Applied Economics PhD Report 2023

Nandan Rao

Advisor: Caterina Calsamiglia
Tutor: Francesc Trillas

December 11, 2023

Research Purpose and Scope

Developing Online Research Methodologies

Randomized control trials can be used to measure the impact of population-level policies and interventions.

By construction, randomized control trials often involve recruiting and gathering data from willing study participants, some of whom will be treated by the intervention.

Online and digital interventions are affordable to scale up quickly. Offline recruitment and data gathering, however, is both expensive and suffers from an external validity problem when attempting to transfer those results to online populations.

Developing Online Research Methodologies

My research develops methodologies for performing end-to-end survey-based research entirely online.

I seek to solve two problems that currently exist for economists across multiple sub-fields:

1. The need to evaluate the impact of an intervention which is performed online, such as an online ad campaign or a mobile app distributed online.
2. The need to run surveys affordably and quickly to gather novel data, or higher-frequency data, that can only be gather through asking people questions.

The second need becomes highly relevant in the context of the first need, however it is more general.

Strategy for Applying Methodologies

Much of the work I present in this dissertation will be based on randomized control trials for impact evaluation, highly collaborative in nature due to their significant budgetary requirements.

What they share, however, are methodologies for recruiting participants and gathering data from them entirely online, affordably and at scale. These methodologies form the core of my scientific contribution in this dissertation.

The usefulness of these methodologies will not be restricted to economics, but they will be of special interest to economists whose work often relies on quantitative data and randomized control trials to measure causal impacts.

Components of the dissertation

My dissertation will consist of three papers:

1. An RCT testing the impact of online ads on attitudes, knowledge, beliefs, and behaviors.
2. An RCT testing the impact of mobile apps on attitudes, knowledge, beliefs, and behaviors.
3. A methodology for dynamically optimizing representative survey recruitment with online advertising and results replicating important population-level social surveys.

Continuation of a Body of Work

My work in the final paper lays out new methodologies that the previous two papers use to perform real-world impact evaluations that would otherwise be difficult or cost-prohibitive.

Beyond the usefulness already proved, however, this also lays the grounds for future work that continues to build on the potential of dynamically optimizing recruitment, paralleling the recent advances in adaptive experimental design[1]. This opens the door for surveying methodologies that maintain cost-effectiveness while improving quality beyond what is possible with contemporary methods, including the ones presented here.

[1] Kasy and Sautmann 2021; Atan, Zame, and Schaar 2020.

Can Facebook Ads Prevent Malaria?

Paper 1: Can Facebook Ads Prevent Malaria?

Status: Near complete.

Presented:

- MIT CODE (Conference on Digital Experimentation / Massachusetts Institute of Technology), November 2021.
- QME (Quantitative Marketing and Economics / University of Chicago) Conference, September 2023.

Is being submitted to journals in Q1 2024.

Paper 1: Can Facebook Ads Prevent Malaria?

Problem: Digital ad campaigns are increasingly used for social good and public health objectives. However, effectiveness measures are limited to direct response metrics (clicks, likes) or short-term "brand-lift" surveys which are limited in scope[2].

Ideally we would have both random assignments of ads and measures of long-term behavior disconnected from the ad itself.

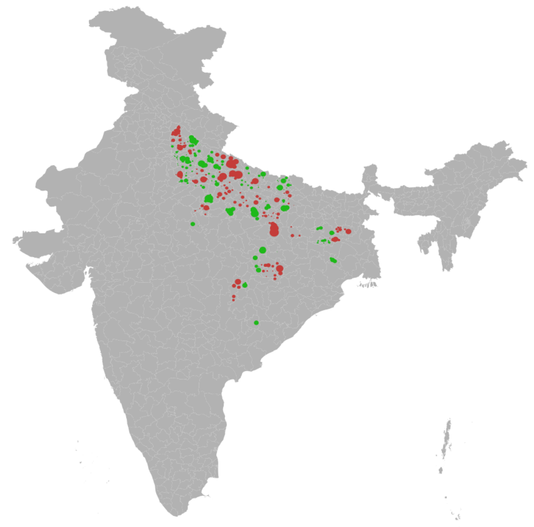
[2] Shawky et al. 2019; Athey et al. 2023.

Paper 1: Can Facebook Ads Prevent Malaria?

This paper contains two randomized control trials:

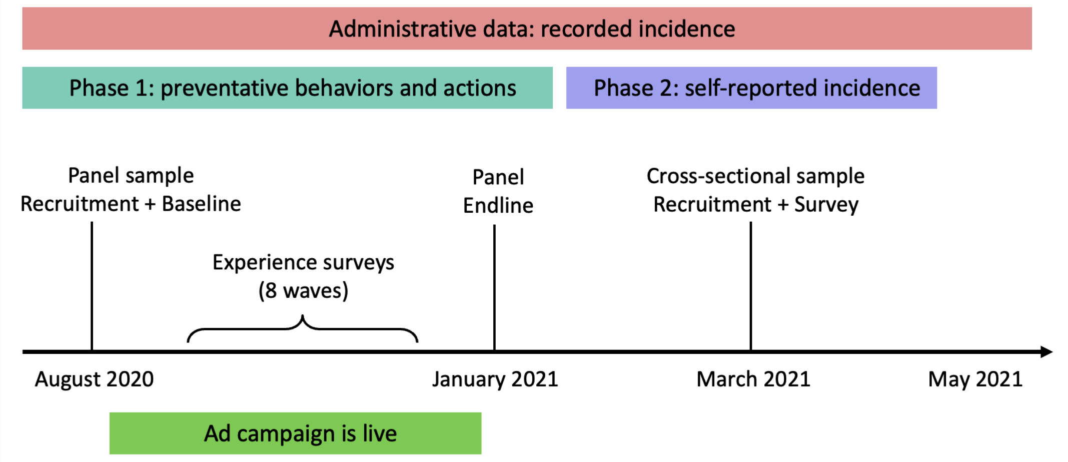
1. Cluster-randomized trial with repeated panel for the duration of the actual, full-scale campaign.
2. Individually-randomized field experiment to test the impact of the ad content from the actual campaign.

Paper 1: Can Facebook Ads Prevent Malaria?



80 districts: 40 treatment and 40 control. Used circles for Facebook targeting, fit within districts centered around populated areas. Control districts will be withheld ("punched out") of the campaign.

Paper 1: Can Facebook Ads Prevent Malaria?



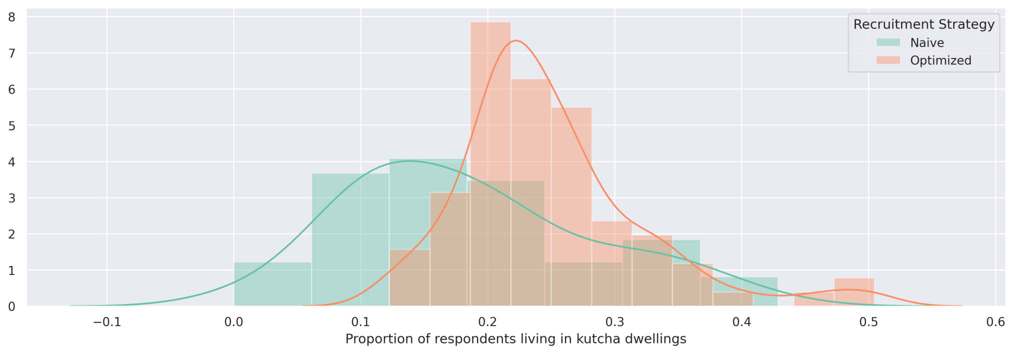
Paper 1: Can Facebook Ads Prevent Malaria?

Problem: Those most at risk for malaria are underrepresented on social media. In particular: those living in non-solid dwellings have malaria incidence at twice the rate of those in solid dwellings.

Solution: Stratify our recruitment by dwelling-type (2), in addition to district (80), creating 160 strata.

Paper 1: Can Facebook Ads Prevent Malaria?

By stratifying by dwelling-type, using Facebook's "Lookalike" audience to target the group, we increase the representivity of that otherwise underrepresented group and ensure we can measure heterogenous effects across this variable of interest:



Paper 1: Can Facebook Ads Prevent Malaria?

We look at several outcomes, but two of special interest:

1. Self-reported behavior, recorded in the repeated panel survey, about whether the reporting individual slept under a mosquito net the previous night, and how many of their household members as well.
2. Administrative data on malaria incidence, collected at district level.

Paper 1: Can Facebook Ads Prevent Malaria?

Table 2: Sleeping under bednet (September 2020 - January 2021)

	Share of times respondent used bednet				Share of HH members who used bednet			
	(1) Full	(2) Full	(3) Full	(4) Solid house	(5) Full	(6) Full	(7) Full	(8) Solid house
T	0.033	0.024*	-0.001	0.051**	0.028	0.022*	-0.017	0.064***
	(0.021)	(0.014)	(0.022)	(0.025)	(0.018)	(0.011)	(0.015)	(0.018)
T × Solid house			0.053				0.082***	
			(0.038)				(0.025)	
Solid house		-0.020	-0.046			-0.042***	-0.082***	
		(0.019)	(0.028)			(0.013)	(0.018)	
Individual controls		✓	✓	✓		✓	✓	✓
Cluster controls		✓	✓	✓		✓	✓	✓
Observations	3900	3900	3900	1789	3900	3900	3900	1789
Adj. R-squared	0.001	0.026	0.027	0.035	0.001	0.047	0.050	0.048
Mean in control	0.634	0.634	0.634	0.603	0.661	0.661	0.661	0.610
SD in control	0.423	0.423	0.423	0.431	0.344	0.344	0.344	0.361

Each observation is an individual. Observations collected from September 1, 2020 to January 31, 2021 are included and are weighted by the number of survey waves the individual completed in this period. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete (i.e., pucca). Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$"

Paper 1: Can Facebook Ads Prevent Malaria?

Table 10: The impact of MNM campaign on malaria incidence rate (cases/1K people)

	Overall incidence		Urban incidence		Rural incidence	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated*Post	-0.0041	-0.0046	-0.0061*	-0.0066**	-0.0008	-0.0008
	(0.0055)	(0.0053)	(0.0035)	(0.0031)	(0.0051)	(0.0053)
District FE	✓	✓	✓	✓	✓	✓
Month and year FE	✓	✓	✓	✓	✓	✓
Subdistrict controls	✓	✓	✓	✓	✓	✓
Cluster controls		✓		✓		✓
Observations	1498	1498	793	793	1338	1338
Adj. R-squared	0.181	0.178	0.446	0.449	0.145	0.142
Treatment Mean at baseline	0.0313	0.0313	0.0160	0.0160	0.0264	0.0264
Treatment SD at baseline	0.0989	0.0989	0.0409	0.0409	0.0982	0.0982

Each observation is a subdistrict-month-year. The panel spans April 2020 to May 2021. Post takes value 1 after August 2020, i.e. when the Facebook campaign started. Standard errors are clustered at the district level. Subdistrict controls include: subdistrict area, share of urban population, (log) elevation, (log) terrain ruggedness and (log) distance to the state capital, all interacted with Post. Cluster controls include baseline survey values of: shares of respondents living in kutcha dwellings, share of respondents with university degree, share of unemployed respondents, share of respondents sleeping under mosquito nets, 5-year malaria prevalence, 2-week malaria prevalence, all interacted with Post. State indicators interacted with Post are included in all regressions. Data come from the Indian Health Management Information System. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In urban areas (where 85% of dwellings are solid), we see a decrease from 16/million to 10/million, or about a 40% improvement.

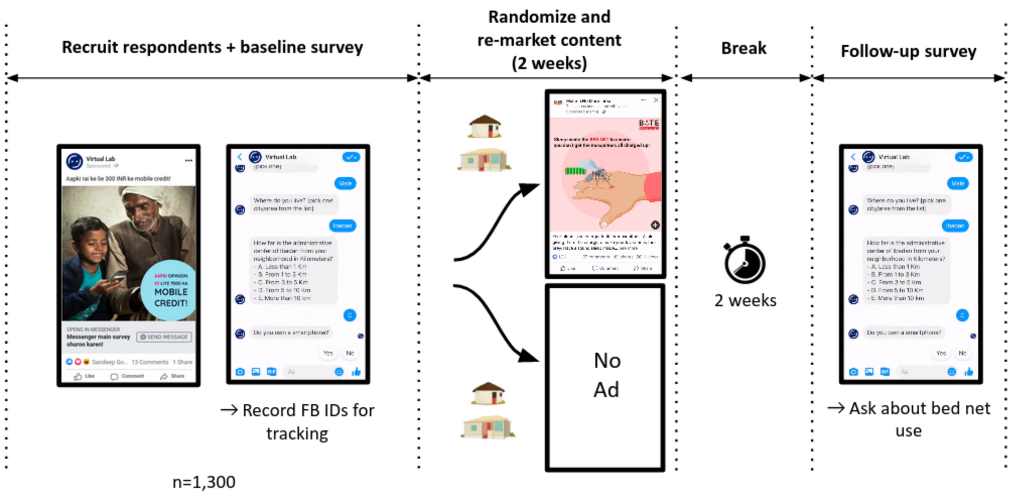
Paper 1: Can Facebook Ads Prevent Malaria?

So we know the ad campaign had an impact on those in solid dwellings and those in urban areas. The impact was significant and cost-effective (estimated at $4.50 per case avoided).

But why did it not impact those in non-solid dwellings? Was it the material, the targetting, or is it not possible to reach them via social media?

To measure the effect of the material itself, we perform a second, individually-randomized experiment.

Paper 1: Can Facebook Ads Prevent Malaria?



Paper 1: Can Facebook Ads Prevent Malaria?

Table 9: Sleeping under bednet (Individual Effect Study)

Y=1 if respondent has slept under bednet last night

	Any dwelling type			Solid house	Non-Solid house
	(1)	(2)	(3)	(4)	(5)
Individual T	0.069**	0.078**	0.072*	0.083*	0.071*
	(0.031)	(0.031)	(0.041)	(0.047)	(0.042)
Individual T × Solid house			0.013		
			(0.062)		
Solid house	-0.115***	-0.085**	-0.092*		
	(0.031)	(0.034)	(0.049)		
T	0.007	0.010	0.010	0.073	-0.039
	(0.031)	(0.031)	(0.031)	(0.048)	(0.041)
Individual controls		✓	✓	✓	✓
Observations	902	902	902	436	466
Adj. R-squared	0.017	0.033	0.032	0.013	0.043
Mean in control	0.657	0.657	0.657	0.594	0.711
SD in control	0.475	0.475	0.475	0.492	0.454

Each observation is an individual. Indivitual T is the individual treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete (i.e., pucca). Heteroskedasticity-robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

When we run campaigns targeted directly at individuals in a balanced sample, we find significant treatment effects across the entire population, and subgroup estimates are of similar size.

Paper 1: Can Facebook Ads Prevent Malaria?

We introduce two novel methods for measuring the impact of an ad campaign: one with stratified sampling and repeated panel surveys ("experience sampling") to measure behavior across groups. The other using remarketing techniques to test the impacts of a real ad campaign on a specific group of individuals.

We find that this content was effective in changing behavior and that the behavior led to impacts of malaria incidence measurable from administrative data. The method was cost-effective, but did not impact the most vulnerable. However, further results show it could have impacted the most vulnerable with more sophisticated (and costly) targeting.

Measuring the Impact of Apps

Paper 2: Apps in Development Contexts

Status: In Progress

Data Collection:

1. First RCT finished: N=1800, Serbia/Bulgaria, initial analysis finished. Working paper submitted for review. Partner: UNICEF.
2. Second RCT in progress: N=4800, Nigeria, Senegal, Bangladesh. Third pilot under way now. Full study launches Q1 2024. Partner: The World Bank.

These studies will be two separate papers and the one included in my dissertation will be determined after they are finished. They have many methodological similarities.

Estimated Finish Date: November 2024

Paper 2: Apps in Development Contexts

Both studies share a similar design:

1. Outcomes of interest are measured both before and after treatment, via online surveys or assessments.
2. An intent-to-treat (ITT) effect will be measured from the impact of random assignment on the difference in outcome before and after treatment (DinD).
3. Takeup among the treated will be measured objectively with app usage data linked to recruitment surveys.
4. Studies are subject to two-sided non-compliance, however takeup is assumed to be monotonically positive under treatment.
5. The treatment effect on the treated (ToT) will be measured with a two-staged least squares model using random assignment as the instrumental variable.

This leads to the following regression model for the ITT effect:

$$y_i - y_i^b = \gamma_1 + \beta T_i + \gamma_2 X_i + \epsilon_i$$

Where y_i represents the outcome of interest for individual i measured after treatment, T_i represents the random treatment assignment, X_i a set of control variables and y_i^b represents the outcome of interest measured before treatment. The parameter of interest will be the treatment effect, β.

Paper 2: Apps in Development Contexts

And the following two-stage model for the ToT effect:

$$y_i - y_i^b = \gamma_1 + \beta \hat{z}_i + \gamma_2 X_i + \epsilon_i$$
$$z_i = \gamma_3 + \gamma_4 T_i + \gamma_5 X_i + \delta_i$$

Where z_i is a binary indicator of takeup based on the recorded app-usage data and $\hat{z}_i$ the predicted takeup based on the first stage regression. Once again, parameter of interest is β

Paper 2: Study 1: Teaching Parenting Skills

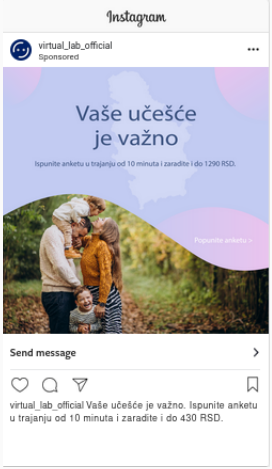
Paper 2: Study 1: Teaching Parenting Skills

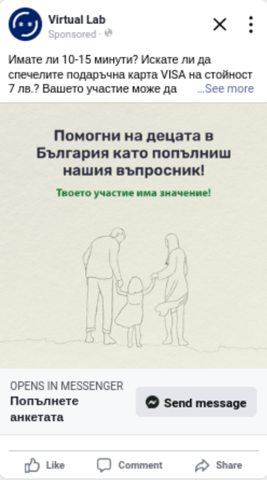
Bebbo is an app for parents, developed by UNICEF. They had no rigorous way of measuring the effectiveness of the app on actually changing parenting knowledge, attitudes, and behavior.

We designed an RCT that could be performed entirely online, in two different countries, to measure the impact the app has on parents who use it.

We developed ads run on social media to recruit study participants in Bulgaria and Serbia.

Recruitment Ads:





Paper 2: Study 1: Teaching Parenting Skills

We recruited 2616 pts in Serbia and 1689 in Bulgaria. Upon recruitment, respondents were administered the baseline survey and the treatment group was invited to download the app via the provided link.

After approximately 4-6 weeks, we resurveyed everyone again. 1327 in Serbia and 750 in Bulgaria completed the second survey, for a retention rate of 50% and 45% respectively.

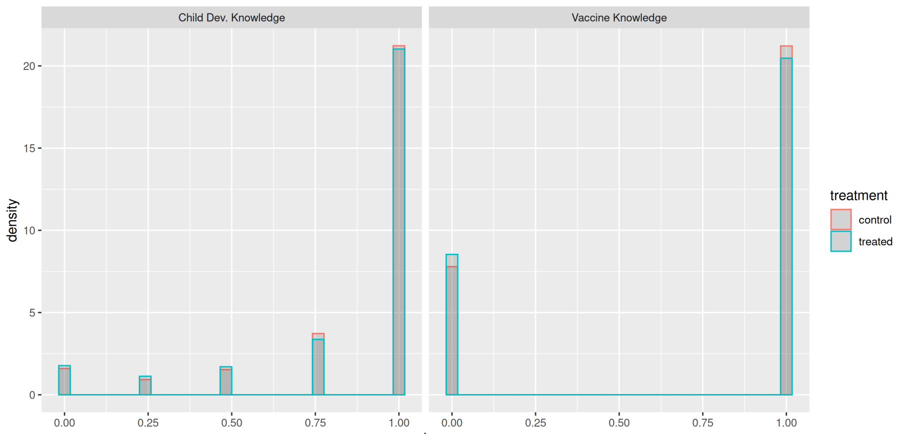
Paper 2: Study 1: Teaching Parenting Skills

We measure 8 distinct outcomes for the experiment, covering knowledge, attitudes, and behavior. After controlling for multiple testing, we do not find any significant impact of treatment, encouragement of downloading the app, Bebbo.

1. Low takeup (28%). Potentially driven by high pre-exposure (55% knew, 23% had used).
2. Relatively low usage among takeup (12% used more than a day).
3. Ceiling effects on some of the outcomes.
4. Survey effects on knowledge outcomes.

Paper 2: Study 1: Teaching Parenting Skills

One of the primary outcomes of interest was the ability for the treatment to impact parenting knowledge. Unfortunately, the questions that made up the two indices in the survey instrument both suffered from quite serious ceiling effects:



Paper 2: Study 1: Teaching Parenting Skills

Table 1: Pooled: OLS - Endline - Knowledge and Awareness

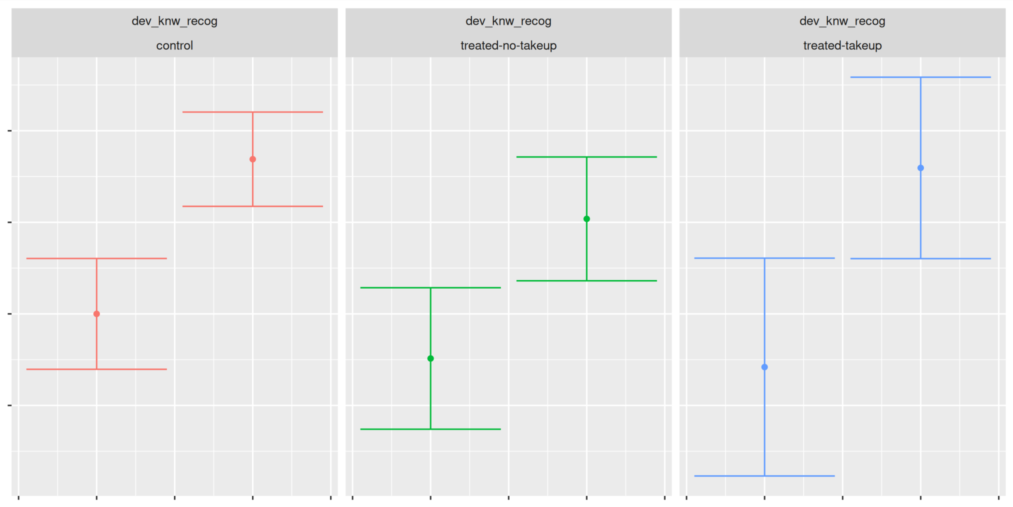
	Dependent variable:	
	Vaccine Knowledge	Child Dev. Knowledge
	(1)	(2)
Treatment	0.03	−0.003
	(0.03)	(0.01)
Adjusted Treatment p-value	0.649	0.793
Observations	691	1,918
R^2	0.01	0.01
Note:		*p<0.1; **p<0.05; ***p<0.01

Regressions do not show any significant effect on knowledge outcomes, outcomes that are most likely to be impacted by the treatment in short time.

To understand what is happening, it can be instructive to look at the raw pre- and post-treatment results for three groups: control, treatment without takeup, treatment with takeup.

What we find is that all three groups improve in their scores between baseline and endline surveys, indicating a potential survey effect: respondents are asked about their knowledge about particular aspects of parenting, and if they don't know, they may go seek the knowledge. Those who download the app might seek the knowledge through the app, but those in the other groups find other means to do the same.

Paper 2: Study 1: Teaching Parenting Skills



Paper 2: Study 1: Teaching Parenting Skills

There is some suggestive evidence that those who choose to download and use the app may be less knowledgeable, in particular regarding vaccines.

However, due to the study design, our ability to measure the impact of the app on parenting knowledge is masked by both the potential impacts of the survey itself along with strong ceiling effects.

Second RCT: Mobile Apps for Early Childhood Literacy

The purpose of this study is to measure the impact of an app for early-childhood literacy developed by the organization Curious Learning. In-field, in-person RCTs have proven effectiveness, but the question remains as to what the effect might be at scale, when delivered online.

Paper 2: Study 2: Teaching Kids to Read

This is designed as a multi-country randomized control trial across three countrie: Nigeria, Senegal, and Bangladesh.

We will recruit 1600 parents with young children in each country to participate in our study. Each parent will pick a child and administer them an online reading proficiency test.

After the first test, the treatment group will be invited to download the reading app and encouraged to offer it to their kid. The control group, "business as usual," will be encouraged to read to their children.

Paper 2: Study 2: Teaching Kids to Read

Parents will be recruited by targeting them with social media ads and offering an incentive. The sample will be representative across geographic by optimizing ad spend across urban and rural geographics to ensure proper representation.

The survey will take place in a chatbot. The chatbot is integrated with the assessment such that high scores will be excluded from the survey, and such that downloading can be tracked in the app usage data.

Assessments will be provided at the beginning and end of a 2-week period starting from initial recruitment.

Paper 3: Recruiting Survey Participants Online

Paper 3: Recruiting Survey Participants Online

Status: In Progress

Data Collection: Starts February 2024.

Estimated Finish Date: December 2024.

Presented:

- MIT CODE (Conference on Digital Experimentation / Massachusetts Institute of Technology), November 2021.
- QME (Quantitative Marketing and Economics / University of Chicago) Conference, September 2023.

Paper 3: Recruiting Survey Participants Online

This paper shows that the process of recruiting a realistic sample to estimate a population parameter of interest with a survey questionnaire can be formulated as an optimization problem subject to budget constraints.

We show the usefulness of this formulation by recruiting a sample online using this methodology. We show that it is both significantly cheaper and faster than traditional methods. We ask our sample select questions of interest from the General Social Survey (GSS), the American Time Use Survey (ATUS), and the Current Population Survey (CPS). We compare our results to the gold standard surveys.

Paper 3: Recruiting Survey Participants Online

Problems:

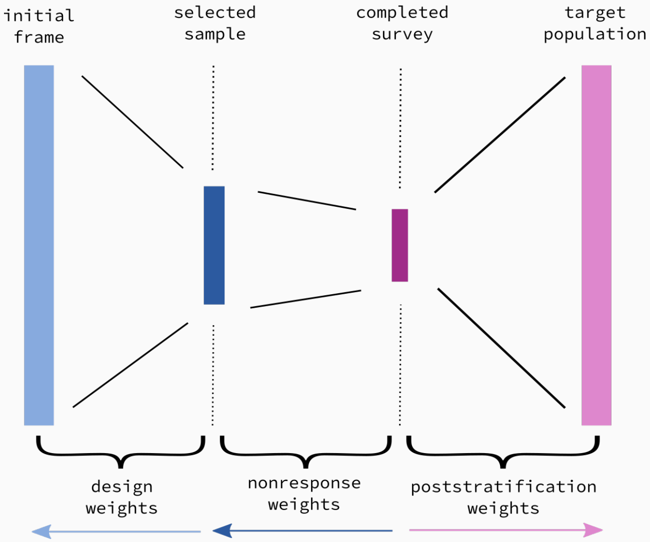
1. Non-response rates are rising for all survey methodologies.
2. Random Digit Dialing was an innovation from the 60s and 70s that brought probabilistic sampling frames at a much cheaper price than previous methodologies, but non-response rates have now reached 91%.
3. Companies build survey “panels” with RDD that can be used for research, however they are not cheap and not available in all places - either all countries or in specific regions.

Paper 3: Recruiting Survey Participants Online

Opportunities:

1. Digital advertising can be used to micro-target ads at people all over the world, in countries and specific regions.
2. Advances in post-stratification weighting have proven the ability to approach similar estimates and probabilistic samples with convenience samples.

Motivation - weighting



Setup

Poststratification weighting can be done in many ways, but we will consider the simplest case.

We want to estimate e a population parameter Y via sampling and measurement. We will assume that the researcher wishes to use the stratified mean for a set of strata $h \in H$ (mutually exclusive cells) with an assigned weight for each stratum W_h, which we will denote $\bar{y}$:

$$\hat{Y} := \sum_h W_h \bar{y}_h$$

Idea

1. Commit to poststratification weighting.
2. Measure response rate dynamically during the surveying process
3. Adjust the selected sample dynamically during the surveying process (dynamic over/under-sampling).
4. Make the adjustment to minimize variance subject to budget constraints.

Setup

The variance of our sample estimate is given by:

$$\mathbb{V}[\hat{Y}] = \sum_h W_h^2 \frac{s_h^2}{n_h}$$

where s_h^2 denotes the variance of the population parameter of interest Y within stratum h. If the outcome was measured during recruitment, we can estimate this stratum-specific variance.

Setup

In practice, many researchers are interested in multiple outcomes, thus the individual variance of each outcome may differ in different ways across strata. While the ideal number of strata can be optimized (TODO!) with a particular outcome, we start with a more general solution by assuming that the variance of the outcome in each stratum is equal (i.e. $s_h^2 = s^2$). Thus:

$$\mathbb{V}[\hat{Y}] = s^2 \sum_h \frac{W_h^2}{n_h}$$

Note that, given a fixed n and the assumption of equal variance across strata, this quantity is minimized when $\frac{n_h}{n} = W_h$, known as the Neyman allocation.

Setup

But we don't have infinite moneys...

Setup

Denote the cost to recruit an individual from stratum h as P_h.

Denote total budget B, such that $B_h \leq P_h n_h$.

Denote desired maximum sample size N_d.

We can then frame the optimization problem of finding the best allocation of budget to minimize the variance of the final estimate as:

$$\underset{n_i,\ldots,n_h}{\text{argmin}} \sum_h \frac{W_h^2}{n_h}$$

$$s.t. \sum_h P_h n_h \leq B$$

$$\sum_h n_h \leq N_d$$

Setup

But we don't know the price per respondent...

How to measure cost?

We can model the inverse cost ($\frac{1}{P_h}$), the number of respondents recruited n_{ht} given budget spend B_h, as a Poisson random variable:

$$\begin{aligned} n_{ht} | B_{ht} &\sim Poisson(\lambda) \\ \lambda &\sim Gamma(\kappa, \beta) \end{aligned}$$

We can use closed-form Bayesian updating to obtain a MAP estimator of λ and the implied mean of the predictive distribution $(1/\lambda)$.

Proving it out

To prove this works, we pick a few questions from the General Social Survey (GSS), the American Time Use Survey (ATUS), and the Current Population Survey (CPS). We will recruit a stratified sample of 1000 adults in the U.S. and ask them the same questions.

We will then analyze the difference in both the sample means and variances for all questions, comparing the difference in our outcomes with the gold standard outcomes, relative to the margin of error in the gold standard surveys.

Research Activities and Go-Forward Timeline

Educational Activites: Seminars

I will participate in seminars in 2024 and 2025, in particular:

1. Online UAB department seminars
2. In-person UAB department seminars (when present)
3. Online global seminars (Causal Inference Seminar, CEPR)
4. In-person department seminars at visiting universities

Educational Activites: International Research Stays

I will engage in two international research stays in the spring of 2024 and again in the academic year 2024-2025.

Places: TBD.

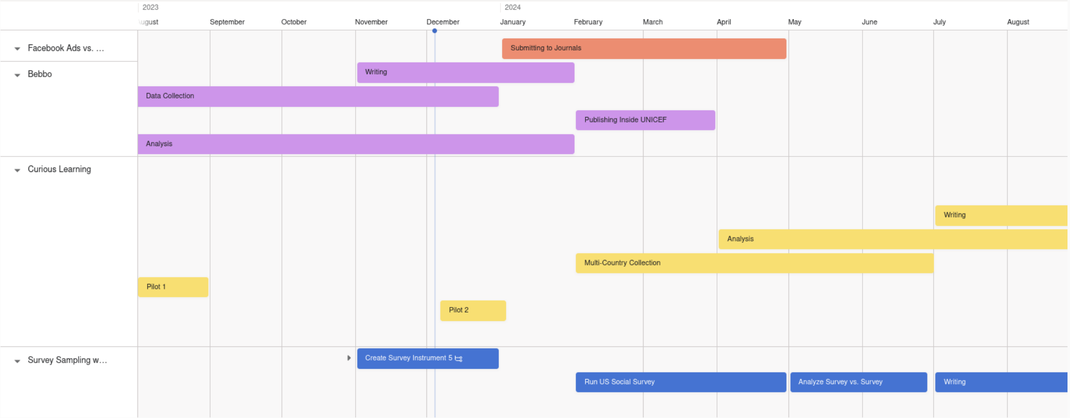
Educational Activites: Summer Courses

I will join the UAB Summer Course in the summer of 2024.

Primary Research

As discussed, my own research consists of (4) papers, 3 of which will be included in my PhD dissertation. What follows is a work timeline with the major stages of each project.

Timeline: 2023-2024



Timeline: 2024-205