Amendment to:
General Equilibrium Effects of Cash Transfers: Pre-analysis plan for Endline 3 (EL3) Child Mortality Analysis[1]

Dennis Egger, Grady Killeen, Johannes Haushofer,
Edward Miguel, Nick Shankar, and Michael Walker

August 16, 2024

Summary: This document outlines outcomes and regression specifications for estimating the effects of unconditional cash transfers on child mortality as part of the General Equilibrium Effects (GE) project in western Kenya. The project is a two-level randomized controlled trial of the NGO *GiveDirectly*'s unconditional cash transfer program. Transfers were distributed from 2014-16, with a first set of findings described in Egger et al. (2022). This amendment builds on an earlier PAP filed prior to analysis of data from a follow-up household census in which we identified births and child survival throughout the 653 villages within the study area. The original PAP describes the analyses that we conduct to understand program treatment effects on child mortality. This amendment describes analyses that we plan to conduct on a new round of household survey data that aims to identify the mechanisms through which cash transfers affected child mortality, and also details several methodological improvements that we have made since the original PAP was filed. This document is part of a series of pre-analysis plans that will be filed as part of the broader third endline (EL3) survey data collection activity for the GE project. We anticipate conducting analyses beyond those pre-specified here, and this document is not meant to preclude additional analyses.

[1] AEA Trial Registry: AEARCTR-0000505, https://www.socialscienceregistry.org/trials/505. Egger: University of Oxford, dennis.egger@economics.ox.ac.uk; Haushofer: Stockholm University, johannes.haushofer@su.se; Killeen: University of California, Berkeley, gkilleen@berkeley.edu; Miguel: University of California, Berkeley, emiguel@berkeley.edu; Shankar: University of California, Berkeley, nick.shankar@berkeley.edu; Walker: University of California, Berkeley, mwwalker@berkeley.edu. We thank Layna Lowe for excellent research assistance in developing this pre-analysis plan. Funding from Open Philanthropy, recommended by GiveWell, is supporting this work.

1 Introduction

This document amends and augments the analysis plan for child mortality data collected between 2014 - 2024 (Baseline, Endline 1, Endline 2, Endline 3), as part of the General Equilibrium Effects (GE) project, a randomized controlled trial of an unconditional cash transfer program by the NGO GiveDirectly (GD). We omit a discussion of the intervention since it is described in detail in the original PAP document (Egger et al., 2023) and elsewhere.

This analysis is primarily focused on data related to child mortality collected in Endline 3 (about 9 years after the transfer) although where appropriate, we will make use of earlier rounds of data as well. As such, this document draws heavily on previously filed pre-analysis plans and existing working papers and published papers, updating earlier regression specifications and outcomes. These earlier plans are described in the original child mortality PAP.

Broadly, there are three components to this analysis:

i. Studying the impacts of cash transfers on child mortality on recipient households (and their neighbors) present in the study area at the time of the intervention,
ii. Identifying the impacts of cash transfers on the causes of child deaths using machine learning methods on verbal autopsy data,
iii. Exploring the mechanisms (economic and non-economic) through which cash transfers affect child mortality.

The original PAP focused on the first two points. This amendment focuses on how we aim to use newly collected data since the original PAP to explore the (economic) mechanisms through which cash transfers affect mortality. This is driven by preliminary analysis of the original PAP outcomes using data from the Endline 3 (EL3) household census which suggest sizable reductions in child and infant mortality among households that received cash at the time of transfer receipt, but does not show evidence of reductions among non-recipients or in the long run.

The primary data source for this analysis of mechanisms is a new round of household surveys in all study villages (the Endline 3 (EL3) Household Survey) and data from health facilities in Siaya County, in particular surveys that we are completing with operational facilities. As described in the initial PAP, we make use of the EL3 Household Census to estimate the core treatment effects on mortality. The EL3 household surveys are from a representative sample of households in study villages. The survey of health facilities visits currently operational health facilities to measure the services that they offer, their staff, and whether they were open during the cash transfer period. We append several administrative and GIS datasets to this survey data, which we describe in Section 2.5.

We will also consider data collected as part of previous survey rounds. The representative household sample includes households that were residents in the study area at baseline, as well as some households that moved into the study area after the intervention (‘new households’). We carefully track and survey households that have remained in the area and those that have moved away. We describe the sample to be used to measure the effects of cash transfers on child mortality in detail in the initial PAP (see Section ‘5.2

Econometric specification of impact analysis' in Egger et al., 2023). We further discuss how we reach a representative subset of this sample below in Section 2.4.

Endline 3 household survey data collection began in June 2024. We are filing this amendment shortly after the launch, with about 44% of villages from wave 1 surveyed (wave 1 is the first half of two representative waves). We have yet to complete revisits and long-distance tracking, so this represents under 20% of total expected observations. Moreover, to date, no member of the research team has linked the data to treatment indicators nor estimated any treatment effects; the only examination of the data has been to check for missingness, outliers and other data collection or processing issues. We also have not yet examined whether all of our primary outcomes (described below) have suitable variation, in part because we need to complete health facility visits before we can construct some measures. If we discover that some outcomes have minimal variation, we plan to drop them from consideration (and exclude them from multiple hypothesis corrections).

After filing this plan, we will begin estimating pre-registered effects for the mechanisms. We will file a separate PAP covering analysis of outcomes not related to child mortality effects or mechanisms, and as such we will only examine effects in the EL3 survey data on the outcomes specified in this amendment until the other PAP is filed.

1.1 Relation to previous work

This amendment builds on the original child mortality PAP filed in 2023 (Egger et al., 2023) and a series of earlier pre-analysis plans filed for the GE project as part of data collection in previous rounds (i.e., short-term/Endline 1: Haushofer et al. 2016, Haushofer et al. 2017a, Haushofer et al. 2017b, Haushofer et al. 2018, Walker 2017; Endline 2: Egger et al. 2021a, b, Orkin and Walker 2021, Egger et al 2022). Moreover, it builds on analyses published in Egger et al. (2022), as well as ongoing analyses of Endline 2 data, which suggests there is some persistence in economic gains for household consumption and assets, among other effects. In this amendment, we err on the side of brevity to avoid repetition, as much of the detail and thought development can be found in these earlier documents and this document is not intended to be read independently from the original PAPs or papers.

2 Research Design

2.1 Sampling and treatment assignment

The GE project takes place in Siaya County, Kenya, a rural area in western Kenya bordering Lake Victoria. Siaya County is predominantly Luo, the second largest ethnic group in Kenya. GD selected both Siaya County and a region within Siaya County based on its high poverty levels and identified target villages for expansion; in practice, these were all villages within the region that a) were not located in peri-urban areas and b) were not part of a previous GD campaign. This gives a final sample of 653 villages, spread across 84 administrative sublocations (the unit above a village), and 3 constituencies.

We use a two-level randomization in order to generate variation that can be used to identify spillover effects. We randomly assigned sublocations (or in some cases, groups of sublocations) to high or low

saturation status. Then, within high saturation groups, we assigned 2/3 of villages to treatment status, while within low saturation groups, we assigned 1/3 of villages to treatment status. As noted above, within treatment villages, all eligible households received a cash transfer.

At baseline, we censused about 65,000 households. At endline 3 we tracked almost 80,000 households and completed censuses with over 62,000 permanent households and 5,000 holiday homes (i.e., residences that are not inhabited regularly but typically only during holiday periods). A random sample of households were drawn for detailed household surveys at baseline; these households were followed up at Endline 1, Endline 2 and will be followed up with at Endline 3. Additionally, at Endline 2, we added a random sample of new households to the sampling frame. We will survey all of these households at endline 3, and also include a random sample of households that were first identified in the Endline 3 census. The addition of new randomly chosen households will allow us to maintain a representative sample of households in the study area.

2.2 Intervention

GD provides unconditional cash transfers to poor households in rural Kenya, targeting (for villages in the study) households living in homes with thatched roofs, a basic means-test for poverty. In treatment villages, GD enrolls all households in treatment villages meeting its thatched-roof eligibility criteria (“eligible" households); approximately one-third of all households are eligible. No households in control villages receive transfers. Eligible households enrolled in GD's program receive a series of 3 transfers totaling about USD 1,000 via the mobile money system M-Pesa[2]. This is a one-time program and no additional financial assistance is provided to these households after their final large transfer. For details on the intervention, see Egger el al. (2022) and Haushofer et al. (2017a).

2.3 Data and Instrument

The analyses outlined in this document will be primarily based on a new round of data collected in 2023-25 (Endline 3), roughly 8-11 years after the GD cash transfers went out, as highlighted in the approximate timeline below:

- 2014 - 2015: Baseline (pre-intervention, BL)
- 2014 - 2016: Intervention (distribution of cash transfers)
- 2014 - 2017: Midline
- 2016 - 2017: Endline 1 (EL1)
- 2019 - 2022: Endline 2 (EL2)
- 2023 - 2025: Endline 3 (EL3)

Activities for Endline 3 will include a household census, enterprise census, household surveys, and enterprise surveys. The household and enterprise censusing activities were conducted at the same time,

[2] The total transfer amount is 87,000 Kenyan Shillings (KES). The average exchange rate from 9/1/14 to 4/30/16 was 97 KES/USD.

and lasted from April 2023 to November 2023. The survey activities will also be conducted concurrently with each other. They began in June 2024 and we expect them to conclude in mid-2025.

We randomly assigned half of households within each village to be visited in wave 1, projected to last from June 2024 through approximately November 2024, or wave 2, projected to last from roughly November 2024 to mid-2025. This ensures that the data from each wave is representative, so we may in the future present initial results using data from wave 1, then add in wave 2 data when data collection is complete.

The child mortality analysis we originally pre-specified primarily made use of data from the household census in our study area, although as mentioned in other parts of the original PAP, we also bring in information from EL2 household surveys. We also stated in the original PAP that we planned to bring in information from the EL3 household surveys, especially to fill in child mortality and verbal autopsy data from households that moved away from the study area (who are not covered in the local household census activities). We encourage readers to examine the original child mortality PAP for a detailed discussion of these components. This amendment focuses on additional analysis that we plan to conduct with the EL3 household surveys, namely analysis of mechanisms generating treatment effects, and also fills in details of the sampling strategy used for the EL3 household survey.

The EL3 household survey has two modules that are focused on child mortality. First, for households missed during the census (such as those that moved out of study villages), we include the full EL3 household census module that captures all births, records under-5 deaths, and conducts verbal autopsies. The original PAP details this component. Second, for all surveyed households from Baseline, we include questions about antenatal, delivery and postnatal care and spending designed to shed light on changes in behavior that could rationalize a treatment effect on child mortality. These questions were added since the original PAP, and this amendment focuses on the outcomes that we plan to construct with those survey questions.

2.4 Endline 3 household survey sample

The EL3 household survey reaches a representative sample of households present in the study area as of the EL3 household census and a representative sample of households present in the study area at baseline. Either of these representative samples may be obtained via appropriate inverse probability weights as described below. We consider the following categories of households present in the endline 3 census:

a) **Baseline households: ~70% of EL3 censused households**
   i) Households that were eligible (for transfers) at Baseline
      - Stayers: Still living in the study area
      - Movers: Now living outside the study area. (These are not covered in the census, but will be tracked as part of the survey activity, allowing us to augment our census-based estimates of infant and child mortality effects).
      - Movers since Endline 2: Now living outside the study area, but moved after Endline 2. As with above, they are not covered in the census since they are no longer living locally.

- Endline 2 movers that returned to the study area: Households that moved out of the study area between baseline and Endline 2, but then moved back to the study area between Endline 2 and endline 3.

ii) Households that were ineligible (for transfers) at Baseline
- Stayers: Still living in the study area
- Movers: Now living outside the study area
- Movers since Endline 2: Same as for eligibles
- Endline 2 movers that returned to the study area: Same as for eligibles

b) **Endline 2 households: ~10% of EL3 censused households**

i) Endline 2 movers into the study area: Households that lived outside of the study area at baseline, but moved into the study area by the Endline 2 census.

ii) Split-off households between BL and EL2: Households where children of baseline households established new households prior to the Endline 2 census, or where the primary male or female established separate / additional households by the Endline 2 census.

iii) "Ghost" households captured at Endline 2: These are defined as households that reported being present in the study area at baseline during the Endline 2 census, but they were not surveyed during the Baseline or Endline 1 data we collected. We will use the same "ghost" term below to refer to households that were missed through Endline 2 but reported being present in the study area by Baseline, Endline 1, or Endline 2, and were not surveyed until the Endline 3 census.

c) **Newly surveyed Endline 3 households: ~20% of EL3 censused households**

i) "Ghost" households captured at Endline 3: These are defined as households that reported being present in the study area at Baseline or Endline 2 during the Endline 3 census, but they were not surveyed during Baseline, Endline 1, or Endline 2.

ii) Split-off households after EL2: Households that split off from Baseline households *after* Endline 2 (category b part ii captures those that split off *before* Endline 2, whereas this category captures those that split off after). We define any household as split-off where children of Baseline households established new households, or where the primary male and female established separate / additional households. This is separated from category b part ii since households that split off before EL2 are included in the panel sample, but those that split off after EL2 are being targeted for surveys for the first time.

iii) New households: Households that moved into our study area after Endline 2 (i.e., were established AFTER Endline 2, but not split off from existing households and are not "ghost" households).

We then consider these categories of households present at Endline 3 plus information from prior rounds to construct the Endline 3 household survey sample. We begin by including all households included in the targeted sample for the Endline 2 household survey. As detailed in the Endline 2 PAPs, this contains a representative sample of households present in the study area at Baseline and at Endline 2 (including households that moved out of the study area between Baseline and the EL2 household survey). The panel data on Baseline households surveyed in each round is the basis of the majority of the analysis in this PAP, while the other categories ensure that we obtain a representative sample of the 2024 population of the study area.

We then draw a representative sample of household types that were newly captured in the EL3 census: Baseline and EL2 ghost households (i.e., those that were not censused at EL2 but claim that they were resident in the area at Baseline or EL2), split-off households since EL2, and new households since EL2. The current project budget allowed us to include 1,500 households from these household types (ghost, split-off households, and new households) in the EL3 household survey sample,. We draw a representative sample of these household types (ghost, split-off or new), stratified by sublocation, by multiplying the share of each household type within the sublocation by the number of additional surveys our budget affords within the sublocation. This ensures that we sample each of the household types that were not included in the EL2 survey with equal probability within a sub-location. The share of ghost, split-off and new households sampled within each sub-location varies since some areas may be more prone to certain types (e.g. migrants), so we wanted to ensure that no types were not undersampled within sublocations where they are prevalent.

We then randomly sort all households of each type (ghost, split-off or new) within the sublocation from the Endline 3 household census and include the first N observations in the sample, where N is the number of households of that type within the sublocation targeted for household surveys. The number of households targeted from each category is listed in the following paragraph.

This approach yields the following overall sampling strategy for the Endline 3 survey:

a) Baseline households: N = 8,409

i. Households that were eligible at Baseline (stayers & movers): N = 5,663
   - Households per village (average): 8.67

ii. Households that were ineligible at Baseline (stayers & movers): N = 2,746
   - Households per village (average): 4.2

We targeted all Baseline households in the target sample for household surveys at Endline 1, unless they withdrew consent during the Endline 2 survey. This includes all households targeted but not surveyed at Endline 1, and households that moved outside the study area since Baseline (more details can be found in Haushofer et al. 2017a).

If households had both a primary male and female, we randomly selected either the male or female to be the “target” respondent; if we could not reach the target respondent, but the spouse/partner was available, we surveyed the spouse/partner. For households that had previously been surveyed (either at Baseline or Endline), we targeted the respondent last (i.e., most recently) surveyed.

We conducted extensive tracking of moved-away households, sending team members across Kenya. For those moving to other countries, or to a few counties within Kenya where the security situation did not permit travel, we attempted to follow-up over the phone.

b) Endline 2 households: N = 2,098

i. **New households established by Endline 2: N = 629**

(a) Households per village (average): 0.96

We targeted a random sample of max(1, 24%) of households newly established between Baseline and EL2 in each village for the EL2 household survey.

ii. **Split households (eligible & ineligible) censused at Endline 2: N = 894**

(a) Households per village (average): 1.37

We targeted all households split-off from an eligible or ineligible household in our Endline 1 target sample at the time of the Endline 2 survey. Since the Endline 1 target sample is representative of all Baseline households, this yields a representative sample of households that split-off as of Endline 2.

iii. **Ghost households reached at Endline 2: N = 575**

(a) Households per village (average): 0.88

We targeted a random sample of max(1, 24%) of ghost households reached at Endline 2 in each village for the EL2 household survey.

c) Endline 3 households: N = 1,494

i. **Households newly established after EL2: N=373**

- Households per village (average): 0.57

See the details of the sampling protocol in the prior section (for this and following groups).

ii. **Endline 3 ghost households: N=599**

- Households per village (average): 0.91

iii. Endline 3 split households: N=522

- Households per village (average): 0.80

Most of our analysis of child mortality will focus on a representative sample of households present at Baseline. We only ask the questions related to child mortality to households that have been tracked since Baseline, so “ghost” households are excluded from most analysis. Since our focus is on the period immediately following the distribution of cash transfers, we also do not ask the questions from “split-off”

households since they were likely part of the original Baseline household during the relevant timeframe, so we focus on the survey of the Baseline household to prevent double counting or data quality issues.

2.5 Data About Health Facilities

In addition to the EL3 household survey, we plan to make use of original data collected about health facilities in Siaya County, Kenya combined with administrative data sources. In 2023, we scraped data from the Kenya Master Health Facility List on all registered facilities operating in Siaya County. We have begun visiting all facilities Level 3 or higher, which covers health clinics and hospitals, including all facilities likely to offer antenatal or delivery services. We collect data on the services related to births that are offered plus information about the staff that work at the facility, and especially if it is staffed by a medical doctor (physician). We also record when the facility opened so we may exclude those which began operating after the period of analysis.

This list of facilities is then populated into the EL3 household survey so that we can identify the specific facilities where individuals received care. We also allow respondents to enter other facilities beyond those on this list and will attempt to survey any that are commonly mentioned but not found in our initial roster of clinics.

Travel times to clinics will be estimated using the GPS location of a facility combined with information about travel speeds. We are currently attempting to track travel speeds on every trip made by our study field team for EL3 data collection to accurately estimate travel times under real-world conditions. However, we may also use OpenStreetMap data, or a shapefile of documented roads before the UCT period, to construct these measures.

We may also complement this data with public information about clinic locations in some instances (e.g. from the WHO or KEMRI) to check for robustness to clinic locations just outside of Siaya County, or to bring in additional information about clinics for exploratory analysis. This data is not used for the primary or secondary outcomes below, so we do not include a detailed discussion here.

3 Refinements to the empirical strategy

In the original PAP (Egger et al., 2023), we discussed that we may make refinements to our empirical specifications to keep up with trends in the econometrics literature. We discuss several such refinements that we plan to make here. All of our empirical strategies are broadly similar, so we view these as relatively minor modifications to improve efficiency rather than departures from the overall pre-specified empirical strategy.

First, we planned to report spatial HAC standard errors with a 10 km uniform kernel (Conley, 1998, 2008). A known limitation of the uniform kernel is that it is not positive definite, so sometimes yields complex standard errors. We found that this happened with some frequency, giving us ill defined test statistics for some outcomes. As a result, we plan to focus on the positive definite kernel

$$K_{ij} = 1(d_{ij} \leq 10) \cdot \left(1 - \frac{d_{ij}}{10}\right)^2 \quad (1)$$

where 1() is the indicator function and d_{ij} is the distance in kilometers between observations *i* and *j*. We found in practice that this typically produces similar or slightly larger standard errors than the uniform kernel, but we will verify that results are not highly sensitive to this choice. The other advantage of this kernel is that we may use the same kernel to conduct inference with the spatial wild bootstrap introduced in Conley et al. (2023), so this allows us to report standard errors in a more consistent manner across methods.

Second, Equations (2) - (4) in the original child mortality PAP indicated that we planned to instrument for the per capita amount of cash received in a village or r-km band with the share of eligible households at baseline that were treated within that village or r-km band. If treatment effects are homogeneous, the efficient vector of instruments includes both the share of eligible households treated and this value times the share of households that were eligible in each village or r-km band, controlling for the share of households that were eligible at baseline. The share of households that were eligible is endogenous, but this value interacted with the share of eligible households treated is exogenous conditional on the control variable being included and this allows us to efficiently model the first stage, generally yielding statistical power gains. Hence, we plan to report results with this vector of instruments, which we call the "interacted" set of instruments for convenience, alongside the original vector. This essentially uses the same experimental variation while better accounting for first stage heterogeneity. The logic is similar to that in Abadie et al. (2023) which shows that similar instruments often yield better asymptotic mean squared error.

Third, we wrote a Python package, called geGMM, which allows for flexible GMM analysis to estimate Equation (4) in the initial PAP.[3] As a result, we may report estimates with efficient GMM in some cases alongside those using two-stage least squares.

We may consider additional methodological refinements in the coming years to keep up with state of the art econometric techniques, but we aimed to highlight these three improvements (since the filing of the original PAP last year) that are likely to be used in the primary analysis.

We plan to report reduced-form effects and instrumental variable estimates on both eligibles and ineligibles, and pooled instrumental variable estimates as outlined in equations (1) - (4) of the original PAP across the outcomes we define below, subject to the refinements discussed in this section. We omit a full discussion of these equations for brevity and refer readers to the original PAP for details (Egger et al., 2023). We may also consider effects over time as outlined in section 5 of the original child mortality PAP.

[3] Instructons for installing and using the code are available at https://github.com/gkilleen33/gkilleen33.github.io/tree/master/packages/geGMM

4 Outcomes of interest

The main measures of interest, as detailed in the initial PAP, are infant mortality and child mortality. In addition, the VA tool employed in this study primarily leads to two distinct outcomes: a determined cause of death for each individual and the cause-specific mortality fraction (CSMF), and we will also use these two measures. These outcomes as detailed in the original PAP remain the core focus of the analysis, and we do not replicate discussion of them here. This amendment discusses the new outcomes that we plan to examine. The outcomes in this amendment are being specified after we have already estimated effects on the outcomes specified in the original PAP, with preliminary results showing large reductions in infant and child mortality among cash transfer recipient households during the period in which cash was distributed, but little in the way of persistent impacts. This has motivated us to add household survey questions and collection of health clinic data regarding potential mechanisms. This amendment discusses the outcomes we plan to construct with this data plus the data about health facilities that we have collected.

The following outcomes will primarily stem from Endline 3 household survey data and the survey of health facilities, although we plan to to integrate this with some aspects of previously collected data. In all cases, due to the focus of this analysis on assessing potential mechanisms behind infant and child mortality impacts, the population to be analyzed in this measure consists of children who were born at least 1 year before the start of Endline 3 census data collection in order to have a consistent population for both the numerator (children who have died) and the denominator (children who have died plus children who have survived until the age of 1) of the infant mortality calculation. As our preliminary results from the Endline 3 census indicate steep declines in infant mortality during the period contemporaneous with the release of cash transfers (2015-17) but negligible reductions in mortality following the release of the transfers (2018-21), we plan to examine the mechanism outcomes below separately for each of these distinct periods, as well as for the pre-period (2011-14) as a falsification check.

4.1 Primary outcomes of interest

The six primary outcomes of interest are largely based on prenatal, delivery, and postnatal care recommendations from the World Health Organization (WHO). As the overwhelming majority of women even in low-resource settings such as rural Kenya receive at least some antenatal and postnatal care (and a substantial majority today deliver in a health facility as opposed to at home), we focus on assessing whether mothers were able to receive enough quality care to align with global health recommendations. We also aim to assess the impact of the transfers on total household spending on the birth, an intensive-margin outcome that may correlate with the quality of care the child received in-utero, during delivery, and postpartum.

We view these outcomes as a distinct family from those filed in the original mortality analysis PAP (Egger et al. 2023) since they examine specific mechanisms, rather than the headline mortality effects. In addition, we have already estimated effects from the original PAP and view these outcomes as designed to test the mechanisms driving mortality reductions that we observe from the census data. We therefore plan to apply multiple hypothesis corrections across the primary and secondary outcomes filed here separately from those in the original PAP.

4.1.1 **Receipt of at least four (4) antenatal care visits**: Birth-level variable that takes the value 1 if the mother received at least four (4) antenatal care (ANC) visits, and 0 otherwise. Receipt of at least 4 ANC visits is the recommendation of the World Health Organization's (WHO) Focused Antenatal Care Model (FANC), which was designed to promote quality antenatal care in resource-constrained settings where the traditional ANC model (comprising 7-16 visits) would be a challenge to implement (Mchenga et al. 2019). While the WHO has since updated its ANC guidelines to recommend 8 visits, only 5% of women in Kenya currently meet that standard and the new WHO standard has been considered by some health professionals as infeasible in many low-resource settings (Jiwani et al. 2020, Mchenga et al. 2019). We thus designate the receipt of at least 8 ANC visits as a secondary outcome.

4.1.2 **Postpartum length of stay in health facility of >24 hours**: Birth-level variable that takes the value 1 if the mother delivered the birth in a health facility and stayed in the facility for >24 hours, and 0 otherwise. This variable accords with the WHO's 2022 Recommendations on Maternal and Newborn Care for a Positive Postnatal Experience. Monitoring of the mother and newborn by health professionals throughout the day following birth is imperative as postpartum complications, many of which cannot be immediately anticipated, are most likely to occur in this period (Campbell et al. 2016).

4.1.3 **Receipt of at least three (3) postnatal care visits within the first six weeks of birth**: At least three (3) postnatal visits within the first 6 weeks of birth for mothers delivering in a facility (and at least 4 postnatal visits within the first 6 weeks of birth for mothers delivering at home) is recommended by the WHO's 2022 Recommendations on Maternal and Newborn Care for a Positive Postnatal Experience. For mothers delivering in a facility, this birth-level variable therefore takes the value 1 if the mother received at least 3 postnatal care visits within the first 6 weeks of birth, and 0 otherwise. For mothers delivering at home, this birth-level variable takes the value 1 if the mother received at least 4 postnatal care visits within the first 6 weeks of birth, and 0 otherwise.

4.1.4 **Delivery in a hospital**: Birth-level variable equal to 1 if the birth was delivered in a hospital and equal to 0 otherwise. Delivery in a clinic or dispensary, which typically do not possess the same level of physical resources and skilled professionals as hospitals, would set this variable equal to 0. We include delivery in any health facility, a less stringent standard of care met by a considerable majority of Kenyans but nevertheless considered as a priority indicator for neonatal health, as a secondary outcome of interest.

4.1.5 **Visit to a facility staffed by a medical doctor (physician)**: Birth-level variable equal to 1 if the mother visited a health facility that employs a doctor for antenatal care, delivery or postnatal care. This variable is constructed by matching the facility the respondent reports visiting to our surveys of health facilities.[4] Recent research, including by Okeke (2023) in a rural Nigerian setting, indicates that health facilities assigned a doctor subsequently saw substantial declines in infant mortality, suggesting that the presence of a doctor may be important for child survival.

[4] One potential limitation of this approach is that we miss any facilities that may have closed prior to 2024. However, doctors are typically employed at large facilities like hospitals that are unlikely to have closed (based on our discussions with health professionals in Kenya). We therefore plan to code this variable to 0 if the facility the respondent visited has closed unless we find information suggesting they had employed a doctor during the relevant time period.

4.1.6 **Total household spending on antenatal, delivery, and postnatal care**: Birth-level variable equal to the sum of the household expenditure oncost of antenatal care for the pregnancy, cost of the child's delivery, and cost of postnatal care for the child. While higher spending on care is not unambiguously a positive outcome (since it could endogenously reflect medical complications or other factors), it is likely generally indicative of better care in a resource-constrained context like the study setting. The variable includes the cost of transportation and all payments to staff and other personnel, but not the cost of health insurance payments. Total household spending on antenatal, delivery, and postnatal care are separately included as three secondary outcomes. (As noted, a potential concern with this outcome measure is that improvements in nutrition, maternal health, or housing quality in treated areas could reduce the likelihood of medical complications, leading to less spending on maternal and infant care. Since expenditures on antenatal care appear more likely to occur prior to the development of any medical complications, we may focus on antenatal expenditures if we see any evidence of this trend.)

4.2 Secondary outcomes of interest

Some secondary outcomes were mentioned above and we re-list them here for completeness, together with several additional secondary outcomes.

Secondary Family 1: Amount and quality of care received

4.2.1 **Receipt of at least eight antenatal care visits**: Birth-level variable that takes the value 1 if the mother received at least 8 antenatal care (ANC) visits, and 0 otherwise.

4.2.2 **Delivery in a health facility**: Birth-level variable equal to 1 if the birth was delivered in a health facility and equal to 0 otherwise. Over 87% of mothers in our Endline 2 survey gave birth in a health facility, so institutional delivery is widespread in our setting. However, institutional delivery is a priority indicator for maternal and child health, and it is possible the transfers enabled some remaining mothers to deliver in an institution.

4.2.3 **Antenatal care at a facility staffed by a medical doctor**: Primary outcome 4.1.5 but restricted to antenatal care.

4.2.4 **Delivery at a facility staffed by a medical doctor**: Primary outcome 4.1.5 but restricted to delivery.

4.2.5 **Postnatal care at a facility staffed by a medical doctor**: Primary outcome 4.1.5 but restricted to postnatal care.

4.2.6 **Delivery at a facility capable of performing a cesarean section (C-section)**: Similar to primary outcome 4.1.4, the survey of clinics will allow us to determine if the respondent gave birth in a facility that can perform a cesarean section.

Secondary Family 2: Investment in healthcare access

4.2.7 **Health insurance coverage**: Birth-level variable equal to 1 if the mother possessed health insurance coverage when pregnant with the child and 0 otherwise.

4.2.8 **Travel time to receive care**: Total travel time to receive care across antenatal, delivery and postnatal services. Estimated by combining information on facility locations with estimates of travel times, then summing over all visits.

4.2.9 **Total spending on care excluding delivery**: Birth-level variable equal to the sum of the household expenditure oncost of antenatal care for the pregnancy and cost of postnatal care for the child. The variable includes the cost of transportation and all payments to staff and other personnel, but not the cost of insurance payments. We include this outcome as a secondary check on the primary outcome 4.1.6, as deliveries are sometimes subsidized by government programs and antenatal/postnatal care may hold outsized importance for child health outcomes relative to their share of total spending on care.

4.2.10 **Total spending on antenatal care**: Birth-level variable equal to the sum of the cost of antenatal care for the pregnancy. In addition to user fees and the cost of drugs and tests purchased, the variable includes the cost of transportation and all payments to staff and other personnel, but the variable does not include the cost of insurance payments.

4.2.11 **Total spending on delivery**: Birth-level variable equal to the sum of the cost of delivery for the child. In addition to the out-of-pocket cost of care, the variable includes the cost of transportation and all payments to staff and other personnel, but the variable does not include the cost of insurance payments.

4.2.12 **Total spending on postnatal care**: Birth-level variable equal to the sum of the cost of postnatal care for the child. In addition to the out-of-pocket cost of care, the variable includes the cost of transportation and all payments to staff and other personnel, but the variable does not include the cost of insurance payments.

4.3 Exploratory outcomes

We will further examine several other potentially important outcomes, which we view as exploratory.

4.3.1 **Miscarriage outcomes**: In the Endline 3 household survey, we expanded the miscarriage module primarily to examine whether miscarriages rates rose (or fell) substantially in cash transfer treatment areas, and/or in places where infant mortality fell. As such, we plan on examining the number of miscarriages, the miscarriage rate, and the survival rate from conception as exploratory outcomes. One note is that measurement issues may exist with tabulating miscarriages, particularly when asking respondents other than the pregnant woman of interest. We thus treat this as an exploratory aspect of the analysis to generate suggestive evidence on a potential confound.

4.3.2. **Travel time to receive antenatal care**: Similar to secondary outcome 4.2.8 but only for antenatal care.

4.3.3. **Travel time to receive delivery care**: Similar to secondary outcome 4.2.8 but only for delivery care.

4.3.4. **Travel time to receive postnatal care**: Similar to secondary outcome 4.2.8 but only for postnatal care.

4.3.5. **Delivery at a publicly-run (vs. privately-owned) facility**: In addition to studying the primary and secondary outcomes of hospital and facility delivery, we intend to examine whether shifts occurred in the public versus private composition of where institutional deliveries occurred.

4.3.6. **Delivery outside of Siaya County**: We intend to explore whether the share of births delivered outside of Siaya County (the study region) changed, with a focus on examining whether more births were delivered in larger cities/towns with more sophisticated health facilities, such as in Kisumu or Busia Town.

4.3.7. **Number of distinct facilities from which care was received:** We intend to explore whether the total number of facilities from which antenatal, delivery, and postnatal care was received changed in treatment areas.

4.3.8 **Total spending on care for deceased vs. non-deceased births**: We intend to examine total spending on care separately for children who passed away (as neonates, within the first year of birth, and within the first 5 years of birth) and for children who remained alive. It is possible that spending rose differentially within the categories of healthy births and less-healthy births, and that the composition of healthy versus less-healthy births also changed due to treatment (i.e., the treatment may have increased the share of healthy births, which could require less expenditure for health care). As a result, we may examine total health spending while re-weighting healthy versus less-healthy births as a way of accounting for the composition of births when assessing impacts on spending.

4.3 Outcomes of interest from other data sources

The aim of the outcomes presented in Sections 4.1-4.2 is broadly to test whether the cash transfers improved access to healthcare near the time of birth. We view this as an important channel through which cash may have affected infant mortality and child mortality, as well as the cause of death. However, we also plan to test for the presence of other mechanisms that could help to explain mortality reductions.

We have already begun to examine, or plan to examine, the role of maternal and infant nutrition (as measured in consumption and food security questions), mother's labor supply during pregnancy (by trimester, from the labor module in the household survey), and housing quality measures as potential channels through which the cash transfers also affected infant and child mortality. These mechanisms are being assessed using Baseline, Endline 1 and Endline 2 household survey data (due to the proximity of these earlier surveys to the period in which cash transfers were distributed), plus administrative data sources in some instances. We may also explore additional channels beyond those already listed here, and

when we do so will attempt to discipline our approach and the number of mechanisms examined by focusing on those with a strong basis in existing medical or public health research.

We view outcomes outside of the clinic and health care access measures detailed above as distinct from those in Sections 4.1-4.2, so plan to treat them as a separate family when applying corrections for multiple hypothesis testing. The measures mentioned above in this subsection (i.e., mother's labor supply) also employ data that we have already been examining –rather than a prospective survey with new data designed to explore pre-specified mechanisms, and this is an important distinction. We do not yet know the full set of outcomes that we may examine or how each measure will be constructed, and as such, outcomes in this section are inherently more exploratory. However, we do not view them as less conceptually important or likely to explain the central infant and child mortality results, and plan to prominently include them in the discussion of findings.

5 Analysis

Other than the relatively minor refinements to the empirical strategy documented in Section 3 above, the proposed econometric analysis follows the approach specified in our initial child mortality PAP, which itself largely follows Egger et al. (2022). As in the original PAP, our examination of potential mechanisms will focus on households that were present at Baseline and that remain present at Endline 3, and we consider effects on eligible recipient households to be the primary effects of interest. We will utilize the primary specifications from our original PAP (i.e., Equations (1), (2), and (3)) as our primary specifications for the analyses here (incorporating the refinements noted above). For more details on these specifications as well as additional details on our econometric approach (such as on balance and attrition and the computation of pooled effects), we encourage readers to examine the original child mortality PAP.

5.1 Multiple inference adjustments

Several of the primary outcomes discussed above (mechanisms underlying infant and child mortality effects) are likely to be correlated. To account for multiple inference for the primary outcomes, we will make use of the Romano-Wolf multiple testing correction, which asymptotically controls the Family-wise Error Rate (FWER). We plan to use a resampling based approach, rather than an analytic approach to limit the FWER, because it is able to mimic the dependence structure between the primary outcomes that we expect to see in the data, yielding better statistical power.

For secondary outcomes, we will follow Haushofer et al. (2017a) for the treatment of multiple inference adjustment, namely calculating sharpened q-values (i) across primary and secondary outcomes (as defined in section 4) following Benjamini, Krieger, and Yekutieli (2006) to control the false discovery rate (FDR) within each family of secondary outcomes (the two families listed in Section 4.2, plus the two families defined in Egger et al., 2023). The FDR controls for the proportion of false positives, which is relevant if one is interested in the proportion of all outcomes affected by treatment. Rather than specifying a single q, we report the minimum q-value at which each hypothesis is rejected, following Anderson (2008). We will report both standard p-values and minimum q-values. We will apply the correction separately for each hypothesis test described in Section 4.1 of Haushofer et al. (2017a). We note that norms around multiple

testing are still evolving in economics, and through the above methods seek to follow current best practices.

As discussed earlier, we view the outcomes in this amendment as a distinct family from those filed in Egger et al. (2023). We therefore plan to apply corrections for multiple hypothesis testing separately for the families of outcomes (mechanisms) included in this amendment. Furthermore, we view the outcomes in Section 4.3 as distinct from those defined in Section 4.1, so we will also apply corrections separately for those outcomes.

5.2 Heterogeneity analysis

Egger et al. (2023) pre-specifies heterogeneity analyses based on child gender, maternal age, and child birth order. In addition to these three dimensions, we will likely examine heterogeneity with respect to the saturation (the share of villages treated) in the sub-location where the child was born. This dimension of heterogeneity, which is exploratory, is designed to test whether spillovers are distinct for those receiving cash in an area that is receiving different amounts of cash transfers overall. This would be relevant if, as one illustrative example, making many households in an area richer improved the quality of healthcare in local clinics, but health care costs prevented all agents from equally accessing the improved services. Of course, there are many other ways in which local cash transfer treatment saturation could affect household outcomes in health or other dimensions.

6 References

Abadie, A., J. Gu, and S. Shen (2023). Instrumental variable estimation with first-stage heterogeneity. *Journal of Econometrics*, 105425.

Abdul Latif Jameel Poverty Action Lab (J-PAL). 2020. "Using cash transfers to improve child health in low- and middle-income countries." J-PAL Policy Insights. Last modified May 2020. https://doi.org/10.31485/pi.2523.2020

Amek, N.O., Van Eijk, A., Lindblade, K.A. et al. Infant and child mortality in relation to malaria transmission in KEMRI/CDC HDSS, Western Kenya: validation of verbal autopsy. Malar J 17, 37 (2018). https://doi.org/10.1186/s12936-018-2184-x

Anderson, M.L., 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool and Early Training Projects", Journal of the American Statistical Association, December 2008, Vol. 103, No. 484: 1481-1495.

Byass, P., Herbst, K., Fottrell, E., Ali, M. M., Odhiambo, F., Amek, N., Hamel, M. J., Laserson, K. F., Kahn, K., Kabudula, C., Mee, P., Bird, J., Jakob, R., Sankoh, O., & Tollman, S. M. (2015). Comparing verbal autopsy cause of death findings as determined by physician coding and probabilistic modeling: a public health analysis of 54 000 deaths in Africa and Asia. Journal of global health, 5(1), 010402. https://doi.org/10.7189/jogh.05.010402

Benjamini, Y., A. M. Krieger, D. Yekutieli, 2006. "Adaptive Linear Step-up Procedures That Control the False Discovery Rate", Biometrika, Volume 93, Issue 3, September 2006, Pages 491–507, https://doi.org/10.1093/biomet/93.3.491

Campbell, O. M. R., Cegolon, L., Macleod, D., & Benova, L. (2016). Length of stay after childbirth in 92 countries and associated factors in 30 low- and middle-income countries: Compilation of reported data and a cross-sectional analysis from nationally representative surveys. *PLOS Medicine*, 13(3), e1001972. https://doi.org/10.1371/journal.pmed.1001972

Chernozhukov, V., M. Demirer, E. Duflo and I. Fernández-Val, 2018. "Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India", NBER working paper.

Conley, T. G., Gonçalves, S., Kim, M. S., & Perron, B. (2023). Bootstrap inference under cross-sectional dependence. *Quantitative Economics*, *14*(2), 511–569. https://doi.org/10.3982/QE1626

Egger, D., J. Haushofer, E. Miguel, P. Niehaus and M. Walker, 2022. "General equilibrium effects of cash transfers: Experimental evidence from Kenya", *Econometrica* 90(6): 2603-2643. https://doi.org/10.3982/ECTA17945

Egger, D., J. Haushofer, E. Miguel, and M. Walker, 2021a. "GE Effects of Cash Transfers: Pre-analysis plan for Endline 2 Household", July 2021. https://www.socialscienceregistry.org/trials/505

Egger, D., J. Haushofer, E. Miguel, and M. Walker, 2021b. "GE Effects of Cash Transfers: Pre-analysis plan for Endline 2 Local Economy Analyses", December 2021. AEA Trial Registry: https://www.socialscienceregistry.org/trials/505

Egger, D., J. Haushofer, E. Miguel, and M. Walker, 2023. "GE Effects of Cash Transfers: Pre-analysis plan for Endline 3 Child Mortality Analysis", May 2023. AEA Trial Registry: https://www.socialscienceregistry.org/trials/505

Gacheri, S., Kipruto, H., Amukoye, E., Ong, J., Mitchell, E. M. H., Sitienei, J., Kiplimo, R., & Muturi, C. (2014). Performance of clinicians in identifying tuberculosis as cause of death using verbal autopsy questionnaires in Siaya County, Kenya. African Journal of Health Sciences, 27, 232-238.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277. https://doi.org/https://doi.org/10.1016/j.jeconom.2021.03.014

Haushofer, J., E. Miguel, P. Niehaus and M. Walker. 2017a. "General Equilibrium Effects of Cash Transfers: Pre-analysis Plan for household welfare analysis", July 2017. AEA Trial Registry: https://www.socialscienceregistry.org/trials/505

Haushofer, J., E. Miguel, P. Niehaus and M. Walker. 2017b. "General Equilibrium Effects of Cash Transfers: Pre-analysis Plan for Targeting Analysis", July 2017. AEA Trial Registry: https://www.socialscienceregistry.org/trials/505

Jiwani, S. S., Amouzou-Aguirre, A., Carvajal, L., Chou, D., Keita, Y., Moran, A. C., Requejo, J., Yaya, S., Vaz, L. M., & Boerma, T. (2020). Timing and number of antenatal care contacts in low and middle-income countries: Analysis in the Countdown to 2030 priority countries. *Journal of Global Health*, 10(1), 010502. https://doi.org/10.7189/jogh.10.010502. PMCID: PMC7101027.

Lee, D. S. 2009."Training, Wages and Sample Selection: Estimating Sharp Bounds on Treatment Effects", The Review of Economic Studies 76 (3): pp. 1071-1102.

Mchenga, M., Burger, R., & von Fintel, D. (2019). Examining the impact of WHO's Focused Antenatal Care policy on early access, underutilisation and quality of antenatal care services in Malawi: A retrospective study. *BMC Health Services Research*, 19(1), 295. https://doi.org/10.1186/s12913-019-4130-1

Murray, C.J., Lopez, A.D., Black, R. et al. Population Health Metrics Research Consortium gold standard verbal autopsy validation study: design, implementation, and development of analysis datasets. Popul Health Metrics 9, 27 (2011). https://doi.org/10.1186/1478-7954-9-27

Nyaguara O. Amek, Frank O. Odhiambo, Sammy Khagayi, Hellen Moige, Gordon Orwa, Mary J. Hamel, Annemieke Van Eijk, John Vulule, Laurence Slutsker & Kayla F. Laserson (2014) Childhood cause-specific mortality in rural Western Kenya: application of the InterVA-4 model, Global Health Action, 7:1, DOI: 10.3402/gha.v7.25581

Odhiambo FO, Laserson KF, Sewe M, Hamel MJ, Feikin DR, Adazu K, Ogwang S, Obor D, Amek N, Bayoh N, Ombok M, Lindblade K, Desai M, ter Kuile F, Phillips-Howard P, van Eijk AM, Rosen D, Hightower A, Ofware P, Muttai H, Nahlen B, DeCock K, Slutsker L, Breiman RF, Vulule JM. Profile: the KEMRI/CDC Health and Demographic Surveillance System--Western Kenya. Int J Epidemiol. 2012 Aug;41(4):977-87. doi: 10.1093/ije/dys108. PMID: 22933646.

Okeke, E. N. (2023). When a doctor falls from the sky: The impact of easing doctor supply constraints on mortality. *American Economic Review*, 113(3), 585-627. https://doi.org/10.1257/aer.20210701

Romano, J. P., and M. Wolf. 2005a. Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing. *Journal of the American Statistical Association* 100(469): 94–108.

———. 2005b. Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica* 73(4): 1237–1282.

———. 2016. Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics and Probability Letters* 113: 38–40.

Semenova, Vira. *Generalized Lee Bounds* (2020). https://arxiv.org/abs/2008.12720

Serina, P., Riley, I., Stewart, A. et al. Improving performance of the Tariff Method for assigning causes of death to verbal autopsies. BMC Med 13, 291 (2015). https://doi.org/10.1186/s12916-015-0527-9

Walker, M. (2017). "Pre-Analysis Plan: Local Public Finance and Unconditional Cash Transfers in Kenya." February 2017. AEA Trial Registry: https://www.socialscienceregistry.org/trials/505

World Health Organization. (2016). WHO recommendations on antenatal care for a positive pregnancy experience. World Health Organization. https://www.who.int/publications/i/item/9789241549912

World Health Organization. (2022). WHO recommendations on maternal and newborn care for a positive postnatal experience. World Health Organization. https://www.who.int/publications/i/item/9789240045989