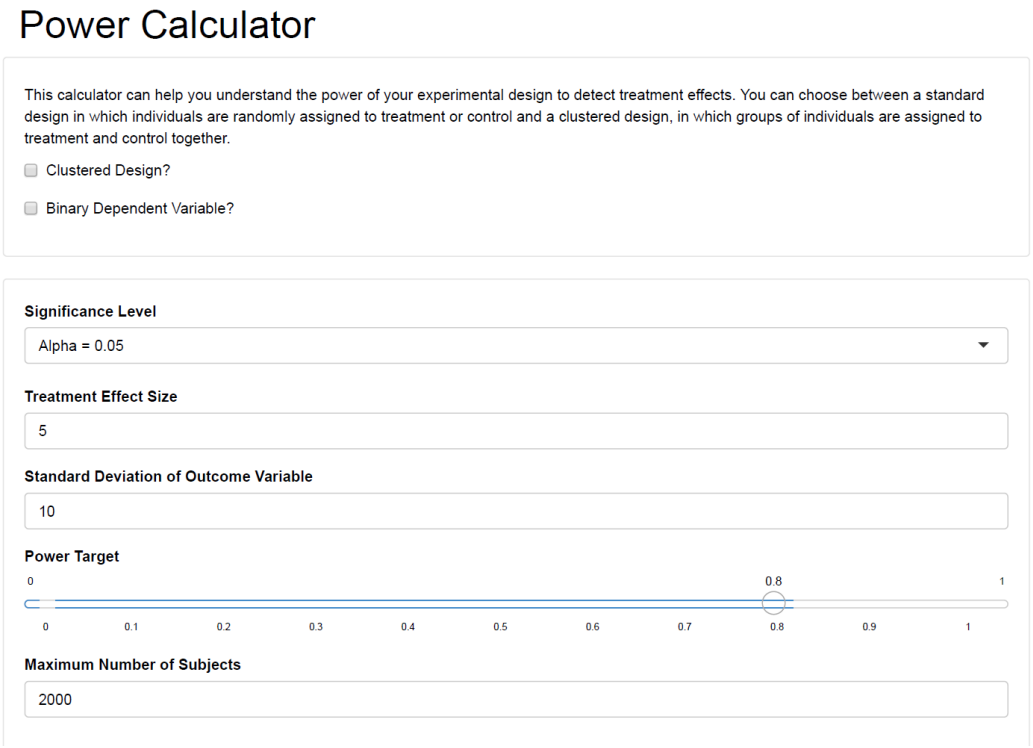
$$MDE > \left(t_{1-\kappa} + t_{\alpha/2} \right) \sqrt{\frac{1}{p(1-p)}} \sqrt{\frac{\sigma^2}{N}}$$

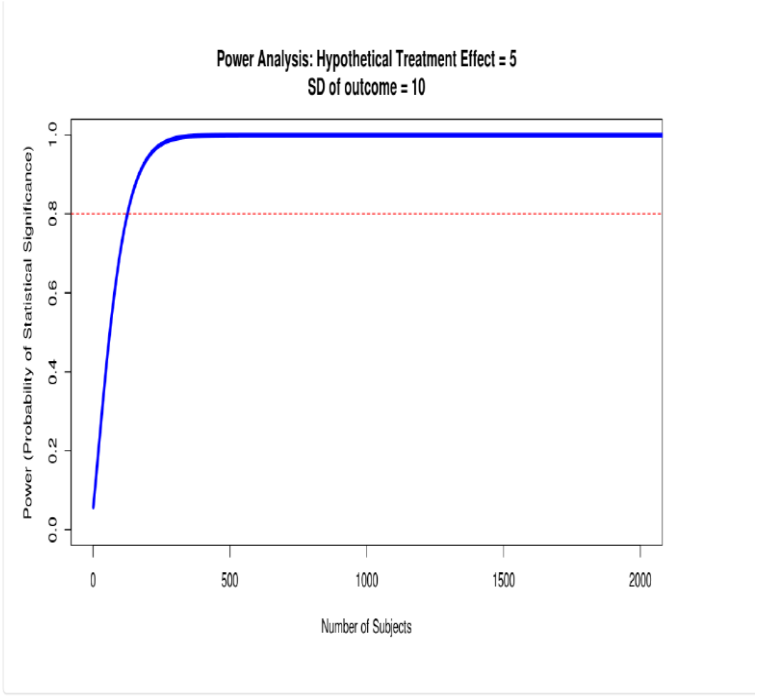
- If clustered design, remember to take into account second level.

$$MDE > \left(t_{1-\kappa} + t_{\alpha/2} \right) \sqrt{\frac{1}{p(1-p)J}} \sigma \sqrt{\rho + \frac{1-\rho}{n}}$$

6.7 Power Calculations

- Feel free to use EGAP calculator and display graphs of MDE at different levels.





https://egap.shinyapps.io/Power_Calculator/

7. Analysis

- If variables will be constructed (e.g. creating index variables), how will they be constructed?
- What are the primary model specifications? Include this in the PAP.
- If multiple hypothesis testing will be done, how will this be accounted for in the analysis?

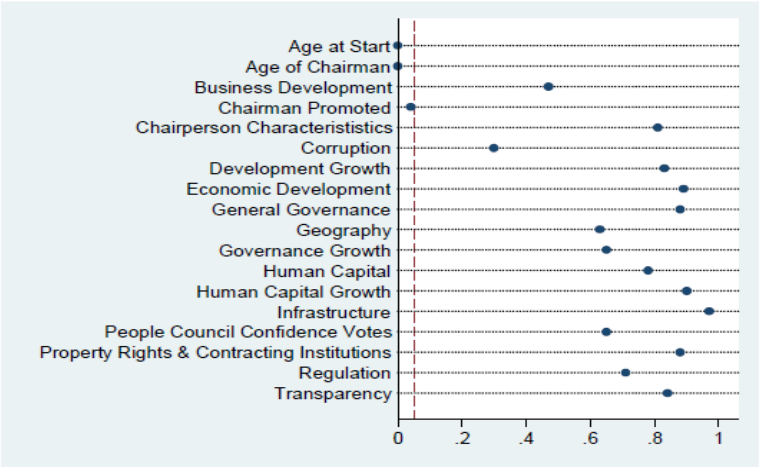
7.1. Balance

- Use preliminary or mock data to calculate prepare balance tests (table or graph)

TABLE 2. Summary Statistics

	(1) Mean in Meeting Group	(2) Difference between Plebiscite and Meeting Group	(3) *p* Value	(4) Num Obs		(1) Mean in Meeting Group	(2) Difference between Plebiscite and Meeting Group	(3) *p* Value	(4) Num Obs
Village characteristics					*Village government characteristics*				
Village population (1,000 inhabitants)	2.401	−0.295	0.625	49	Village head age	45.935	2.368	0.443	47
	[2.726]	(0.598)				[8.370]	(3.059)		
Agricultural wage (1,000 Rupiah)	21.023	−1.061	0.466	43	Village head years of education	11.645	−1.409	0.081*	47
	[5.892]	(1.443)				[2.026]	(0.788)		
Percent village roads that are asphalt	0.305	−0.042	0.507	49	Number of village head candidates in last village head election	2.207	0.304	0.432	44
	[0.269]	(0.062)				[1.013]	(0.383)		
Number of hamlets per village	4.813	−0.633	0.142	49	More than one candidate in last village head election	0.724	0.089	0.449	44
	[1.839]	(0.423)				[0.455]	(0.116)		
Number of churches and mosques per village	2.438	−0.220	0.698	49	Share of population that voted in last village head election	0.888	−0.004	0.910	43
	[1.933]	(0.563)				[0.100]	(0.031)		
Distance to subdistrict capital (km)	5.766	3.548	0.109	49	Village head's margin of victory in last election (if challenger)	0.263	−0.011	0.870	33
	[6.509]	(2.173)				[0.262]	(0.069)		
Village ethnic fragmentation	0.268	−0.075	0.190	49	Number of village government executive branch members	8.516	−0.616	0.386	47
	[0.250]	(0.056)				[2.850]	(0.703)		
Village religious fragmentation	0.106	0.011	0.827	49	Share of hamlets represented in village executive branch	0.853	0.043	0.442	47
	[0.137]	(0.051)				[0.240]	(0.056)		
					Number of people in village parliament	7.750	−0.976	0.249	36
						[3.627]	(0.832)		
Survey respondent characteristics									
Survey respondent predicted log per capita expenditure	11.505	0.034	0.602	224	Share of hamlets represented in village parliament	0.843	0.054	0.339	36
	[0.279]	(0.066)				[0.202]	(0.056)		
Survey respondent years education	8.925	−0.519	0.404	244	Number of village parliament meetings in last year	5.714	−1.853	0.041**	44
	[3.088]	(0.616)				[4.689]	(0.878)		
Survey respondent is female	0.431	0.025	0.292	245	Village parliament district system (1 = district, 0 = at large)	0.241	0.081	0.587	45
	[0.497]	(0.023)				[0.435]	(0.148)		
Survey respondent age	41.700	1.896	0.271	245	Number of previous KDP projects	1.875	−0.239	0.455	49
	[12.021]	(1.701)				[0.976]	(0.318)		
Survey respondent is farmer	0.594	−0.052	0.541	245					
	[0.493]	(0.084)							

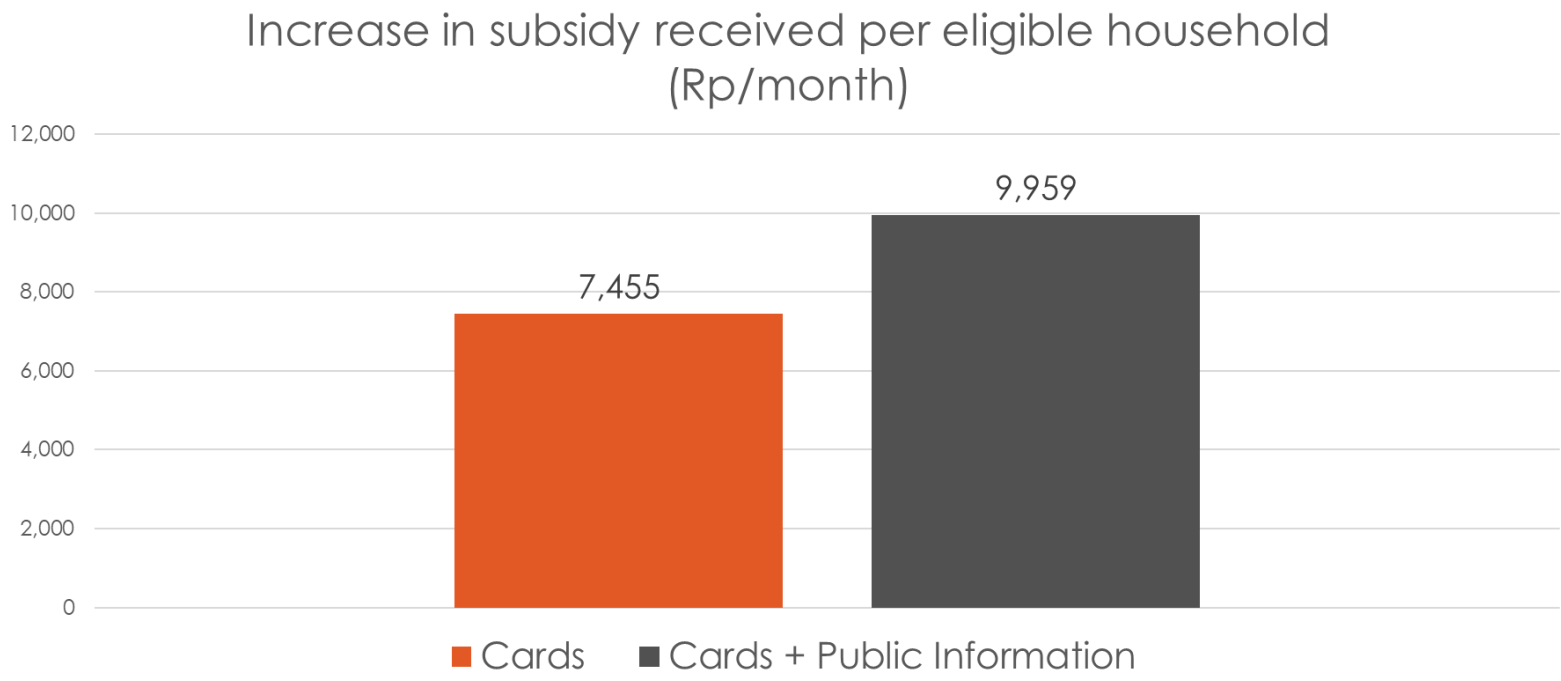
Notes: Column (1) presents the mean of the listed variable in the meeting villages, with standard deviations in brackets. Column (2) presents the difference between election and meeting villages, estimated with wave fixed effects, with robust standard errors in parentheses clustered at the village level. Column (3) shows the *p* value from a test of the null hypothesis that the listed variable is not different between elections and meeting villages. Column (4) shows the number of observations of the listed variable.
* significant at 10%; ** significant at 5%; *** significant at 1%.



Note: Blue dots represent p-values from MANOVA analyses of grouped variables. The y-axis supplies the title of each grouping. A full list of indicators under each title can be found in Online Appendix B. Dashed line represents p=.05 from the MANOVA analysis. For dots below that number, we reject the null hypothesis that the treatment and control are different on that set of criteria.

7.2. Prepare Descriptive Analysis

- Use mock data to create table or graph.



7.3. Econometric Analysis

- Pre-State Equations and Prepare Mock Tables

$$\Pr(Outcome_i = 1) = \beta_0 + \beta_1 CitizenCard_i + \beta_2 FirmCard_i$$
$$+ \beta_3 Fulltime_i + \beta_4 CNo\min ated_i + \beta_5 VoteShare_i + \lambda_p + u_i$$

Mock Table 1: Effects of Experiment on Responsiveness of Delegates

	DV: Made up Mind=1			DV: Sufficient Info.=1		
Dependent variable	No Controls	Blocking Variables	Province FE	No Controls	Blocking Variables	Province FE
	(1)	(2)	(3)	(5)	(6)	(7)
Citizen Treatment=1						
Firm Treatment=1						
Full Time=1						
Centrally Nominated=1						
Delegate Vote Share (%)						
Constant						
Provincial FE	No	No	Yes	No	No	Yes
Observations	485	485	485			
Clusters	63	63	63	63	63	63
R-Squared						
RMSE						

Linear probability model (OLS) with standard errors, clustered by province, in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Panel 1 studies whether delegates had made up their mind on Education Law in VNA library Survey. Panel 2 studies whether delegate said more constituent information was necessary .Equations 1 & 4 are unadjusted, Equations 2 & 5 control only for blocking variables, and Equations 3 & 6 introduce province fixed effects.

Mock Table 2: Effects of Experiment on Responsiveness of Delegates (Provincial Level Analysis)

	DV: Mention in Caucus=1			DV: Count in Caucus			DV: Mention on Floor=1			DV: Count on Floor		
Dependent variable	No Controls	Blocking Variables	Regional FE	No Controls	Blocking Variables	Regional FE	No Controls	Blocking Variables	Regional FE	No Controls	Blocking Variables	Regional FE
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Citizen Treatment=1												
Firm Treatment=1												
Full Time Share												
Centrally Nominated Share												
Delegate Vote Share (%)												
Provincial GDP												
Provincial Population												
Central Transfers												
National Level City=1												
Constant												
Regional FE	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Observations	63	63	63	63	63	63	63	63	63	63	63	63
R-Squared												
RMSE												

Standard errors, clustered by province, in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Panel 1 studies whether delegates speak in provincial caucus (OLS). Panel 2 studies number of speeches in caucus (Poison). Panel 3 studies whether delegates speak on floor (OLS). Panel 4 studies number of speeches on floor (Poison). Equations 1 are unadjusted, Equations 2 control only for blocking variables, and Equations 3 introduce province fixed effects.

8. Conclusion

- Potential limitations
  - Anticipate them and describe how they will be solved.
- Scaling Up and Extensions
  - What will happen if initial experiment successful
- Policy Implications
  - Describe cost benefit analysis based on minimum detectable effect.