Don’t Follow the Crowd: Incentives for Directed Spatial Sampling

Nathan Jensen[1], Russell Toth[2], Yexiang Xue[3], Rich Bernstein[4], Eddy Chebelyon[5], Andrew Mude[6], Christopher Barrett[7], and Carla Gomes[8]

Selected Paper prepared for presentation at the 2017 Agricultural & Applied Economics Association Annual Meeting, Chicago, Illinois, July 30-August 1

[1] Mara House, International Livestock Research Institute, 30709 Naivasha Rd., Nairobi, Kenya. n.jensen@cgiar.org
[2] Room 363, Merewether H04, The University of Sydney, NSW 2006 Australia. russell.toth@sydney.edu.au.
[3] 344 Gates Hall, Cornell University, Ithaca, NY, USA. yexiang@cs.cornell.edu
[4] 344 Gates Hall, Cornell University, Ithaca, NY, USA. rab38@cornell.edu
[5] 9500 Gilman Drive, University of California San Diego, La Jolla, CA, USA. echebelyon@ucsd.edu
[6] Mara House, International Livestock Research Institute, 30709 Naivasha Rd., Nairobi, Kenya. a.mude@cgiar.org
[7] 301G Warren Hall, Cornell University, NY, USA. cbb2@cornell.edu.
[8] 353 Gates Hall, Cornell University, Ithaca, NY, USA gomes@cs.cornell.edu

Don't Follow the Crowd: Incentives for Directed Spatial Sampling[1]

Nathan Jensen[2], Russell Toth[3], Yexiang Xue[4], Rich Bernstein[5], Eddy Chebelyon[6], Andrew Mude[7], Christopher Barrett[8], and Carla Gomes[9]

March 2016

Abstract: The rapid spread of mobile phones globally provides new opportunities to collect data directly from citizens to inform decision making. However such "crowdsourcing" can suffer from a number of sampling problems. We address the issue of sub-optimal sampling of citizen reports through a new method that uses small incentives to induce respondents' willingness to supply information of a given quality, enabling a social planner to induce a more optimal sampling distribution. We report the results of a field experiment that sought to induce pastoralists that live in the drought-prone rangelands of central Kenya to provide reports on localized vegetation conditions. We begin by estimating the reward elasticity of response to spatially-varying incentives, which is positive and shows that individuals are willing to shift their submission patterns in response to small spatial and temporal variation in rewards for their submissions. We then show that our dynamic rewards successfully increase submissions from regions that were initially under sampled. Finally, we use a value of information model to determine that the dynamic rewards led to increased net value of information for cost, as the net effect of our dynamic reward structure was to reduce costly redundant submissions and increase more informative submissions from never or rarely visited locations. Our results have implications for a range of applications of "big data" sources in the developing and developed world.

***** PRELIMINARY VERSION: PLEASE DO NOT CITE OR CIRCULATE *****

[1] First version: December 2015.We thank seminar audiences at the Sydney Experimental Brownbag Seminar (University of New South Wales), University of California at San Diego (BDES), the Helmholtz-Centre for Environmental Research (Leipzig, Germany), the International Livestock Research Institute, and Cornell University for helpful comments. We acknowledge research collaboration and fieldwork support from colleagues at the International Livestock Research Institute (Nairobi, Kenya), Cornell University and UC San Diego. We acknowledge the Atkinson Centre for a Sustainable Future's Academic Venture Fund and Australian Aid. This research was also supported by National Science Foundation (0832782, 1059284, 1522054), ARO grant W911-NF-14-1-0498. The authors thank the International Livestock Research Institute for generous hospitality. The usual disclaimer applies.
[2] Mara House, International Livestock Research Institute, 30709 Naivasha Rd., Nairobi, Kenya. n.jensen@cgiar.org
[3] Room 363, Merewether H04, The University of Sydney, NSW 2006 Australia. russell.toth@sydney.edu.au.
[4] 344 Gates Hall, Cornell University, Ithaca, NY, USA. yexiang@cs.cornell.edu
[5] 344 Gates Hall, Cornell University, Ithaca, NY, USA. rab38@cornell.edu
[6] 9500 Gilman Drive, University of California San Diego, La Jolla, CA, USA. echebelyon@ucsd.edu
[7] Mara House, International Livestock Research Institute, 30709 Naivasha Rd., Nairobi, Kenya. a.mude@cgiar.org
[8] 301G Warren Hall, Cornell University, NY, USA. cbb2@cornell.edu.
[9] 353 Gates Hall, Cornell University, Ithaca, NY, USA gomes@cs.cornell.edu

Introduction

The rapid and global spread of mobile phones provides new opportunities for low-cost data collection. There are close to 4 billion unique mobile phone users in the world, with about 83% of those connections in the developing world (GSM Assocation, 2015), with access now expanding rapidly even in rural and remote areas. This provides opportunities both for the passive collection of administrative call data records, and directly tapping the local knowledge and information of users, who are spatially dispersed and hence well-placed to generate reports at places and times that are otherwise costly to sample. The latter could include anything from reporting on local traffic conditions and accidents, to online ratings of restaurants and hotels, to providing information about conditions on the ground during the onset of disasters. Such information could be particularly valuable for improving access to information and decision making in the developing world, where gaps in infrastructure and government capacity particularly hinder the collection of useful data for decision making.

Mobile phones raise promising opportunities to collect new data that could inform policy and business decision making. However efforts that rely on contributions from volunteers, commonly known as 'crowdsourcing', generally face a crucial trade-off: reduced cost of collecting information in return for reduced control over data collection. For many of the applications that utilize such data, we are not concerned with the goal of representative sampling, rather to have sufficient data on target events. For example, in an application of crowdsourcing to traffic accidents we would not be seeking a representative set of reports from all drivers in the city. Rather we would simply want a sufficient density of reports on each accident, which is an idiosyncratic set of localized events at any given time. The crowdsourcing problem, then, is often less about uniformity of sampling than simply ensuring that we receive sufficient reports on each event of interest. This may take on a temporal dimension, if we want to track the persistence of the event over time. A similar intuition would hold for events from crimes, to restaurant quality, to disasters.

In this paper we present the results of a field experiment that demonstrates a method to quantify in monetary terms the extent to which respondents are willing to provide desired reports. These estimates could be used to design an incentive system that generates a closer-to-optimal sampling distribution for a given budget, by nudging users to provide more useful reports through monetary rewards. Specifically, we focus on a spatial-sampling problem in a drought-prone region of central Kenya, where it would be desirable for a variety of policy applications to have better micro-data on vegetation conditions 'on the ground,' particularly to identify "bad events" (locations with particularly poor conditions). These conditions are a leading indicator for localized risks of welfare and economic losses due to drought, the major source of humanitarian risk in this setting.. Existing approaches to collect such data face drawbacks in terms of cost (i.e., data collection by scientists) or accuracy (i.e., remote sensing data from satellites, which are often aggregated at the level of tiles that can be thousands of square meters in size).

We consider the possibility of collecting data on localized vegetation conditions at high spatial and temporal resolution from local citizens through smartphones. We ask participants to characterize the forage conditions throughout their day using a short, visually-oriented and geo-located survey, and then take a picture of the location they are characterizing. The crowdsourcing challenge in this setting is that participants are likely to provide highly clustered observations while landscape-scale forage modeling requires more data from locations where participants visit under typical conditions and from locations that are less frequently visited, which might be used drought periods.

The project was implemented among pastoralists in a drought-prone region in northern Kenya. We use a baseline treatment, which offers participants a uniform payment ("piece rate") per report. Such an intervention should not incentivize respondents to deviate from their natural movement patterns, identifying locations that are frequently and rarely visited at baseline. We then employ a series of randomized incentive treatments over the course of 12 approximately 9-day periods. We randomize 113 participants into one of six treatment groups across five training sites (villages). In each weekly period, two of the treatment groups face the same uniform piece rate incentives as in baseline. This allows us to control for temporal variation in movement patterns and participation. The other four treatment groups face a version of spatially-varying incentives, which are implemented and reported to participants through digital maps that attach different rewards to reports from different locations. Through this approach we receive about 95,000 valid survey submissions over the treatment period.

First, we estimate an individual level 'reward elasticity of supply' – the percent increase in the propensity of each individual to increase submissions in a region due to an increase rewards attached to that region. The estimates are identified off the extent to which participants change the location from which they make submissions when under spatially-varying incentives than under uniform incentives. Nearly all participants express positive elasticities, reflecting a willingness to change their location in response to small changes in rewards. We then look at the sample average elasticity, as an indicator of the elasticity of the aggregate distribution of submissions, finding that at the median reward level, a 50% increase in rewards leads to approximately a 13.6% increase in (aggregate) submissions from that region.

We then estimate the aggregate 'reward elasticity of supply', capturing the extent to which the aggregate spatial distribution of submissions responds to relatively small changes to incentives. Such a parameter estimate is a crucial input to designing an optimal spatial incentive scheme for the key application we have in mind, and for a much broader class of applications. To do so, we identify those regions that are under sampled at baseline. We find that adjusting the rewards to favor under-sampled regions increases the likelihood that a submission is from an under sampled region by a factor of at least 1.5 and that the *number* of reports per day from under sampled regions more than doubles under spatially-varying incentives. While these results are based on the starkest definition of under sampled – seeing zero submissions at baseline – we additionally show that the results are robust to a range of approaches to defining under sampled regions.

In addition to these main results, we provide an alternative way to quantify the effects of spatially varying incentives through an ad hoc value of information calculation. The idea is to suppose that the rewards assigned to regions are proportionate to the value of obtaining surveys from those regions, as would be the case under an incentive scheme for a particular application. The value of submissions in each location is discounted at a rate inversely proportional to the number of submissions previously recorded at that location on that day. Under this metric we find that spatially-varying incentives lead to a 63% increase in the value-to-cost ratio of the information collected.

Taken together, these results provide a proof-of-concept that large-scale crowdsourcing of potentially policy-relevant information from low-literacy/low-education, remote populations can be feasible. Furthermore, we show that participants naturally respond to incentives, increasing our confidence that optimally-designed incentive schemes could be used to systematically gather policy-relevant information in such settings.

Our spatial sampling problem is one of a broader class of research problems that are increasingly relevant in the digital world. The principles we develop extend broadly, for example for other ongoing projects to use volunteered local information for monitoring, such as the Global Polio Eradication Initiative, which uses mobile phones to monitoring vaccination programs and polio outbreak in remote regions of Nigeria and Pakistan (Kazi & Jafri, 2016), or for improving platforms that seek to coordinate optimal spatial provision of services, from Uber to the eBird project.[1] The most relevant literature concerns the latter, from emerging literatures in Computer Science on 'citizen science' (Bonney, et al., 2009) and 'directed learning,' which have primarily focused on aggregate-level allocation and coverage issues (e.g., (Marshall, Kleine, & Dean, 2012; Xue, et al., 2013; Ho, Slivkins, & Vaughan, 2014; Yang, Xue, Fang, & Tang, 2012). Our main contribution to this literature is to present an approach and methodology for experimental design to identify incentive responses and behavioral parameters at the level of the individual participant. These insights have the potential to inform incentive design for a broad class of activities aimed at effecting aggregate level outcomes by changing individual behavior.

The paper proceeds as follows. In the following section we give an overview of the crowdsourcing project that underpins our application. We then present the experimental design, and discuss our empirical strategy. A subsequent section gives an overview of our data and some descriptive evidence. Finally, we present our main results, covering our core estimates around the extent to which spatially-varying incentives can shift sampling to under sampled regions, and a number of additional questions.

Context and Implementation

The Crowd Sourcing Rangeland Conditions (CSRC) project[2] aims to improve data on vegetation and rangeland conditions in a remote and sparsely populated part of the world, through crowdsourcing information from local rangeland users—pastoralists (livestock herders). Such information can be difficult to obtain, because standard approaches can be too costly to carry out at scale or fine time resolution (sending in experts such as professional rangeland ecologists), or too crude (remote sensing data from satellites, which is often reported at the level of tiles that are hundreds of meters wide or more). Specifically, the project seeks to collect and use geo-located, ground-level information on forage conditions from remote locations in central Kenya.

Through a series of focus group discussions and field pilots, we developed a simple survey to standardize the data collection process. The surveys were developed to be flexible enough to cross ethnic boundaries and did not require literacy or familiarity with standard symbols or metrics. The survey collects information on aspects of vegetation conditions that are especially challenging to discern using data from remote sensing platforms, such as plant vegetation

[1] http://ebird.org/content/ebird/.
[2] https://www.udiscover.it/applications/pastoralism/

palatability. In short, each survey is geo-located and consists of a phone identifier, a photo, a set of four questions on vegetation conditions and palatability, one question on carrying capacity of the vegetation in the immediate vicinity, one question on distance to water for livestock, and one question asking which types of animals the participant has grazing near that location.[3]

We launched the vegetation survey as a mobile application for smart phones using the open source platform ODK Collect. Smartphones with GPS receivers were used so that each survey was geo-located at the point of completion, even in the absence of cellular service. Completed surveys are stored on the mobile device until the participant is in a location with cellular coverage, at which time it is sent to a served, Google's App Engine in this case. The mobile platform also minimizes the burden of participation, allowing participants to complete a survey in 1-2 minutes with nothing more than their phone. The project provided all participants with training, smart phones, solar phone chargers and phone credit required to submit the surveys.

A 150 km by 155km region north of Mount Kenya in central Kenya was selected for this study. This region was ideal for this pilot because, relative to other pastoralist regions, it offered fairly easy access and extensive phone coverage, which was important as we worked to address technical challenges associated with piloting software and minimize costs. Volunteers were recruited by local field staff. Our main selection criteria was that the participants currently spent much of their time herding livestock, were geographically dispersed and could attend the training sessions. In total, 113 volunteer participants were trained at 5 training sites in the Isiolo, Laikipia, and Samburu districts of northern Kenya (Figure 1).

Figure 1. Kenya with the study region in red

[3] See the Appendix A for a full description of the survey and a mobile application.

A baseline survey was successfully collected from 111 of the 113 participants. All of the participants were males, most were in their 20s, had a great deal of herding experience and nearly all (84%) continued to herd livestock as a primary livelihood activity (Table 1). All of the participants had access to phones before the project launched and many of them reported a great deal of phone use.[4] Over 40% report using cell phones to learn about rangeland conditions in other areas. This is all to say, our participants were generally familiar with mobile phones.[5] Parenthetically, three of the participants had smart phones before the project started, but for most, smart phones represented a technology that they were aware of but had little experience with.

Table 1. Summary statistics of participants.

Variable	N[1]	mean	sd	min	Max
Age (years)	111	22.37	3.52	18	35
Education Completed[2]	111	2.09	0.84	0	5
Year of herding experience	111	13.32	5.53	2	27

[4] Although the maximum number of SMSs and calls may seem high, these values do not represent extreme outliers in the data. In addition, data from an end-line survey (not reported in this manuscript) reports even greater phone use by the end of the project. But, in both cases rounding becomes quite coarse as frequencies get high.
[5] Although we do not know this to be the case, it seems very likely that the participants were, one average, more familiar with mobile technology than the population average of the region.

Primary activity (1=yes):					
- Livestock Herding	111	0.84	-	0	1
-Livestock other	111	0.01	-	0	1
-Farming	111	0.02	-	0	1
-Casual labor (not herding)	111	0.03	-	0	1
-Domestic	111	0.06	-	0	1
-Business Owner	111	0.04	-	0	1
- Other	111	0.01	-	0	1
Number of people talked to on the phone per day	111	20.0	15.2	2	100
Number of calls made and received per day	111	21.2	16.1	2	100
Number of SMS/WhatsApp messages made and received per day	111	87	128	0	700
Use cell phones to get information on rangeland conditions for locations that you are not at? (1=yes)	111	0.41	-	0	1

[1] Baseline data was not collected for two participants. [2] 0=none, 1=primary education, 2=secondary, 3= College (Professional certificate), 4= College (Diploma), 5=Graduate.

The data collection phase of the project was launched on March 16th, 2015 and ran for 148 days, ending on August 12th, 2015. The participants were trained on phone use, the survey application, the survey protocol, and *Ntioto*—a project-developed mobile application used to communicate rewards.[6] Surveys were to be collected only during daylight hours and submissions were required to be a minimum of 60 minutes apart. This indirectly limits the maximum number of valid surveys by an individual in a day at about 12 in the study region. No other protocol was used to invalidate submissions, but participants did receive extensive training on the characteristics of a high quality survey.[7] In all cases, participants were rewarded for each valid survey submitted.

Participation was voluntary and there were no penalties for non-participation. Formal attrition, those that returned their phones early and officially dropped out of the project, over the study period was 9.7%. Nine of the 11 participants that left the study dropped out due to employment opportunities and/or moving. One participant lost his phone mid-way through the project. One individual was formally dropped by our field staff for behavioural issues during the trainings. There are also cases of informal attrition, where participants never make submissions or make submissions early in the study but then stopped participating completely. Finally, in ten cases, participants reported issues with their phones that we could not solve remotely. In those cases we replaced the phones, but acknowledge that the duration from phone failure to replacement was occasionally substantial and non-random.

This is all to say that our initial sample, which was not random to begin with, did experience attrition that may not have been random. To minimize the impact of prospective attrition bias, we maintain all individuals in the data for the full study period, recording zero submissions for all days after they dropped. But, as will be discussed in the next section, that attrition is unlikely to be systematically related the treatments considered here.

[6] The mobile application *Ntioto* was developed by Rich Bernstein. *Ntioto* is a Samburu word used for things that point towards money. The application is described in more detail in the next section.

[7] Survey quality training focused on three components: the region to report on, the survey questions themselves, and how to take photos that would be most helpful to us. Most important for this research, participants were instructed to report the conditions of their immediate vicinity that was captured by their photo.

Research Design

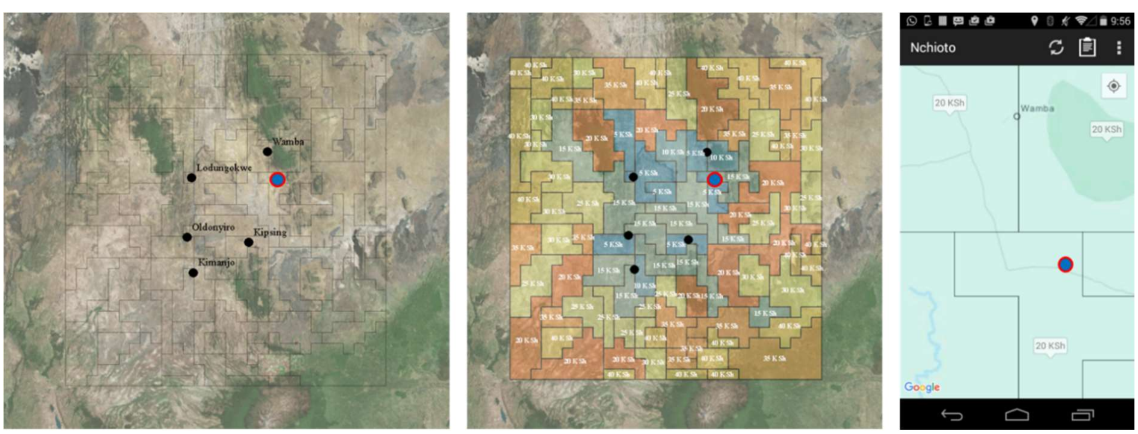
The study was designed to tackle a common challenge with open crowdsourcing: an undesirable sampling distribution. In our setting, respondents might not naturally provide surveys in locations where we would like to collect data, because of innate preferences for certain locations, and costs of movement. Foreseeing issues with suboptimal clustering of submissions, we implemented a series of randomized control trials (RCTs) that introduced spatial heterogeneity in rewards in order to determine participants' responsiveness to spatially-varying response incentives and to spatially disperse the submissions. In order to disentangle the fixed cost of providing a survey from the variable cost of reporting from a particular location, we also implement treatments with uniform (spatially invariant) rewards.

The spatially-varying incentive schemes (SV) were developed by partitioning the 23,250 km^2 project area by landcover, into 96 regions.[8] Each region was designated a reward, which was varied across stages of the study.

Participants used a mobile phone application called *Ntioto* that worked offline, to learn about the rewards. The application contained a map containing all of the regions, the associated rewards and their own location. The maximum/minimum reward offered was 40/5 KSH (about 0.40/0.05 USD) per submission, with the values constrained so that the average value across the 96 regions was 20 KSH. Control treatments with uniform incentives involve providing payments of 20 KSH per survey, regardless of location. As a reference point, the minimum daily wage for unskilled worker in the agricultural industry is 228.30 KSH (WageIndicator Foundation, 2015).

Figure 2 illustrates how the study region was divided into incentive regions by landcover, with the five training sites indicated (left panel), and an example of how they were then designated rewards (middle panel). The participants could zoom to see a higher resolution map, as is shown by the screen shot of the *Ntioto* application on the right. As an illustration, the location of the phone is indicated by the blue circle with a red boarder in all three screens.

Figure 2. Study site regions (left), regions with an example incentive scheme (middle), and an example screen shot from the mobile application (right).



Notes: Phone location is indicated by blue circle with a red border.

[8] The regions were designed by hand from a grid of 930, 5km x 5km tiles. Using satellite imagery, like (based on groundcover, elevation, etc) and near tiles were grouped into 96 regions composed of between three and 35 tiles.

The participants were divided into five training sites geographically. A factorial design within training site was used to implement the incentive experiments described in this manuscript and a monitoring experiment, which is examined in Jensen et al. (2017).[9] All incentive groups see the same treatments and the monitoring treatments are equally represented in each of the incentive groups, so there is no reason to believe that the two experiments would interact systematically. Thus, the monitoring treatments left out of the analysis included in this manuscript. The allocation of participants into incentive groups is described in Table 2.

Table 2. Allocation of participants into incentive groups across training sites.

Incentives	Wamba	Mabati	Kipsing	Kimanju	Oldonyiro	Total	%
Incentive Group 1	5	5	3	2	3	18	0.16
Incentive Group 2	5	4	3	3	3	18	0.16
Incentive Group 3	5	4	3	2	3	17	0.15
Incentive Group 4	6	5	2	3	3	19	0.17
Incentive Group 5	5	4	3	4	5	21	0.19
Incentive Group 6	6	6	2	3	4	20	0.18
Total	32	27	16	17	21	113	

The sequence of events during implementation are as follows. Two initial days of training were provided at each training site. Following the training, all participants faced a single uniform reward—20 KSH/submission—for 23-27 days.[10] Because rewards were spatially invariant, this was meant to identify participants' home movement range and to familiarize them with the survey and submission process. The optimal response is clearly to provide regular reports without deviating from where one would have otherwise moved in absence of the intervention and when the opportunity cost in time or inconvenience is not so high as to rule out reports altogether. We use the reports from this first phase to learn about the spatial distribution of submissions in the absence of our intentional distortions, which come in the experimental phases.

For the following 6 days, we ran a test of the process used to spatially-vary incentives. This six day period was intended to help the participants identify issues with the mobile application used to communicate incentives (*Ntioto*) so that we could then address in the subsequent training. The data from this period are not included in the analysis as there were many issues with the application and some participants did not see the reward scheme that they were facing at all. A one day training followed the 6 day test period.

For the remaining duration of the project, participants cycled through phases in which they were treated with three incentive maps, each for about a 9-day period—one with uniform incentives (U) and two different spatially-varying (SV). The distinction between within phase SV rewards is in the algorithm that used the distribution of submissions from all previous

---

[9] Participants were randomly assigned to available openings in one of three monitoring groups within each training site. Incentive groups are intended to be near uniform in size and determined using a random without replacement process within training site and within monitoring group.

[10] The implementation was rolled out by site in order to account for differences in training dates and personnel constraints. The result is that the participants were not all in phase one for the same period or in the baseline for exactly the same duration.

phases to set rewards for the following phase.[11] Because the algorithms themselves are not relevant for this research, we do not distinguish between them in this paper.

All households went through four such phases. Incentive group designation determined the order that participants faced the three maps within each phase. The experiment schedule and order that each group faced a specific incentive map is found in Table 3.

Table 3. Experiment schedule by incentive group

Incentive Group	Phase 1 (27 days)	Phase 2 (6 days)	Phase 3 (27 days)			Phase 4 (27 days)			Phase 5 (27 days)			Phase 6 (27 days)		
Time Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	U	SV2	U	SV3	SV3	SV4	SV4	U	SV5	U	SV5	SV6	U	SV6
2	U	SV2	U	SV3	SV3	SV4	SV4	U	U	SV5	SV5	U	SV6	SV6
3	U	SV2	SV3	U	SV3	SV4	U	SV4	SV5	U	SV5	SV6	SV6	U
4	U	SV2	SV3	SV3	U	SV4	U	SV4	U	SV5	SV5	S6	U	SV6
5	U	SV2	SV3	U	SV3	U	SV4	SV4	SV5	SV5	U	SV6	SV6	U
6	U	SV2	SV3	SV3	U	U	SV4	SV4	SV5	SV5	U	U	SV6	SC6

The spatially-varying maps were generated a few days before each phase started, drawing on all of the submissions to date and designed to draw submissions towards regions with low submission densities. All SV maps maintained the same partitions, maximum, minimum, and mean (40, 5, 20 KSH respectively). The maximum was intentionally set fairly low so as to reduce the risk of incentivizing participants entering restricted areas or changing their livelihood strategies dramatically.

Empirical Strategy

The objective of this work is to learn about participants' responsiveness to spatially-varying rewards, particularly to understand the extent to which respondents are willing to sample from otherwise under sampled regions. To do so, we examine changes in the distribution of submissions coinciding with shifts in rewards, as outlined above.

Although the specific empirical strategies used in this research are discussed in detail in each section of the analysis, they all rely on exogenous changes in rewards. Recall that during each phase, each participant faces uniform rewards for a short period of time, which is especially important for our empirical strategy for three reasons:

1. The participants are pastoralists that are known to be mobile. Such movement would results in changes to the distribution of submissions over time, which we do not want to mistake for effects of changes to rewards.
2. Seasonal changes to labor demand would affect the time allocation to data collection.
3. The enthusiasm that participants have for this project and skill with which they navigate the map application will change with time.

[11] Although not the focus of this paper, one incentive scheme was determined by the inverse of the density of submissions in each location to date. The second was developed by a machine learning algorithm.

The repeated implementation of the uniform incentives, which participants face in a random order within each phase, helps us to control for these time trends and allow us to control for individual level fixed effects.

In addition, our analysis is restricted to phases 3-6 to avoid the trainings and early stages when participants were still learning how to use the applications and when we were fixing bugs in the process.

Data

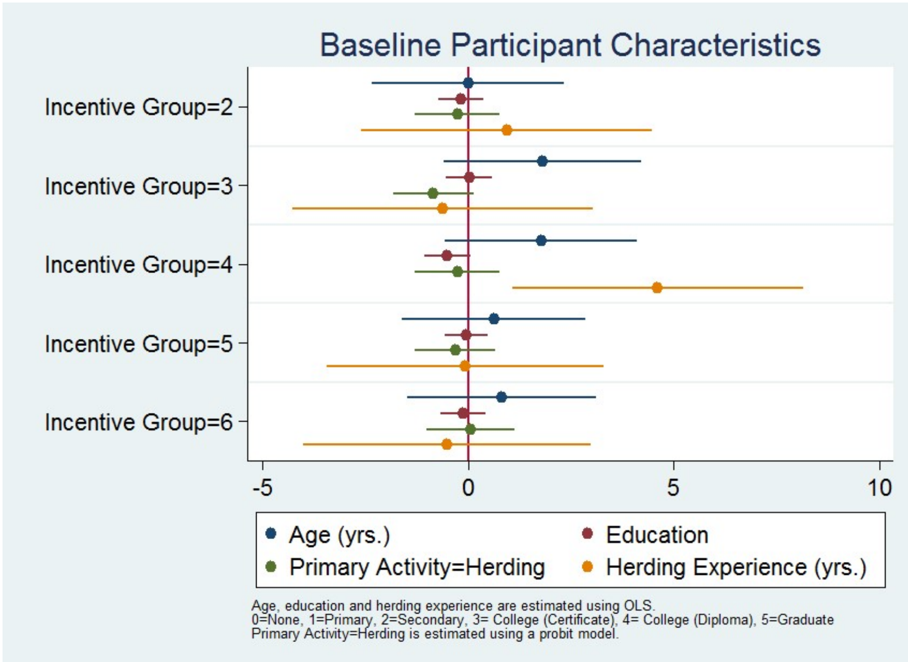
Balance Tests

The participants were randomly assigned into one of six incentive groups. To test for baseline balance across the six incentive groups, we regress baseline characteristics onto incentive group dummies in an OLS specification. Here, we are essentially comparing each of the incentive groups against the omitted group—one. By this method, a lack of balance between group one and the others is expressed by coefficient estimates that are statistically different from zero. And the differences between groups 2-6 can be examined by examining the confidence intervals around the means.

Considering our small sample size (N=113) and 6 incentive groups, we expect small differences to arise, which is marginally true for education and primary activity (Figure 3). If we adjust p-values to account for the multiple hypotheses (20, to be exact) represented by Figure 3, there is no question of balance.[12]

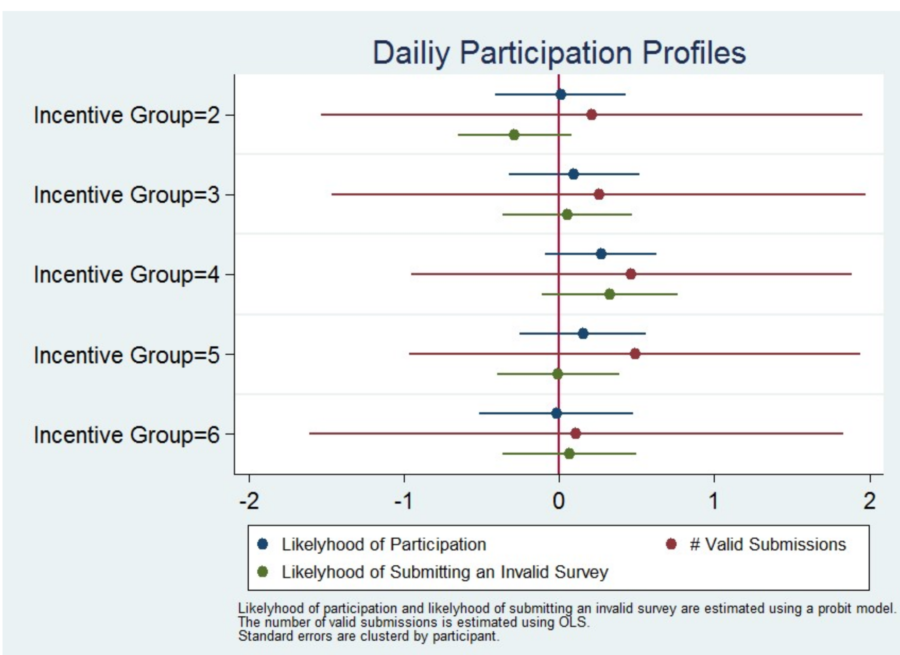
Figure 3. Differences in average group characteristics from Incentive Group=1 at baseline. The dots represent the estimated mean that the lines represent the 95% confidence intervals.

[12] For example, the unadjusted p-value Herding Experience (yrs.) for the fourth Incentive Group, which is the most statistically significant of the above estimates, is 0.011. The Bonferroni adjusted and Holm adjusted p-value is 0.22 (analysis available upon request).



We also use submissions from the first phase to regress daily rates of participation, likelihood of daily participation, and daily likelihood of submitting an invalid submission onto incentive group dummies. As a reminder, all participants were treated equally and faced a single, identical and uniform incentive until the end of the first phase. Field staff were unaware of incentive group designations until Phase 2 was completed. Similar to above, the groups are quite balanced (Figure 4).

Figure 4. Daily difference in participation characteristics from Incentive Group=1 during phase one.

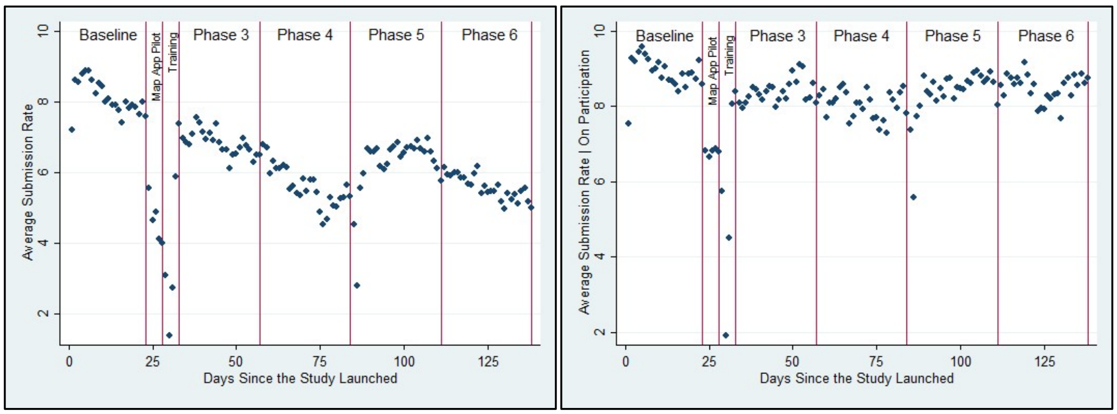


All in all, the incentive groups seem to be remarkably well balanced in both characteristics and submission profiles before treatment considering their small size. We proceed assuming that randomization was successful.

Submissions

Over the study, we received 107,308 survey submissions, 94,590 of which were valid.[13] Figure 5 illustrates daily average submission rates across the study period.

Figure 5. Daily average valid submission rates (left) and daily average valid submission rates among those that participated that day (right) across the study period.



There has been a clear decline in participation rates (Figure 5, left), but the decline is mostly due to an increase in non-participation, rather than lower rates of submission by those participating (Figure 5, right). We assume that this is mostly due to the aggregate effect of the 9 formal drop-outs as well as additional informal drop-outs as excitement for the project waned or as participants faced labour constraints. In addition, the map application pilot phase and the training phases, which are indicated by "Map App Pilot" and "Training" in Figure 5, had large depressive impacts on participation rates. We focus on data from Phases 3-6 for the analysis in this paper, all of which take place after the training.

Participation Profile

It is important to note that during Phases 3-6, the period in which our analysis takes place, participants faced spatially-varying incentives for twice as many days as they faced uniform incentives. In addition, on any particular day, there are twice as many participants facing SV than U. We account for this in our analysis, most often by focusing on daily statistics at the individual level.

Table 4 provides some basic summary statistics of the characteristics of submissions, dividing the data into baseline (Phase 1), those collected under U during phase 3-6 and those collected under SV during phases 3-6.

Table 4. Summary statistics of participation.

	Phase 1 (baseline)	Phase 3-6

[13] Allowing for error on the side of the participant, payments were made for submissions that were 50 or more minutes apart and were collected during daylight hours. We use this standard to identify valid submissions. In addition, we have identified 3,888 repeated submissions, which are dropped.

Rewards:	Uniform		Uniform (U)		Spatially-Varying (SV)		
	Stat.	Std. Dev.	Stat.	Std. Dev.	Stat.	Std.	Diff [#]
Total submissions	22,289		24,215		44,268		
Likelihood of participation/day	0.813	0.390	0.695	0.460	0.657	0.475	0.04***
Average submissions/person/day	7.23	4.32	5.86	4.82	5.37	4.87	0.49***
Average rewards/submission	20	0	20	0	20.4	13.2	-0.42
Average rewards/person/day	145	86.5	117	96.4	110	146	7.60
Average # of regions/person/day	1.25	0.896	0.965	0.848	0.906	0.857	0.06***

Note: Phase 2 submissions are not included in this table. [#] Participant clustered standard errors.

The above seems to indicate a small difference in behaviour comparing SV to U. Although the change is in the direction of reduced participation and fewer regions visited per day SV, the incentives may still have succeeded in changing submission patters to one more favourable to our objective of increasing sampling in regions were under sampled.

Results

The primary objective of varying the incentives spatially is to effect the distribution of submissions so that it is more aligned with our objectives. In this case, we are generally aiming to maximize our value for information on the vegetation conditions by increasing submissions in regions that we have the least information—the under-sampled regions. We take three empirical approaches to testing for success in this regards. We first determine if individuals respond to changing rewards by examining the reward elasticity of submitting a survey while controlling for latent costs and benefits associated with each location. We then test the aggregate distribution of submissions to see those individual level changes in behavior resulted in more ideal distributions. Finally, we end with a value of information analysis to ensure that we were successful in increasing the net value of information per cost.

Elasticity of submission rate with respect to rewards

Formally, assume that individual i maximizes daily utility by choosing to be in discretized locations ($l \in [1, \dots, L]$) during each period (t) over T periods in each day (d). Let $v(l_{idt})$ be a function that translates individual and location specific attributes into a common scale, which then enters the utility function (equation 1).

(1)

$$U_d = \max_{l \in [1,\dots,L]} \sum_{t=1}^{T} u(v(l_{idt}))$$

The result is a daily distribution of visits by each individual across space. Assuming additivity in the value function, let the value of each location be a function of costs ($cost_{lidt}$) and benefits ($benefit_{lidt}$) so that $v(l_{idt}) = cost_{lidt} + benefit_{lidt}$. Thus, an individual's day can be described by a set of location visits, which in turn can be described by a set of costs and benefits associated with each location.

We use ρ_{id}^* to signify an individual's original (undistorted) optimal distribution of location visits on day d, which is a vector of the ratio of the day spent in each location, the components ($\Pr(l_{idt}^*)$) of which sum to one for each individual across all the locations each day d. This ratio is a function of the relative costs and benefits of the location, which are in the same units as the components for convenience (equation 2).[14]

$$(2) \qquad \rho_{id}^{*} = [\Pr(l_{id1}^{*}), \ldots, \Pr(l_{idt}^{*}), \ldots, \Pr(l_{idT}^{*})]$$

where:

$$\Pr(l_{idt}^{*}) = cost_{lidt} + benefit_{lidt}$$
$$\textstyle\sum_{t=1}^{T}\sum_{l=1}^{L}\Pr(l_{idt}^{*}) = 1$$
$$cost_{lidt} = cost\ for\ individual\ i\ to\ be\ at\ location\ l\ in\ period\ t\ on\ day\ d$$
$$benefit_{lidt} = benefit\ for\ individual\ i\ to\ be\ at\ location\ l\ in\ period\ t\ on\ day\ d$$

The project introduces an incentive scheme that provides a reward ($reward_{ld}$) for providing a survey from a location on a specific day, which is scaled in the distribution function to the cost and benefits units by the factor α_{idt} (equation 3). In addition, submitting a survey requires some effort ($effort_{i}$), which is constant across locations and time and is scaled to the cost and benefit units by the factor α_{i}^{2}.

$$(3) \qquad \rho_{id}^{r} = [\Pr(l_{id1}^{r}), \ldots, \Pr(l_{idt}^{r}), \ldots, \Pr(l_{idT}^{r})]$$

where:

$$\Pr(l_{idt}^{r}) = cost_{lidt} + benefit_{lidt} + \alpha_{idt}^{1} * reward_{ld} + \alpha_{i}^{2} * effort_{i}$$
$$\sum_{t=1}^{T}\sum_{l=1}^{L}\Pr(l_{idt}^{r}) = 1$$

Assume that the project implements a spatially uniform treatment (U), which offers rewards for submissions that are uniform across space ($reward_{d}^{u}$). In this case, we us the syntax $\rho_{id}^{*} = \rho_{id}^{u}$, where ρ_{id}^{u} is the utility maximizing distribution of locations under uniform rewards (equation 4).

$$(4) \qquad \rho_{idt}^{*} = \rho_{id}^{u}$$
$$= [\ldots, \Pr(l_{idt}^{u}) = cost_{lidt} + benefit_{lidt} + \alpha_{idt} * reward_{d}^{u} + \alpha_{i}^{2} * effort_{i}, \ldots]$$

where:

$$reward_{ld} = reward_{d}^{u}\ \forall\ l\ if\ d \in U$$
$$\sum_{t=1}^{T}\sum_{l=1}^{L}\Pr(l_{idt}^{u}) = 1$$

Now, assume that the project implements a spatially varying treatment (SV) in which rewards for submissions that vary across locations l. The change to the distribution of submissions while going from uniform (U) to specially varying (SV) will identify the individual's spatial response to incentives as long as we can control for what the distribution of submissions under U would have been, since ρ_{id}^{u} controls for the costs and benefits of each location and the costs of

[14] For example, if individual i has the option of spending her time in five location and has four periods in a day, but prefers to spend half of her time in only the first two locations, $\rho_{id}^{*} = [0.5, 0.5, 0, 0, 0]$.

providing surveys. In that case, α_{idt}, which is the marginal change in the ratio of time spent in location *l* due to a change in the reward, is identified even though we cannot observed the latent costs and benefits associated with locations.

(5)
$$\begin{aligned}\rho_{id}^{sv} &= [\dots, \Pr(l_{idt}^{sv}) = cost_{lidt} + benefit_{lidt} + \alpha_{idt} * reward_{ld} + \alpha_i^2 * effort_i, \dots] \\ &= [\dots, \Pr(l_{idt}^{sv}) = \Pr(l_{idt}^{u}) + \alpha_{idt} * (reward_{ld} - reward_d^u), \dots]\end{aligned}$$

where:

$$\sum_{t=1}^{T}\sum_{l=1}^{L}\Pr(l_{idt}^{sv}) = 1$$

Empirically we can never observe an individual under both U and SV treatments at the same instant. But, any systematic change to the cost and benefits of visiting a location are addressed by the randomized order that individuals face reward maps within each phase, so that we can compare the distribution of each individual's submissions while facing SV to their distribution of submissions facing U within each phase.

Explicitly, within-phase randomization between U and SV addresses any potential bias to estimates due to systematic and unobserved variation in benefits associated with a location ($benefit_{lidt}$), the cost of visiting a location ($cost_{lidt}$), or the scaling factor between latent net benefit of a location and cash from the reward (α_{idt}).

Thus, we can rewrite equation (5) with individual, but phase (*p*) constant (rather than day constant) costs, benefits and scaling factor as follows.

(6)
$$\begin{aligned}\rho_{id}^{sv} &= [\dots, \Pr(l_{idt}^{sv}) = cost_{lidt} + benefit_{lidt} + \alpha_{idt} * reward_{ld} + \alpha_i^2 \\ &\qquad * effort_i, \dots] \\ &= [\dots, \Pr(l_{idt}^{sv}) \approx cost_{lip} + benefit_{lp} + \alpha_{ip} * reward_{ld} + \alpha_i^2 * effort_i, \dots] \\ &= [\dots, \Pr(l_{id}^{sv}) \approx \Pr\left(l_{ip}^{u}\right) + \alpha_{ip} * (reward_{ld} - reward_d^u), \dots] \\ &= [\dots, \Pr(l_{id}^{sv}) \approx \Pr\left(l_{ip}^{u}\right) + \beta_i^0 + \alpha_{ip} * (reward_{ld}), \dots]\end{aligned}$$

where:

$$p = phase$$

$$\sum_{t=1}^{T}\sum_{l=1}^{L}\Pr(l_{id}^{sv}) = 1, \sum_{t=1}^{T}\sum_{l=1}^{L}\Pr(l_{id}^{u}) = 1$$

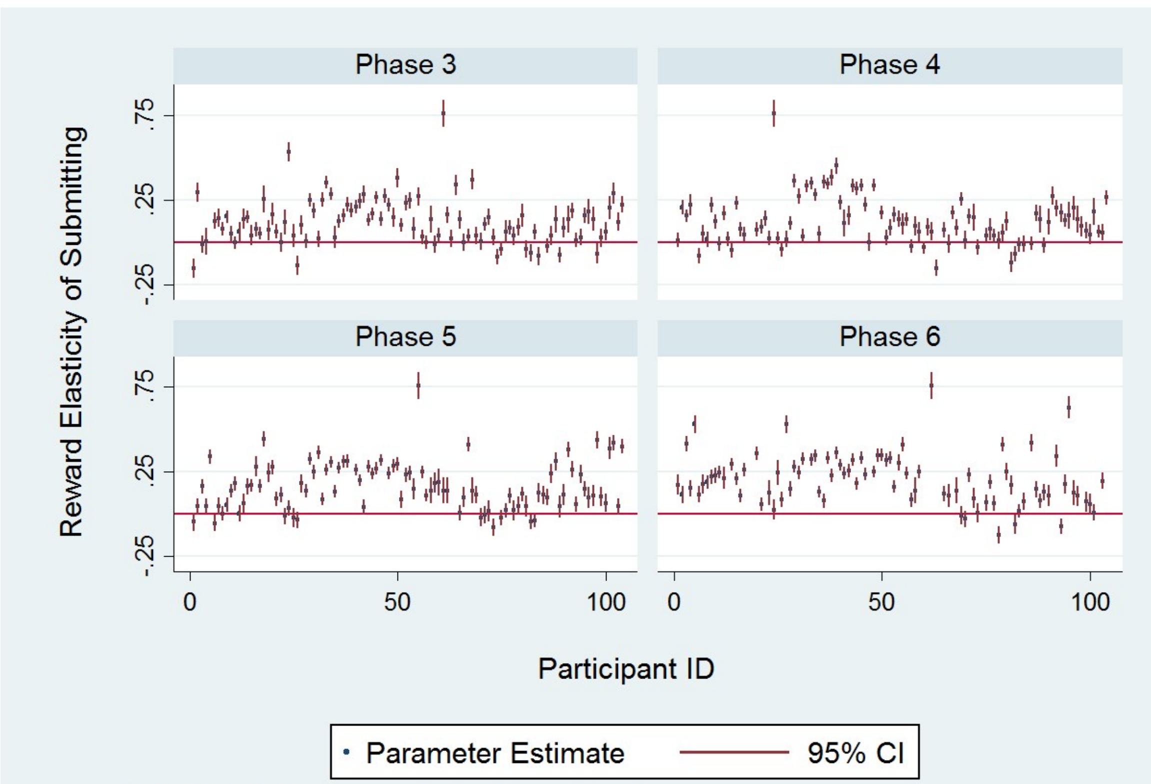
We observed the distribution of submissions under U (ρ_{id}^{u}) and SV (ρ_{id}^{sv}) and can econometrically estimate the fourth line in equation (6) in each location using an OLS regression ($\Pr(l_{id}^{sv}) = \beta_i^0 + \Pr\left(l_{ip}^{u}\right) + \alpha_{ip} * reward_{ld} + \varepsilon_{lid}$), for each phase ($p \in [3,4,5,6]$), where $\Pr(l_{id}^{sv})$ is the likelihood that a submission from participant i is from location l on day d in phase *p* during SV treatments. In addition, taking the natural log of the rewards and of the ratio of submissions in a location offers an estimate of the rewards elasticity of the ration of submissions in a region. Following our model's specification, the parameter for $\Pr\left(l_{ip}^{u}\right)$ is constrained to one.

The process for this regression is as follows. Within each phase, all of an individual's submissions in each region while facing uniform treatments are tallied. The tally is then divided by that individual's total submissions under uniform treatments in that phase to get the ratio of within-phase submissions in each region under uniform rewards (ρ_{ip}^{u}). We consider this variable to be a participant-phase specific control for the latent net benefit ($cost_{lip} + benefit_{lp} + \alpha_i^2 * effort_i$) of submitting from a region in the absence of rewards. We then follow a similar process, but at a daily temporal resolution, to generate the daily distribution of submissions (ρ_{id}^{sv}).

The components of that daily distribution of submissions ($\Pr(l_{id}^{sv})$) are regressed on our proxy for latent regional net benefits and fixed costs ($\Pr(l_{id}^{u})$) and the rewards associated with the specific region on that day. The results include 385 elasticity estimates, which is less than the total possible 452 (113x4) because we cannot estimate the elasticities for person-phases without submissions.

The estimates from the regression are illustrated in Figure 6. Each dot is a person-phase specific estimate of their reward elasticity of the ratio of their daily submissions to a region. For example, if an individual had an elasticity of 0.12 (the mean) in phase 3, a 10% increase in the reward offered for submitting in a specific region and *ceteris paribus*, would increase the ratio of submissions by that individual in that region by and expected value of 1.2%. That same participant might have a very different elasticity in other phases.

Figure 6. The participant and phase specific reward elasticity of submission ratio in a region



Notes: Standard errors are clustered by participant.

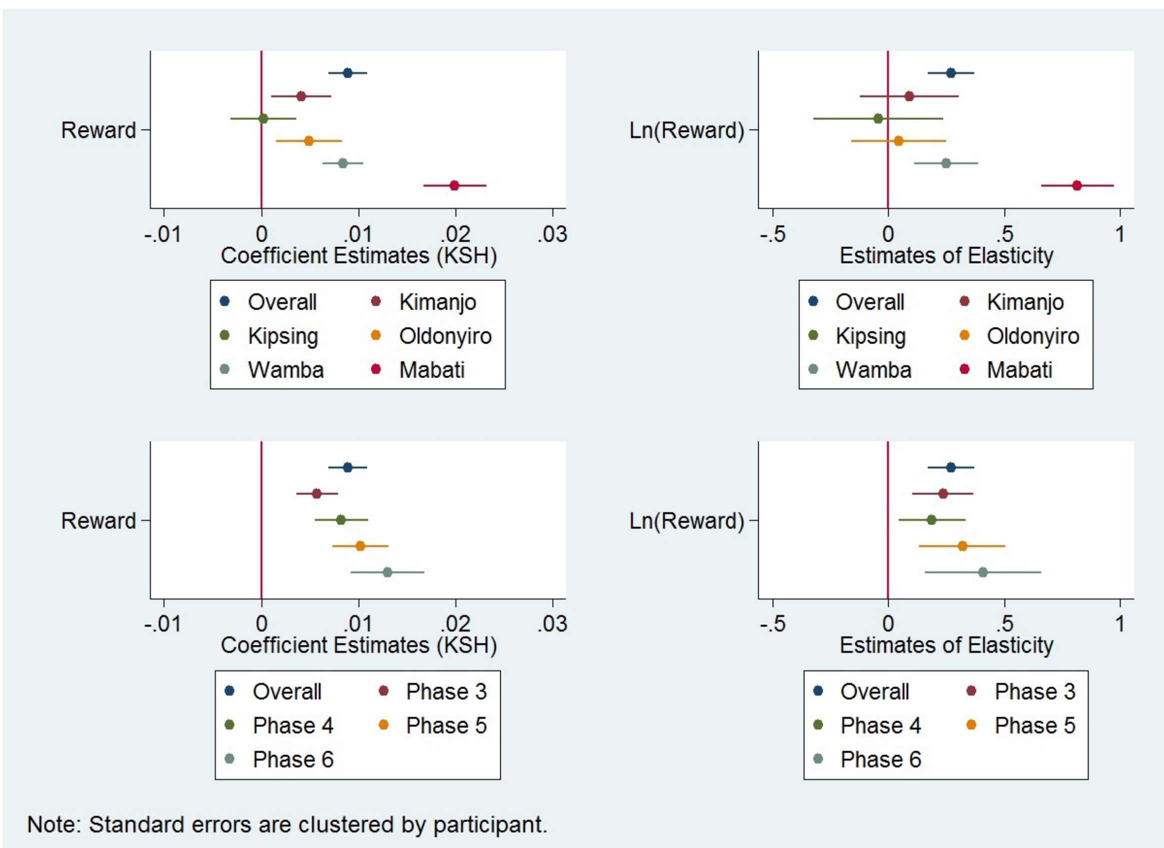
Most of the estimates (74%) are statistically greater than zero. These individuals are positively responding to rewards, at least in those phases. A few, 5%, are statistically less than zero. These individuals appear to be repelled by higher rewards, but come at a rate that is low.

From a program perspective, we are more interested in the average response of the crowd, rather than that of an individual. Thus, we estimate the average elasticity (α^1) across the crowd and across the phases, controlling for site (α_s^2), phase (α_p^3) and their interactions (equation 7).[15]

$$\ln(\Pr(l_{id}^{sv})) = \Pr(l_{ip}^{u}) + \beta_i^0 + \alpha * \ln(reward_{ld}) + \beta_s^1 + \beta_p^2 + \beta_s^1 * \beta_p^2 + \varepsilon_{lid} \tag{7}$$

To better understand the potential time and spatial trends, the analysis is also done allowing for variation between sites—estimating β_s^1 (top panels, Figure 7)—and then allowing for variation between phases –estimating β_p^2 (bottom panels, Figure 7). Once again, the parameter for $\Pr(l_{ip}^{u})$ is constrained to one.

Figure 7. Coefficient and reward elasticity of region submission ratio



---

[15] Site*phase interactions are an attempt to control for general

Notes: The *Overall* estimates include interacted phase-site dummy variable controls. Phase controls are included in the top panels and site controls are included in the bottom panels. Standard errors are clustered at the participant level and robust.

The sample-level elasticity estimate is 0.272. Thus, we could expect that increasing a reward in a region from its uniform rate of 20KSH to 30KSH would increase the ratio of submissions from that region by 13.6%.

Clearly there are some differences between locations. Participants from Mabati responded most strongly to the rewards while participants from Kipsing did not seem to respond at all. The authors have no hypothesis as to what causes heterogeneity between sites, but controls for simple household characteristics and participant fixed effects do not qualitatively change the estimates. In an endline survey of a subsample of participants, participants in Kipsing most frequently reported having issues with network coverage and the incentive map application, which injects noise into the system through delayed map (incentive) updating. But, because the subsample available for the endline survey was quite small, controlling for those reported network issues results in extremely large standard errors.

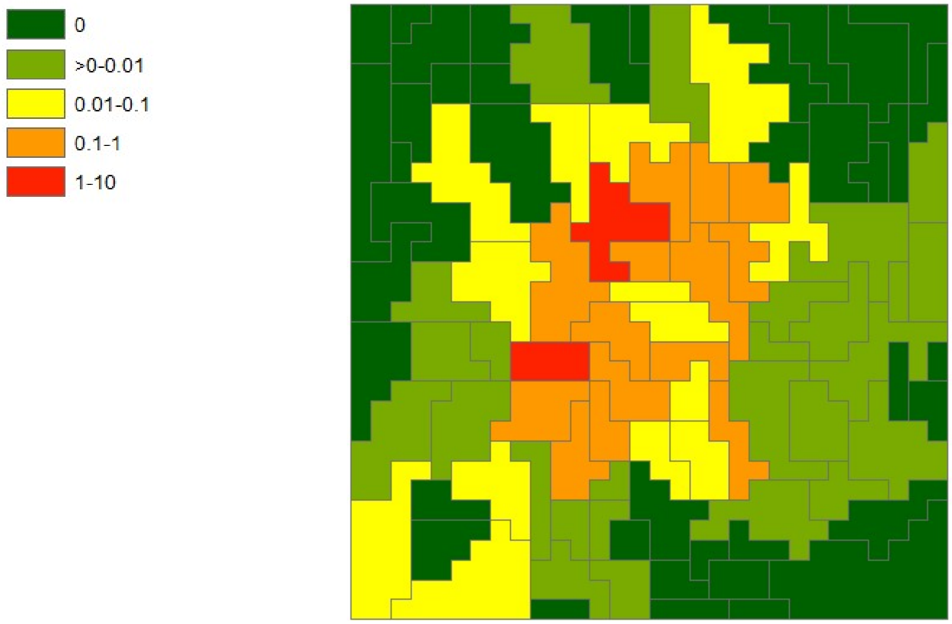
There is a clear increase in response to rewards over time. The observed drop in participation rates over time (Figure 5) provides a plausible hypothesis—non-random attrition. It is reasonable to assume that the participants that are least enthusiastic about the program are the most likely to stop participating. That same lack of enthusiasm may have made them less responsive to the rewards in periods that they did participate. In addition, we should expect that enthusiastic participants improve their skill at capturing high rewards with time and practice.

One concern with the above analysis is that it does not control for rewards associated with neighboring regions. For example, it is reasonable to think that participants might be willing to travel longer distances to reach a cluster of high reward regions. Alternatively, participants might have a tendency to move to adjacent regions if they have higher rewards. Although we do find some evidence that participants do avoid making submissions from particularly low rewards regions, accounting for rewards in adjacent regions has no impact on our estimates of elasticity. See Appendix C for analysis.

Sampling in under sampled Regions

We now examine the impact that our treatments had on improving the information generated by the project by drawing submissions into otherwise under-sampled regions. Under sampling can be understood as an extensive margin issue—regions in which there is no sampling versus regions where there is any sampling—or an intensive margin issue—regions that receive proportionally less sampling than others. Focusing on the extensive margin, 39 of 96 regions are unsampled after the baseline (dark green regions in Figure 8).

Figure 8. Distribution of submissions per region per day per participant across the 96 study regions during baseline.



We used spatially varying rewards that favored the green regions, especially the dark green regions, to attract submissions. Because we expect some degree of variation in motion and location, to test if the SV treatments worked, we need to compare them to the contemptuous U treatments though, not the baseline distribution. To do so, we will examine the impact that SV has on the daily rate of both the intensive and extensive margins. We define these terms contemporaneously, taking into account all the submissions up to that beginning of each phase. This updating approach is consistent with how the SV incentives were designed and mimics how a project could use an updating approach to achieve dynamic objectives.

To start with, we examine impact of SV on the extensive margin, identifying all those regions that had received zero submissions by the end of the prior phase. The change in sampled to unsampled regions across phases is displayed in table 5.

Table 5. Sampled and unsampled regions to date

Phase	Sampled (≥1 submissions)	Unsampled to date (zero submissions)	Percent unsampled to date (zero submissions)
1	-	-	-
2	57	39	40.6%
3	61	35	36.5%
4	71	25	26.0%
5	74	22	22.9%
6	82	14	14.6%

There are 96 total regions.

To test the impact of the SV rewards on participants' likelihood of visiting an unsampled region, we regress an indicator that an individual submitted from an unsampled region onto an indicator for reward scheme type (indicator=1 for SV). Under U the likelihood that a submission is from an unsampled region is 1.3%. SV increase the likelihood by at least 1.5 times in all specifications (Table 6, columns 1-3) and nearly quadruples it in our preferred specification with participant fixed effects (Table 6, columns 3).

In addition, we test if participants are making more submissions from unsampled regions per day. We find that under spatially varying rewards participants submit at least 0.09 more submissions per day from unsampled regions than they do under uniform rewards (Table 6, columns 4-7), which represents a more than doubling of the uniform rate of 0.097 submissions per participant from an unsampled region per day.

Table 6. Average marginal effect (AME) of SV on the likelihood that a submission is from an unsampled region and the number of submissions each day from an individual in unsampled regions.

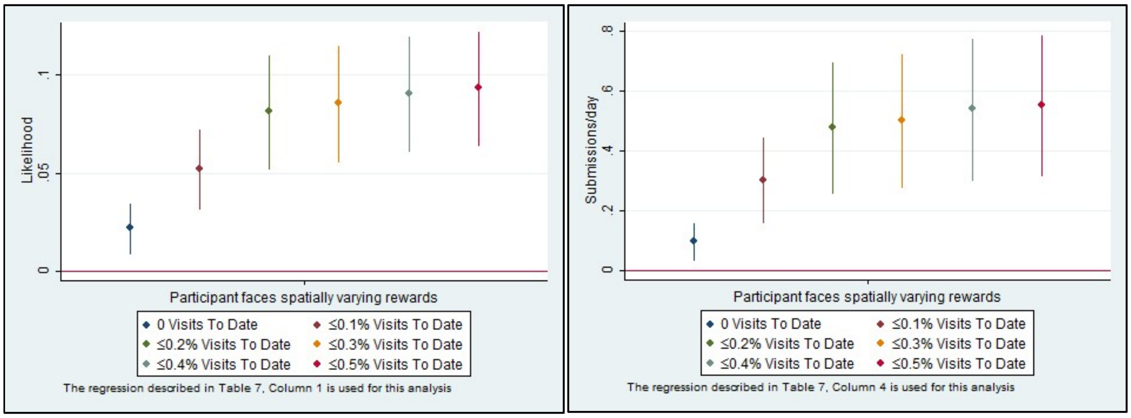
	Submission is from an Unsampled Region (AME from Probit)			Unsampled Submissions Each Day (OLS)[1]			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Spatially-Varying Rewards	0.0235***	0.0239***	0.0498***	0.0984***	0.0977***	0.0999***	0.0999**
	(0.00695)	(0.00712)	(0.0114)	(0.0312)	(0.0317)	(0.0321)	[0.0389]
Site*Phase Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Participant Characteristics	No	Yes	No	No	Yes	No	No
Participant Dummies	No	No	Yes	No	No	Yes	Yes
Observations	59,465	56,526	26,663	12,377	12,158	12,377	12,377
R-squared	0.298	0.297	0.344	0.113	0.110	0.171	0.171

[1] The dependent variable is participant-level count of valid submissions from unsampled regions each day.
Participant clustered standard errors in parentheses. Participant and date clustered standard errors in brackets.
*** p<0.01, ** p<0.05, * p<0.1

But, we are also interested in increasing submissions from under sampled (rather than only unsampled) regions because the regions are unlikely to be homogenous across space and time, and as a tool validating submissions. To test for impacts on under sampled regions we identify those regions that received fewer than the ratio χ of the total submissions to date, where $\chi \epsilon [0.001, 0.002, 0.003, 0.004, 0.005]$. If submissions were distributed uniformly across the 96 regions, each region would have received about 1% (χ =0.01) of the submissions.

The SV treatments have a large positive and statistically significant impact on the likelihood and rate of submissions from under sampled regions. We illustrate these results in Figure 9, which shows the coefficient estimate of the SV treatment on i) the likelihood of submitting a survey from an under sampled region, and ii) the number of submissions from undersampled region. Our results are robust to a wide range of definitions for identifying under sampled regions. Note that we include the extensive definition of undersampled—zero percent—and that those estimates are consistent with those found in table 6.

Figure 9. The impact of spatially-varying rewards on submissions from under sampled regions, where under sampled regions are those regions that received fewer than 0.1%-0.5% of the total submissions to date.



Submissions under the uniform rewards during baseline were highly clustered, leaving many regions under sampled and some completely unsampled. The above analysis shows that spatially varying rewards to favour those un (under)-sampled regions successfully attracted more submissions than does continuing with uniform rewords.

Value of Information

Having confirmed that the reward elasticity of submitting from a location is positive and that our spatial varying rewards successfully exploited this responsiveness to increase information coming in from under-sampled regions, we now turn examine if these efforts translated into improved value of information for cost. For our purposes every submission has value, but the value is decreasing in spatial and temporal density, since in the broader applications we have in mind the value of information would follow a strictly increasing and strictly concave function. For example, the 1st submission from a region is more valuable than the 10th. In our data, there more than 100 events in which the number of submissions from a region in a day surpasses 100 submissions. At the same time, on any given day about 64 regions are unsampled.

As a preliminary approach to exploring the effectiveness of the spatially-varying rewards on reducing expenses on costly, nearly redundant submissions, we place a value equal to the maximum reward in our system—40KSH—for the first submission within each region each day and 40KSH multiplied by the inverse of the number of preceding submissions for subsequent submissions from that region that day. For example, two submission in one region in one day are worth a total of 60KSH (40KSH/1 + 40KSH/2) while one submission per day in the same region for two days are worth a total of 80KSH (40KSH/1 + 40KSH/1). The value of each submission is then divided by the reward distributed for the submission. The ratio is the value of the submission to the cost. This is admittedly an ad hoc functional form assumption, yet it provides a reasonable first proxy to what the value function might look like in a real spatial mapping application.

This value structure acknowledges that there is always positive marginal value to additional information, but that the additional information that a survey contains is decreasing in daily frequency within each region. Admittedly, the initial value of 40KSH is rather arbitrary. Ideally, the maximum value and decay rate would be set internally, relying on actual value of information estimates. In this case, 40KSH was set as the maximum rewards for a single submission so it seems like a reasonable value to use. In addition, we do not allow information to extend beyond region or daily boundaries, which is unrealistic.

The maximum value to cost ratio that we observe is 8 (1st daily submission in a region with rewards equal to 5KS) and the minimum is 0.007874 (the 127th daily submission in a region with rewards equal to 40KSH).

The daily average value to cost ratio is regressed onto rewards scheme, estimating that the SV rewards increase our value to cost ratio by 0.197, an increase of about 63% from 0.319 under uniform (Table 7).

Table 7. The impact of spatially-varying rewards on the value of information to program cost.

	Daily Average Value to Cost Ratio	
VARIABLES	(1)	(2)
Spatially-Varying Rewards	0.197***	0.197***
	(0.0103)	(0.0104)
Constant	0.319***	0.556***
	(0.00589)	(0.0178)
Phase Dummies	No	Yes
Observations[1]	224	224
R-squared	0.600	0.684

[1] The number of observations is larger than the number of days because at any given day, uniform and spatially-varying reward schemes are active. Date clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Conclusion

In this paper we speak to literature in economics, computer science, and statistics, exploring the microeconomics of inducing optimal sampling in a crowdsourcing setting. We present an experimental design and empirical strategy to measure important behavioral parameters at the individual and group level that can inform the design of incentive systems to induce optimal spatial sampling. These issues are increasingly important in light of the explosion of digital technologies and the opportunities that this brings about.

Our study provides a proof of concept that an incentive scheme can be implemented in a remote setting with participants with low literacy and education attainment. Furthermore, we show that participants respond to incentives in sensible ways, responding to spatially-varying incentives by increasing submissions from high-reward but more distant areas. We show that in principle this could increase the net value of information collected in a particular application, and estimate the aggregate elasticity of response to spatially-varying incentives in terms of sampling from previously under sampled areas. These findings, and more importantly the underlying approach, foreshadow a range of new research problems that arise in the presence of efforts to harness emerging digital and communication technologies to improve decision making and policy design.

References

Africa Pay (2015). Minimum Wages in Kenya with effect from 01-05-2015. Accessed at Minimum Wages in Kenya with effect from 01-05-2015 on February 25, 2016)

Bonney, R., Cooper, C., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K., & Shirk, J. (2009). Citizen science: a developing tool for expanding science knowledge and scientific literacy. *Bioscience, 59*(11), 977–984.

GSM Assocation. (2015). *The Mobile Economy.* GSM Assocation.

Ho, C.-J., Slivkins, A., & Vaughan, J. W. (2014). Adaptive Contract Design for Crowdsourcing Markets: Bandit Algorithms for Repeated Principal-Agent Problems. *15th ACM Conf. on Economics and Computation (EC).*

Laffont, J.-J., & Martimort, D. (2009). *The Theory of Incentives: The Principal-Agent Model.* Princeton University Press.

Marshall, N., Kleine, D., & Dean, A. (2012). Coralwatch: education, monitoring, and sustainability through citizen science. *Frontiers in Ecology and the Environment, 10*, 332-334.

Pew Research Center. (2014). *Emerging Nations Embrace Internet, Mobile Technology: Cell Phones Nearly Ubiquitous in Many Countries.*

Xue, Y., Dilkina, B., Damoulas, T., Fink, D., Gomes, C., & Kelling, S. (2013). Improving Your Chances: Boosting Citizen Science Discovery. *First AAAI Conference on Human Computation and CrowdSourcing (HComp).* AAAI.

Yang, D., Xue, G., Fang, X., & Tang, J. (2012). Crowdsourcing to Smartphones: Incentive Mechanism. *Proceedings of the 18th annual international conference on Mobile computing and networking*, (pp. 173-184).

Appendix A: Vegetation Survey

Survey Text

Survey Meta Data:

- Phone EMEI
- Start: date, time
- End: date, time

Survey Questions:

- Photo
- GSP location

Screens 1a-3c: Questions are asked using only icons. Screens *Yb* and *Yc* are only asked if the response to *Y* is not absent, where *Y*=[1,2,3]. Participants may select more than one animal for screen *Yc*.

1a. Are the trees in this area dense, sparse or absent?
    b. Are their leaves Green/Brown/absent?
    c. Are they palatable for camel, cattle, goats, or none?

2a. Are the shrubs in this area dense, sparse or absent?
    b. Are their leaves Green/Brown/absent?
    c. Are they palatable for camel, cattle, goats, or none?

3a. Is the grass in this area dense, sparse or absent?
    b. Are they Green/Brown/absent?
    c. Are they palatable for camel, cattle, goats, or none?

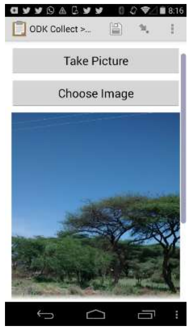
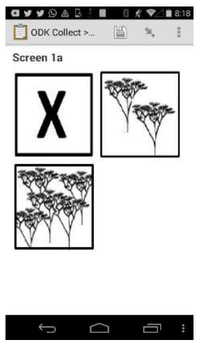
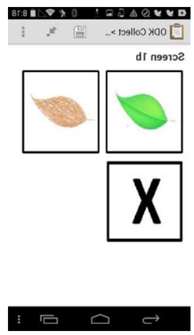
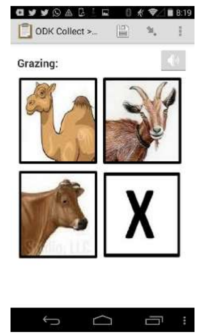
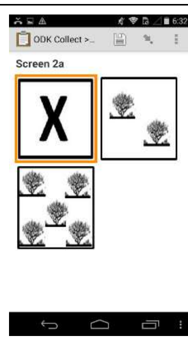
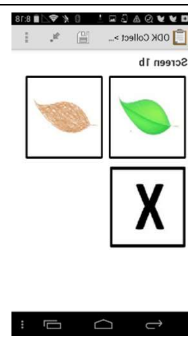
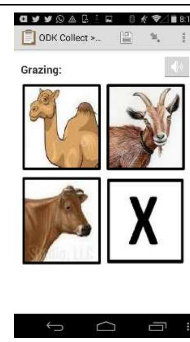
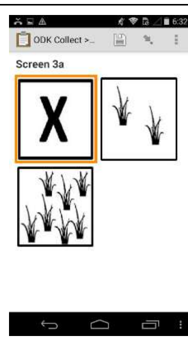
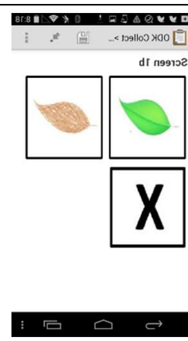
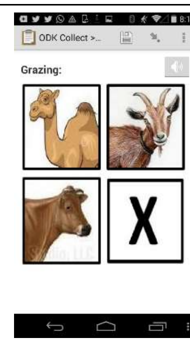
Screen 4-6: All these questions are accompanied by audio. The English translations of the audio content are in quotations. The audio was in Samburu for the study participants.

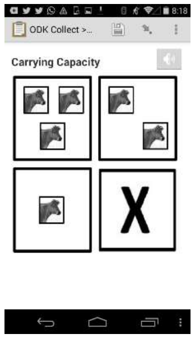
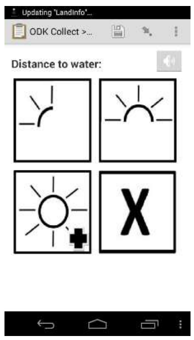
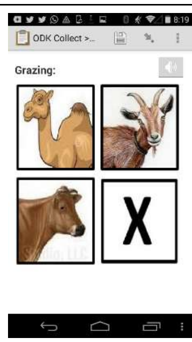
4. *Carrying Capacity 2*: "Think about the forage that is within twenty steps of you in every direction. How many cows could it feed for one day? Zero cows, one cow, two cows, or three cows or more. Please respond by touching the correct number of cows." Icons are an X for zero cows, an icon of one cow, an icon of two cows, and icon of three cows.

5. *Distance to water:* "From this location how long would it take to walk to the nearest water for livestock? Please respond by selecting the X icon if you are at a water point, the icon of the quarter sun for an hour or less walk from here, the icon of the half sun for half of a day, and the icon of the full sun if the nearest waterpoint is a day or more walk from here."

6. *Grazing:* "Are your livestock grazing here now? Respond by indicating the types of animals that you have grazing here. Select the X if none." Icons are: X, camel, cow, goat.

Survey Screen Shots

Topic	Screen shot a	Screen shot b	Screen shot c
Location			
Trees			
Shrubs			
Grass			

Topic	Screen shot	Audio in recorded in Samburu
Carrying Capacity		"Think about the forage that is within twenty steps of you in every direction. How many cows could it feed for one day? Zero cows, one cow, two cows, or three cows or more. Please respond by touching the correct number of cows."
Water point		"From this location how long would it take to walk to the nearest water for livestock? Please respond by selecting the X icon if you are at a water point, the icon of the quarter sun for an hour or less, the icon of the half sun for half of a day, and the icon of the full sun if the nearest waterpoint is a day or more walk from here."
Herding		"Are your livestock grazing here now? Respond by indicating the types of animals that you have grazing here. Select the X if none."

Appendix B: The Impact of Uniform Rewards.

This research assumes that uniform rewards have little or no impact on the locations that individuals visit or duration that they spend at those locations. One way that the uniform rewards could change location or duration is if the rewards were sufficient that they reduced the marginal utility of cash to the extent that submitting surveys substituted for other livelihood activities. But, the maximum rewards that a participant could generate in a day was 240 KSH, or about the minimum daily wage for an unskilled agricultural worker in the area. Although that is a considerable amount, the substitution also includes a substitution across time because participants were paid about once every 2 months by the crow sourcing project. That it, submitting surveys generated future income, not current income. So the only concerned is that the promise of future rewards reduced the current marginal utility of cash significantly, which seems unlikely.

A second way for the uniform treatments to effect participant location is if submissions required a great deal of time, so that they also entered the utility function through a time constraint. The time required to submit a survey is very small, less than 5 minutes, and so there should be very little substitution due to time constraints.

Assuming that participation is not bumping up against a time constraint and that the rewards are not sufficient to substantially reduce the marginal utility of income, we could assume that the decision to visit a location can be modelled using a multinomial logit model. Here, the probability of individual i being in location l in absence of rewards can be written as follows.

$$\Pr(l_{idt}^{*}) = \frac{e^{[cost_{lidt}+benefit_{lidt}]}}{\sum_{l=1}^{L} e^{[cost_{lidt}+benefit_{lidt}]}}$$

With uniform rewards the probability of individual i being in location l can be written as follows.

$$\Pr(l_{i11}^{u}) = \frac{e^{[cost_{lidt}+benefit_{lidt}+\alpha_{idt}*reward_{d}]}}{\sum_{l=1}^{L} e^{[cost_{lidt}+benefit_{lidt}+\alpha_{idt}*reward_{d}]}}$$

$$= \frac{e^{[cost_{lidt}+benefit_{lidt}]} * e^{[\alpha_{idt}*reward_{d}]}}{\sum_{l=1}^{L} e^{[cost_{lidt}+benefit_{lidt}]} * e^{[\alpha_{idt}*reward_{d}]}}$$

$$= \frac{e^{[cost_{lidt}+benefit_{lidt}]}}{\sum_{l=1}^{L} e^{[cost_{lidt}+benefit_{lidt}]}}$$

$$= \Pr(l_{idt}^{*})$$

The above is a well know aspect of multinomial models, which are often used to model discrete choice.

Appendix C: Neighbor Analysis

It is likely that participants include the rewards in other regions in their decision of where to participate. *A priori* it is ambiguous if a region with neighbors that have high rewards will make that region more appealing. Neighboring high rewards may act as a substitute or a compliment. We are ambiguous in our predictions of how that relationship will play out but acknowledge that rewards in surrounding regions will likely play a role and should be accounted for. Here we account for the other nearby options by including three additional covariates, all with respect to a region's immediate neighbors. We include the maximum reward on an adjacent region, the average reward of all adjacent regions, and the number of adjacent regions. The first should capture if participant intentionally avoid submitting from a region with a high-reward neighbor in order to submit form the high-reward neighbor. The second controls the appeal of moving clusters of high reward regions. The third is to control for any sort of systematic bias in average adjacent rewards or maximum adjacent rewards associated with having greater or fewer neighbors.

The results are found in Table C1. There does not seems to be any systematic relationship between submitting a survey in a region and the number of adjacent regions or the mean and maximum rewards offered in those regions. But there is some indication that participants leave especially low regions if they have neighboring regions with higher rewards. Importantly, accounting for these spatial components of the data does not change the estimates of responsiveness to rewards.

Table C1. The impact of a region's rewards and those if its neighbors on the log likelihood of submitting in that region

VARIABLES	(1) Ln(pr(region))	(3) Ln(pr(region))
Ln(reward)	0.354***	0.347***
	(0.0430)	(0.0563)
Ln(Number of adjacent regions)		-0.186
		(0.197)
Ln(Maximum reward in adjacent regions)		-0.0849
		(0.162)
Ln(Minimum reward in adjacent regions)		0.225**
		(0.0886)
Ln(Mean rewards of adjacent regions)		-0.0713
		(0.182)
Constant	-0.466***	0.0563
	(0.145)	(0.528)
Observations	5,472	5,472

Participant clustered and robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1