# 1 Context

The Kenya Universal Basic Income experiment randomly assigned subjects to one of four main conditions: control, lump-sum transfers, short-term streams of transfers, and long-term streams of transfers. Randomization was conducted at the village level in Bomet and Siaya counties. Randomization was conducted and outcomes were measured by a research team at Innovations for Poverty Action; transfers were implemented by the NGO GiveDirectly.

This document describes a set of outcomes we will examine and methods we will use to examine them. It is not intended to be exhaustive of either outcomes or methods, nor to describe the analysis that will ultimately be included in any particular paper or policy brief. Rather, we commit to producing and making publicly available in a single, easily-accessible place the results of all the analysis described here.

# 2 Outcomes

We define outcome families, and one or more outcomes within each, which we plan to examine. This list is not intended to be exhaustive of all the outcomes we may examine, but rather to reduce degrees of freedom where possible by committing in advance to (i) reducing the number of outcomes through aggregation, either by adding dollar values or by creating indices, and to (ii) pinning down the way we will make certain judgment calls, e.g. valuing household labor when calculating agricultural profits. For several families we include both a broader measure that is conceptually important (e.g. net worth) and also a narrower measure that we expect will be better-measured (e.g. the value of durables acquired in the last year). In the case of outcomes measured at the individual level we refer to individuals who were eligible at baseline to receive transfers.

- From the income family, **total annualized non-transfer income**. This includes (wage) labor income (household survey, section G.1), agricultural income (including the value of any output consumed) (household survey, section G.3), non-agricultural self-employment income (enterprise survey), non-enterprise capital income - return on financial assets (balances in household survey, section C - financial assets, multiplied by standard rates we can look up), interest on casual loans (household survey, section E), and rent (household survey, section C - land and real estate), etc. It does not include transfers from government (household survey, section D), transfers from NGOs (household survey, section D) including those from GiveDirectly, and transfers from other households (household survey, section D). Here and elsewhere, we will value agricultural output when necessary using the median of available local estimates of its price.

- From the consumption family, **the value of monthly food consumption** and **total annualized consumption**. The former includes the value of food, beverage and related items acquired to be consumed, consumed from own production, or consumed from gifts received in the last 7 days converted to a monthly basis (household survey, section B1), as well as expenditure on food and beverages outside the home in the last 30 days (household survey, Section B2, excluding Tobacco / Stimulants and Other Items). Total annualized consumption includes food consumption as well as all other consumption expenditure on all durable and non-durable consumer goods (household survey, section B), except that to avoid double-counting we will drop from this measure any items which also appear in the assets module (and this appear in net worth). It does not include taxes or contributions paid or transfers to other households (household survey, section D).

- From the wealth family, **the value of major assets acquired in the previous year** and **net worth**. We measure the former as the sum of the amounts spent on all household assets the household reports acquiring in the previous year which cost more than Ksh. 5,000 (household survey, section C), the analogous sum for farm assets (household survey, section G.3), and the analogous sum for enterprise assets (enterprise survey, section D, questions 11-13). We measure the latter as the sum of the value of household assets (household survey, section C), agricultural assets (household survey, section G.3), and non-agricultural enterprise assets (enterprise survey, Section D) plus the net value of all outstanding loans issued, i.e. loans issued minus loans received (household survey, Section E).

- From the labor supply family, **individual labor supply over the past year** including wage labor (household survey, section G.1), self-employment in agricultural enterprise (household survey, section G.3), and self-employment in non-ag enterprise (enterprise survey, Section C), and **individual time spent on working** broadly construed to include both remunerated work (including wage and self-employment) and unremunerated work (chores, shopping, childcare, etc.) from time use data.

- From the output family, **enterprise revenue** and **enterprise profit** from the enterprise census for non-agricultural enterprise and from the household survey (Section G.3) for agricultural enterprise..

- From the prices family, a **consumer price index** defined as the weighted sum of the prices of commodities sold in local markets (market survey) weighted by control group expenditure (household survey, Section B).

- From the subjective well-being family, the **CES-D depression index** based on household survey, section J, questions 1-20.

- From the family relationships family, **a domestic violence index** equal to the sum of normalized indicators for whether female respondents reported having (i) been threatened or (ii) been hit or had something thrown at them during the past 30 days (gender relations survey).

- From the community relationships family, a **social integration index** constructed from the household survey, section L, as the normalized sum of (i) the total number of items 1, 2, 3, 6, 7, and 8 to which the household responded either Yes, participant or Yes, position holder; (ii) the response to question 10, and (iii) the number of all yes responses to questions 12 and 13.

- From the children family, **an index of child anthropometrics** defined as the average of the weight-for-age and height-for-age z-scores of all children present and aged 2-12 years at baseline, and **years of schooling completed** for children present at baseline who would have made a primary to secondary transition during the period from baseline to endline if on normal grade progression.
- From the public goods family, **total contributions to public officials** over the past 12 months, equal to the sum of questions 5b, 5d, 7a, 7b, and 7c in the household survey, Section D: Remittances.
- From the externalities family, **a crime index** equal to the normalized sum of the responses to questions 3, 6, 9, and 11 in the village elder survey, Section F, and the value of **gross remittances sent** in the past year including money, goods or gifts (household survey, Section D).
- From the empowerment family, **willingness to accept to work on public works** during the off-season (household survey, Section G.4) and **an aspirations index** (household survey, section K) which is the Anderson weighted index of (i) the average of responses to question 1a on aspirations for kids educational attainment, (ii) question 2c on level of assets the respondent wants to achieve in 10 years, and (ii) question 3b on the level of income they want to achieve after 10 years.

Generally speaking, we express transfer amounts and monetary outcomes in PPP USD. If at any point we compare monetary values across survey rounds we will adjust these for inflation. Where we need to aggregate flow outcomes that we measured with different recall periods, we annualize. We will not trim outcomes to remove outliers; we will use validation and supervisor scrutiny during the data collection process to check on any values that look extreme. To create indices, we will use standard indexing procedures if these are available in the literature (e.g. for the CES-D), and otherwise will take the unweighted average of outcomes normalized by the control group mean and standard deviation.

# 3 Analytical methods

## 3.1 Sample definition and weighting

We estimate ITT effects, meaning that a units regressors are defined based on the treatment to which we think it should have been assigned (as opposed to the treatment GD delivered to it). These disagree in a few cases because research and GD in a few cases drew village boundaries differently. We report rates of disagreement but do not calculate IV TOT effects.

We plan to measure but omit from the analysis one village (the "pilot village) which started well before the rest of the project (in October 2016), was used by GiveDirectly as an operational pilot, and received substantial media coverage.

We survey and study at endline households that we were able to survey at baseline, including 10.5% that replaced households originally sampled for the study but which we were unable to locate or which declined to participate. Baseline replacement does not appear to have been substantially imbalanced on observables or differential by treatment arm. Where applicable, we will weight observations by their inverse sampling probability to obtain estimates that are representative for the larger frame from which they were sampled.

One of the comparisons we make is between households that received long-term as opposed to short-term transfers. The intent of this comparison is to isolate the effect of *anticipating* future transfers, but in some cases this may be confounded by the fact that children aged 15-17 at the time of initial enrollment were subsequently enrolled to receive transfers after turning 18 in the long-term, but not in the short-term arm. We will exclude households with children in this age range by default when making this comparison, and examine sensitivity to including them.

### 3.2 Unit of analysis

For outcomes measured at the level of the individual, we estimate effects at the individual level. For outcomes measured at the household level, we estimate effects at the household level. The interpretation of the latter effects is potentially complex because household composition may have changed since baseline, and these changes may be related to treatment status. If we find a significant rate of compositional change we will explore methods for adjusting for this and/or bounding treatment effects.

### 3.3 Migrant data

We plan to track migrants and gather analogous data on their earnings, expenditures, etc. For migrants living alone, we will add their reported quantities (e.g. income, expenditure) to the household total. For migrants living with another household (e.g. students living with a relative while attending school) we will attempt to measure their share of host household outcomes and add this to the sending household total. If feasible we will adjust these quantities for differences in the cost of living by location. If we decide to selectively track a random sample of migrants intensively, we will upweight their outcomes by these sampling probabilities.

### 3.4 Estimating equations

For count outcomes such as revenue or profit measured at the enterprise level, we first aggregate these to the level of the unit of agglomeration in which we measured them, which is either the village, the market, or the shopping center. We then model causal effects as follows.

For outcomes measured at the individual, household, and village level we will first estimate a simple model that includes only indicators for the treatment to which the corresponding village was assigned (LT, ST, or LS, with C the omitted category), conditioning on the randomization stratum of the village. For outcomes measured at the household level, we will also condition on indicators for the number of adults in the household at baseline and interact these with own-village treatment indicators, as impacts likely vary with the number of treated family members.

We will then estimate an augmented spatial model which also includes the share of households (i) nearby (defined below) and (ii) part of the randomization that were assigned to each treatment arm, conditioning on the share in each randomization stratum.[1] For these and other neighborhood specifications, we will also explore more flexible models in which outcomes depend (possibly non-linearly) on the number of treated households as well as or rather than the share, while controlling as flexibly as possible for the number of total households.[2]

For outcomes measured at the market or shopping center level, we will estimate spatial models as above but without own treatment status indicators (as markets and shopping centers were not assigned to treatment). We will check whether the existence of markets and shopping centers is endogenous to treatment and if so will need to further adjust this approach.

We define nearby in two ways. First, we will estimate models defining it as being within 2km of the unit of observation. For any observations with no neighbors inside that band, we replace the regressor with its mean and include an indicator for this. Second, we will explore data-based approaches to selecting among potential spatial models, noting that these approaches must also provide a method for inference after model selection.

In the case of a few outcomes (earnings, food consumption, and aspirations) where we expect the mean may not summarize well the distributional impacts, we will also examine the latter using median regression and distribution plots.

For all outcomes, we will condition on the baseline value of the outcome (if an exact or reasonably analogous variable is available), or its mean if missing (with indicator for missingness).

### 3.5 Aggregate effects and hypothesis tests

We will use estimation results to calculate the following aggregate effects:

- The own-village effect of each treatment (LT, ST, LS) vs nothing: this is simply the coefficient on the own-village treatment effect indicator, estimated with or without also modelling the neighborhood effect. This may be a biased estimate of the total effect to the extent there are cross-village spillovers.

- The total effect of each treatment (LT, ST, LS) vs nothing: the own-village effect of the treat-

[1] This criterion excludes from the calculation households in villages which were omitted because they were too large or too small, and 9 which were omitted because of high rates of baseline non-response. For any villages for which we lack the coordinates for specific households we will use reasonable approximations, e.g. the village center.

[2] For example, it may not be feasible to estimate dummies for every value of the number of neighbors but may be possible to dummy for fairly fine bins.

ment in question (where applicable) plus the average neighborhood effect of that treatment, i.e. the estimate coefficient on neighborhood treatment intensity times the mean treatment intensity (setting to zero the neighborhood effects of all other arms).

We will also calculate differences in the above measures between (a) LT and ST arms and between (b) ST and LS arms. The former tells us whether expectations of the future matter, while the latter tells us how different it is to get money in a stream as opposed to a lump sum.

We may find that some treatment conditions affect other resource flows into treated areas, e.g. other NGO activity, participation in government programs such as Inua Jamii, etc. If we find effects on these we will report them in order to help interpret the ITT effects.

### 3.6 Inference

We conduct inference using Conley standard errors, which allow for spatial correlation in the residuals. We may also explore approaches based on design-based uncertainty. We limit the number of tests by pre-specifying focal outcomes and aggregating many outcomes into sums or indices.

### 3.7 Heterogeneity

We will examine heterogeneity of impacts on our household outcomes along the following dimensions:

- Baseline gender of household head
- Baseline household wealth above / below median
- Baseline age of household head above / below median, as theory says spending patterns should vary over the life cycle.
- Baseline ownership of any non-agricultural enterprise
- County (Siaya vs Bomet)

## 4 Nudges

Within the main experiment, we also assigned households within each treatment arm to one of three “nudge conditions: a control arm, a “planning nudge, and a “savings nudge. The “planning nudge encouraged people to think about what they wanted their life to look like in the future and to describe, either verbally or in writing, goals for the next few years and plans for achieving those goals. The “savings nudge informed people about the availability and features of “M-shwari savings accounts linked to their M-Pesa mobile money accounts.

We will focus our analysis of the effects of these nudges on a list that includes (a) the focal outcomes defined above, and (b) if not otherwise included, additional outcomes that seem a priori likely to respond to the given nudge. Specifically, we will examine the impact of the savings nudge on net worth and its components (i.e. assets and liabilities), on liquid savings specifically, and on M-shwari savings specifically. We will examine the impact of the planning nudge on net worth and its components (i.e. assets & liabilities), measures of childrens educational status / attainment / performance, and measures of lean times from the food security & resilience module.

To estimate the effects of the nudges, we will regress outcomes on an indicator for the nudge assigned, spatial measures of neighborhood treatment intensity (as defined above), and the interaction between the two. We will then calculate the average total effect of the nudge, defined as the main effect plus the interaction with neighborhood treatment intensity multiplied by mean treatment intensity. We will conduct this analysis separately within each transfer arm.