Short-term impacts of a livestock transfer and training program in rural Nepal

Sarah A. Janzen
Montana State University
sarah.janzen@montana.edu

Nicholas Magnan
University of Georgia
nmagnan@uga.edu

William M. Thompson
University of Georgia
willt78@uga.edu

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

Short-term impacts of a livestock transfer and training program in rural Nepal

Sarah Janzen[1], Nicholas Magnan[2], and William M. Thompson[2]

[1]Montana State University

[2]University of Georgia

May 24, 2017

Abstract

This study evaluates the short term impacts of Heifer International's (HI) livestock transfer and training program in Nepal using a randomized control trial (RCT). The RCT assigned three variations of a HI program in order to capture differential effects of various program components. We also evaluate a unique "pay it forward" program rule where recipients are encouraged to share what they have learned and even share some of their newly accumulated wealth, in the form of a productive asset, to other households in need. After 1.5 years, we find financial inclusion increases by 0.2 standard deviations, and empowerment increases by 0.3 standard deviations among direct beneficiaries. We observe "pay it forward" effects to indirect beneficiaries in financial inclusion and empowerment that are of approximately the same magnitude as direct beneficiaries. These observed short-term impacts are similar across the different program variations.

1 Introduction

It is often argued that the rural poor largely lack access to the productive assets and human capital necessary to be successful entrepreneurs. Productive asset transfer programs, which typically include a training component, are one way non-governmental organizations (NGOs) and governments try to alleviate these constraints thereby facilitating a permanent transition out of poverty. These programs are popular among donors who adhere to the well-known "teach a man to fish" mantra. In some cases, they also include a "pay it forward" component where recipients share what they have learned and even share some of their newly accumulated wealth, in the form of a productive asset, to other households in need. Rigorous impact evaluations of combined asset transfer and training programs, particularly evaluations designed to measure impacts on indirect beneficiaries, are few and far between.[1]

In this paper we evaluate the short-term (1.5 year) welfare impacts of Heifer International's (HI) Smallholders in Livestock Value Chain Program in rural Nepal using a randomized controlled trial (RCT). Like similar programs, the program targets poor households in rural areas, particularly women, and seeks to provide a sustainable livelihood and a pathway out of poverty for its beneficiaries. The standard intervention in Nepal provides a package of benefits that includes group formation, livestock (in this case two female goats), technical trainings on improved animal management and entrepreneurship, and values-based trainings. The values-based training encourages beneficiaries to "pay it forward" by providing technical training and giving the first-born female offspring of their received livestock to another poor individual in their community.

This paper contributes to the literature in three important ways. First, we add to a small but growing body of empirical evidence on the *overall* positive impact of livestock transfer programs worldwide (Bandiera et al., 2013; Banerjee et al., 2015; Darrouzet-Nardi et al., 2016; Jodlowski et al., 2016; Kafle, Winter-Nelson, and Goldsmith, 2016; Miller et al.,

[1]Banerjee et al. (2015) test for spillover effects of a program without a "pay it forward" component and find none.

2014; Rawlins et al., 2014). Second, our evaluation is carefully designed to estimate the aforementioned "pay it forward" indirect effects on members of the same targeted community who were not initially targeted by the implementing partner. Measuring the strength and persistence of this element of the program design is crucial to understanding the overall program impacts. Third, our evaluation includes three unique treatments in order to unpack the welfare impacts of different program components. In the first treatment arm, beneficiaries received a complete package that included a livestock transfer, skills-based technical trainings and values-based non-technical trainings. In the second treatment arm, beneficiaries received skills-based technical trainings and values-based non-technical trainings, but *not* livestock. In the third treatment arm beneficiaries received a livestock transfer and skills-based technical trainings, but *not* values-based non-technical trainings. To our knowledge, previous studies in this area do not attempt to disaggregate the impacts of a bundled treatment.

Our total baseline sample is 3,283 women across 60 treatment clusters (village development committees, or VDCs) stratified by region and ethnic composition. Between July and December 2014, HI administered training and delivered goats to randomly selected targeted beneficiaries. Various additional trainings continued throughout 2015. Shortly after HI delivered training and livestock to the original beneficiaries of the project, a devastating earthquake struck Nepal. The earthquake adversely affected ten villages originally included in the evaluation, spread evenly across treatment groups and control. These ten VDCs were dropped from the RCT so that HI could provide disaster relief as deemed appropriate. The remaining sample size comprises 50 villages and 2,724 individuals from which follow-up data was collected in June-July 2016. Although a definite concern, updated power calculations suggest the study remains sufficiently powered to capture overall treatment effects.

Our hypotheses, along with detailed plans for handling the data and analysis, are documented in a registered pre-analysis plan available at `http://www.socialscienceregistry.org/trials/1504`. Because we have a rich dataset informing numerous hypotheses regarding behavioral change and improved welfare across several dimensions, we employ summary in-

dices to capture broad program impacts and reduce the number of hypotheses tested. These include income, asset holdings, expenditures, financial inclusion, physical health, mental health, aspirations, and women's empowerment. Although we greatly reduce the number of hypotheses tested by using these indices, we still account for multiple hypotheses testing by controlling for the false discovery rate using the Benjamini and Hochberg (1995) step-up method.

We estimate intent-to-treat (ITT) effects of the overall program on directly targeted beneficiaries as well as beneficiaries brought into the program through the "pay it forward" process, whom we term indirect beneficiaries. In the short-term, we find the intervention causes a statistically significant 0.31 standard deviation increase in our index of financial inclusion, and a 0.24 standard deviation increase in empowerment among direct beneficiaries. Perhaps surprisingly, we do not observe statistically significant differences in outcome indices across treatments for direct beneficiaries, suggesting we are either underpowered to capture small differences between treatments, or the combination of activities is not necessary to achieve the desired impact in the short run (though perhaps in the long run it is).

We also observe substantial indirect effects. Those who live in the same community as direct beneficiaries but who were not targeted as direct beneficiaries themselves experience similar and statistically significant increases in financial inclusion (+0.21 standard deviations) and empowerment (+0.29 standard deviations) as those observed for directly targeted beneficiaries. These results are impressive given the relatively short time horizon over which to observe an impact on indirect beneficiaries. Notably, this indirect effect vanishes when the values-based trainings are withheld, suggesting the "pay it forward" encouragement (a critical component of the values-based training) helps successfully achieve a broader impact.

This paper focuses on short run impacts. We do not observe statistically significant changes in income, asset holdings, or expenditures in the short run. However, taking into account goat gestational periods and kid growth rates, livestock sales of transferred goats' kids are largely unanticipated within the timeframe of this study. In a program that targets

livestock production as an income-generating activity, economic impacts within this short time horizon would be unanticipated. Future work will analyze the long-term impacts of the program on these economic outcomes.

2 Background

2.1 Evidence on asset transfer and training programs

Asset transfers, particularly livestock, have been conducted in poor countries since at least 1944, when HI sent 17 cows from Arkansas to Puerto Rico. Since then HI has expanded its reach to over 125 countries. Numerous NGOs and even governments have embraced livestock transfer and training programs as a strategy for fighting poverty (World Vision, BRAC, Save the Children, Oxfam, and the Government of Rwanda are a few examples).

Despite the long history and prevalence of livestock transfer and training programs, until recently there was very little rigorous empirical evidence of their effectiveness (DFID 2014). Recent papers have found that these programs increase income (Bandiera et al., 2013; Banerjee et al., 2015), expenditures (Bandiera et al., 2013; Banerjee et al., 2015; Jodlowski et al., 2016), savings (Bandiera et al., 2013; Banerjee et al., 2015), overall food consumption (Bandiera et al., 2013; Banerjee et al., 2015; Kafle, Winter-Nelson, and Goldsmith, 2016), dairy and meat consumption (Banerjee et al., 2015; Rawlins et al., 2014), dietary diversity (Darrouzet-Nardi et al., 2016; Jodlowski et al., 2016; Kafle, Winter-Nelson, and Goldsmith, 2016; Rawlins et al., 2014), food security (Bandiera et al., 2013; Banerjee et al., 2015), and anthropometrics (Miller et al., 2014; Rawlins et al., 2014). Evidence of impacts on emotional well-being and women's empowerment have been mixed (Bandiera et al., 2013; Banerjee et al., 2015; Roy et al., 2015).

Most notably, Banerjee et al. (2015) evaluate the impact of BRAC's graduation program, a large asset transfer and training program in six countries. Their study finds after three years the graduation program has significantly positive impacts on consumption, food

security, assets, finance, time use, income and mental health, but no indirect effects for non-beneficiaries. They observe positive impacts on women's empowerment in the short but not long run.

Several major differences exist between the graduation program and the one evaluated here. First, beneficiaries of the graduation program chose an asset (or bundle of assets) from a list of productive assets. Although livestock was the most common choice, there were alternative options. The value of the productive asset transfer was always higher than the one evaluated here, and beneficiaries in their study also received a regular transfer of food or cash for a few months or even up to a year. In another significant deviation, beneficiaries of the livestock transfer and training program we study here were encouraged to "pay it forward," as described below. This encouragement is a central component of all livestock transfer and training programs implemented globally by Heifer International (HI). To our knowledge, no study has evaluated the impact of a program with this type of encouragement.

2.2 Heifer International's "Pay-it-forward" livestock transfer and training program in Nepal

The intervention we evaluate replicates HI's Smallholders in Livestock Value Chain (SLVC) Program in rural Nepal. Like similar programs, the program targets poor households in rural areas, and seeks to provide a sustainable livelihood and a pathway out of poverty for its primarily women beneficiaries The standard HI intervention in Nepal provides of a package of benefits that includes formation of women's self help and savings groups, technical trainings on improved animal management and entrepreneurship, values-based trainings, a productive asset transfer (in this case goats), and encouragement to "pay it forward".

The process is as follows: After identifying a location to receive the intervention, HI recruits an original group of direct beneficiaries. Direct beneficiary groups typically consist of close neighbors and often include most or all of the households in a given neighborhood. As a rule, HI considers all the households in a targeted area to be objectively poor and there-

fore eligible for the program, allowing for the possibility that a considerable range of relative wealth and poverty might exist within a group. Once selected, direct beneficiaries within a ward are organized into a self-help group (SHG). Over a period of months all SHG members participate in a series of trainings. Trainings include (1) technical training on improved animal management, fodder/forage development, entrepreneurship, human and animal nutrition, and home gardening, and (2) HI's values-based training on topics of accountability, sharing and caring, sustainability, self-reliance, income management, environmental stewardship, spirituality, self-help group management, gender justice, and encouragement to pay it forward.. The trainings culminate with the beneficiaries receiving a transfer of livestock which includes two doe goats for each beneficiary and a single buck of improved stock (to facilitate a breeding program) for the SHG.

A unique component of HI's model is that it encourages members to "pay it forward" by recruiting additional community members into the program, giving a gift of livestock (of equal value to what was received), and passing down all knowledge that was gained through participation in the programs. HI facilitates values-based empowerment training for both direct and indirect beneficiaries (albeit separately and at different points in time), while all other "pay it forward" trainings are implemented by direct beneficiaries with minimal support from HI. In this way, what might typically be deemed a spillover effect is actually an important program component. The program we evaluate follows an innovation to the basic HI pay-it-forward model, in which each direct beneficiary SHG is tasked with recruiting up to five indirect SHGs, with the goal of full saturation and complete adoption of improved practices and technologies within a community in a relatively short time frame.

3 Experimental design

To establish a causal relationship between the program and changes in outcomes, this study uses a cluster randomized controlled trial (RCT). A cluster design was employed for two reasons. First, group membership is a key component of the program design. Second,

indirect effects are anticipated. As described below, we will seek to estimate both direct and indirect effects.

Nepal comprises 75 districts. Districts are further subdivided into village development committees (VDCs), which can be thought of as clusters or groupings of villages within a district. Every VDC is split into nine wards, and each ward might include multiple *toles*, or communities. A typical *tole* in the study area has approximately twenty to thirty households; a typical ward has roughly 150 households.

Nepal-based HI staff first identified 60 VDCs in which they had never worked, but that would be good candidates for an asset transfer and training program. Before assigning treatments, HI also identified a central ward and targeted *tole* within the selected central ward for each of the 60 selected VDCs. The expectation was that if assigned to treatment, everyone residing in the targeted *tole* would be targeted by the program, and therefore likely to enroll as a direct benificiary. Through this process, HI pre-identified all targeted beneficiaries (but not necessarily actual beneficiaries) who were later encouraged to form SHGs. Following treatment assignment, these SHGs formed in treated VDCs but not control VDCs. In this way, the individuals in the control arm are directly comparable with those in the treatment arms.

Although indirect effects are expected, we do not anticipate contamination. To an extent, the isolation of rural communities in Nepal provides a natural impediment to such contamination. This is especially true in the Middle Hills (home to about two-thirds of our sample), where lower population density, rugged terrain, poor roads, and inferior cellular connectivity cause communities to be especially cut off. Nevertheless, communities are linked by family and commercial ties. Fewer natural barriers against contamination exist in the Terai, the densely populated plain along the Indian border where about one third of our sample resides. Apart from naturally occurring geographic and social barriers to contamination, we also buffered treated wards from each other and from control VDCs by selecting the 'central' ward within a VDC to be the targeted ward. In this way, we ensure an additional

degree of isolation and further reduce the prospect of unintentional spillovers that could bias results.

To improve balance across treatment and control VDCs (and between the various treatment VDCs) we stratified by geography and caste/ethnic composition. First we divided the sample of VDCs into four pools based on district groupings (Hills (2), Middle Hills (1), and Terai (1)). These clusters contained 15, 15, 10, and 20 VDCs respectively. Using administrative data, we then calculated the proportion of residents in each VDC from each of 39 caste/ethnic groups. Within each district grouping we ordered VDCs by the most prevalent caste/ethnic group, then second most prevalent caste/ethnic group, and so on through the ninth most prevalent caste/ethnic group.[2] This created new groups within the district groupings based on rank prevalence of caste/ethnicity. Within these groups, we ordered VDCs by the proportion of the most prevalent caste/ethnicity, then second most prevalent, and so on. From this ordering we established 16 bins.

Within each stratification bin, we then randomly assigned the 60 VDCs to one of three treatment arms or pure control.[3] All three treatments share some common features. First, HI facilitates the formation of women's self-help groups, so all beneficiaries are expected to acquire some level of social capital by belonging to a group. Group members are encouraged to contribute to group savings accounts with a goal toward increasing financial inclusion. Finally, all beneficiaries are trained on a variety of technical topics including nutrition, home gardening, fodder and forage development, and improved animal management. In addition, all beneficiaries are provided a small amount of cash support for home garden and fodder/forage production. We'll call these common features the basic intervention.

In order to 'unpack' the benefits of various program components, two additional programmatic elements vary across across treatment arms: a productive asset transfer and

[2]Only two of 60 VDCs had more than 9 caste/ethnic groups represented.

[3]Because of the unequal number of VDCs in each district grouping there was one bin with two VDCs, two with three VDCs, and 13 with four VDCs. Because of the uneven bin sizes, random treatment assignment within bins resulted in 16 VDCs in treatment 1, 16 VDCs in treatment 2, 15 VDCs in treatment 3, and 13 VDCs in the control group. To obtain equally sized treatment arms we randomly drew one VDC each from treatments 1 and 2 to be placed in the control group.

additional values-based trainings. The productive asset transfer included two doe goats and cash support for goat shed improvement to each individual beneficiary, as well as a shared buck of improved breeding stock for the self-help group. The values-based trainings cover the 12 HI Cornerstones which include: passing the gift; accountability; sharing and caring; sustainability and self-reliance; improved animal management; nutrition and income; gender and focus on the family; genuine need and justice; improving the environment; full participation; training, education, and communication; and spirituality. Notably, the values-based training encourages beneficiaries to "pay it forward" by providing technical training and giving the first-born female offspring of their received livestock to another poor individual in their community.

The treatment arms can be described as follows:

1. *Full Treatment* (FT): basic intervention, values-based training, and a productive asset transfer.
2. *No Goats* (NG): Identical to FT, but without the productive asset transfer.
3. *No Values-based Training* (NVT): Identical to FT, but without values-based training.

A fourth arm was randomly selected as pure control. Figure 7 summarizes the elements of each treatment arm.

4 Data

We collected panel household survey data from nearly 3,300 rural women eligible to participate in an asset transfer program across three regions of Nepal in June-September 2014 and 2016. Project implementation began in mid to late 2014 and continued throughout 2015. The main data used for this analysis was collected in June-July 2016, approximately 1.5 years after initial enrollment in the program.

There are two types of respondents in the sample used for this analysis: targeted direct beneficiaries and prospective indirect beneficiaries. Specifically, our sample of targeted direct beneficiaries consists of all households in each of the targeted *toles* (around 25 per ward). In addition, after removing households from the targeted *tole*, we selected a random sample of 15 potential indirect beneficiaries from a complete roster of all households in the central ward. Because of the aggressive nature of the "pay-it-forward" encouragement, we expect that many (if not most) of these households will actually become PIF beneficiaries. Although no intervention took place in control VDCs, sampling in these VDCS occurred in exactly the same manner as in treatment VDCs: 25 individuals from pre-determined targeted *toles*, and 15 individuals from a complete roster of all households in the central ward.

Our total baseline sample is 2,375 women, including 1286 targeted for direct treatment, and 1089 households from the central ward likely to receive indirect treatment through the "pay-it-forward" mechanism. Shortly after HI delivered training and livestock to the original beneficiaries of the project, a devastating earthquake struck Nepal. The earthquake greatly affected the 10 VDCs belonging to the 'Middle Hills' stratification pool, and were therefore spread evenly across treatment groups and control. We made the decision to drop these from the RCT so that HI could provide earthquake relief in whatever manner they deemed appropriate. The remaining sample consists of 50 VDCs and 1,829 households, including 1,031 from targeted *toles* and 798 from the central ward more broadly.

The data includes information on basic household demographics and a variety of outcomes. We group these outcomes across 8 dimensions: asset ownership, income, non-food consumption, financial inclusion, physical health, mental health, aspirations, and women's empowerment.[4] Multiple indicators exist for each dimension. For the purpose of analysis, these subindicators are then aggregated into a primary summary index for each dimension of welfare. Details regarding how these outcomes are measured are discussed in section 1 of the online appendix.[5]

[4]The PAP outlines ten dimensions. However, the data on two of the ten dimensions, time use and food security, were deemed unusable for analysis.

[5]This paper considers dozens of outcomes and includes scores of tables containing hundreds of individual

As is appropriate for an RCT, we check for balance across the various treatments and control. The results of the balance checks are reported in Section 3. We do observe some imbalance at baseline. Where imbalance exists for a baseline level of an outcome or a demographic control with respect to any comparisons of interest, we include the baseline level of the imbalanced variables in the econometric specifications described in sections 5.1.1 and 5.1.2.

Ignoring the sample puposefully removed following the earthquake, we observed XXX% attrition between 2014 and 2016. To assess whether the observed attrition is systematic in a way that might bias our results, we employ the same three approaches outlined in the pre-analysis plan of Haushofer and Shapiro (2013), and adapted to our design. For ease of notation, the approach presented below considers the full sample, but we also assess attrition separately for direct and indirect households.

First, equation 1 estimates whether attrition rates differ across treatment types and control households, where $attrit_{hv}$ is a binary variable indicating that a household was surveyed at baseline but is missing from the endline data set.

$$attrit_{hv} = \beta_0 + \beta_1 T1_{hv} + \beta_1 T2_{hv} + \beta_1 T3_{hv} + \varepsilon_{hv} \tag{1}$$

After estimating equation 1, we do not observe any significant treatment effects on attrition status. Next, we assess whether attrition rates differ across households with respect to a set of baseline characteristics. To do this we regress a variety of baseline outcomes on attrition status as estimated in equation 2:

$$y_{hv} = \beta_0 + \beta_1 attrit_{hv} + \varepsilon_{hv} \tag{2}$$

We estimate equation 2 for each of the indices and sub-index outcomes as well as a set of

regressions. In order to keep this paper compact and focused we leave much of the detail to an extensive online appendix: https://wm-thompson.com/basishi-results-online-appendices.

demographic variables, and correct for false discovery rate. While we do find scattered individual cases where attrition status correlates with a baseline characteristic, these instances do not appear to be systematic or to threaten the integrity of our results. Finally, equation 3 estimates the extent to which baseline characteristics of treated households differ from control households, after restricting the sample to attrited households:

$$(y_{hvB}|attrit_{hv} = 1) = \beta_0 + \beta_1 T_{hv} + \varepsilon_{hv} \tag{3}$$

The results of all attrition tests are available in online appendix section 4. In general, we find no noteworthy attrition effects from specification 3.

5 Estimation

Our main questions are: (i) across a range of dimensions of household and individual welfare, what are the overall impacts of a productive asset transfer that simultaneously develops human, physical, and social capital, (ii) what are the specific impacts of each aspect of the intervention, (iii) within a treated village, do treatment effects spillover to subsequent generations of beneficiaries, and (iv) which package of benefits results in the most cost-effective improvements to household and individual well-being? Our specific hypotheses, along with detailed plans for handling the data and analysis, are documented in a pre-analysis plan prepared and registered (`http://www.socialscienceregistry.org/trials/1504`) before any analysis took place. In some instances we deviate from this plan, and will specify when this is the case.[6]

[6]Notably, the PAP also proposes to analyze heterogeneous treatment effects; we leave this to future work.

5.1 Estimation of treatment effects

5.1.1 Estimation of direct effects–treatment wards

First, we examine program effects for intended direct beneficiaries. We use two specifications to estimate these impacts. The first specification follows the PAP and estimates the intent to treat (ITT) impact for each of the three treatment groups relative to a common control:

$$y_{hvm} = \beta_0 + \beta_1 T1_{hv}^{DIR} + \beta_2 T2_{hv}^{DIR} + \beta_3 T3_{hv}^{DIR} + \delta y_{hvb} + \mathbf{X_{hvb}}\gamma + \mathbf{S_{vb}}\rho + \varepsilon_{hv} \qquad (4)$$

where y_{hvm} is the outcome of interest for household h in village v, measured at midline ($t = m$). Treatment indicator variables ($T1_{hv}^{DIR}$, $T2_{hv}^{DIR}$, and $T3_{hv}^{DIR}$) take a value of 1 for directly targeted households in wards selected to receive any of the previously described treatments ("treated households"), and a value of 0 otherwise. The omitted category is control intended-direct households located in pure control villages ("control households"). In order to improve statistical power, we conditioned on baseline ($t = b$) levels of the outcome of interest y_{hvb}, a vector of control variables X_{hvb} for which imbalance at baseline was observed across treatments, and a vector S_{vb} of stratification bin dummies. Finally, ε_{hv} is an idiosyncratic error term. We cluster errors at the VDC level, as this is the level of treatment.

The treatment effects of interest are β_1, β_2, and β_3. β_1 represents the "intent to treat" (ITT) treatment effect on households selected to directly receive the full treatment package (T1, FT) when compared to intended-direct households in pure control VDCs, β_2 identifies the ITT treatment effect on households selected to directly receive the no-goats package (T2, NG), and β_3 identifies the ITT treatment effect on households selected to directly receive the no-values-based-training treatment package (T3, NVT). An important aspect of our evaluation is to test whether the treatments effects vary across treatment type. Therefore, we conduct Wald tests for $\beta_1 = \beta_2$, $\beta_1 = \beta_3$, and $\beta_2 = \beta_3$.

The second specification, which was not included in the PAP, aggregates the three treatment arms and compares the aggregated treatment to a common control:

$$y_{hvm} = \beta_0 + \beta_1 T_{hv}^{DIR} + \delta y_{hvb} + \mathbf{X_{hvb}}\gamma + \mathbf{S_{vb}}\rho + \varepsilon_{hv} \tag{5}$$

We do this to increase power, noting that for the most part there are not substantial differences in treatment effects across groups. Although initial power calculations were conducted to determine an appropriate sample size, a major earthquake struck Nepal one year after baseline, and one year before the midline data was collected. This earthquake had devastating consequences to roughly 20% of our sample. Rather than preserve the experiment, we opted to drop households severely affected by the earthquake from our analysis in order to allow Heifer International to intervene as they saw fit. This certainly affects the precision of our results. A second issue is lower-than-anticipated takeup. Recruitment across treatments of the intended direct beneficiaries varied from 62% for the NVT treatment, to 68% for the FT package. This suggests that ITT effects may be very conservative. We take up a more detailed discussion of the implications of this lower than expected compliance in section 6.3.2.

5.1.2 Estimation of indirect effects–treatment wards

Under HI's pay-it-forward model, HI expects some fraction of potential indirect households within the same ward to receive similar, or even identical, treatment as the direct members spread benefits by training their peers and giving them livestock. For the SLVC program, and the parallel evaluation, HI implemented a novel exponential recruitment strategy, where each group of direct SHG is tasked with forming five additional indirect groups. The goal of this strategy is for HI's impact to occur more broadly and rapidly (more households reached in a short period of time) than under the traditional HI (non-exponential) model. However, the exponential model may also be a less intense treatment because households receive benefits over a longer period of time (livestock get passed down much more slowly). Given the time

between when direct households received benefits and when midline data was collected, it is possible that very few of the potential indirect households will have received any livestock.

Even if potential indirect households had not yet received any benefits, spillover effects may have occured through a second channel: households may simply observe or discuss techniques or other concepts learned through the trainings. If households replicate these techniques, they may benefit indirectly from the HI trainings, even if they receive no goats or formal training. Because this second type of spillover effect is possible, estimation of local average treatment effects (LATE) is not preferable. We therefore estimate ITT effects, keeping in mind that they may be very conservative, especially using outcomes at midline. We also calculate and report the proportion of potential indirect households actually receiving benefits by midline data collection. The regression model for indirect treatment effects within treatment wards compares potential indirect treatment households to the corresponding indirect control households (and excludes all direct and spillover households). The regression specification is:

$$y_{hvm} = \beta_4 + \beta_5 T1_{hv}^{PIF} + \beta_6 T2_{hv}^{PIF} + \beta_7 T3_{hv}^{PIF} + \delta y_{hvb} + \mathbf{X_{hvb}}\gamma + \mathbf{S_{vb}}\rho + \varepsilon_{hv} \tag{6}$$

Here again we will condition on baseline levels of the outcome of interest y_{hvb}, a vector of control variables X_{hvb} for which an imbalance at baseline was observed across treatments, and vector S_{vb} of stratification bin dummies. We cluster standard errors at the VDC level. In addition to equation 6, we also estimate the indirect-beneficiary analog to equation 5.

In this specification, β_5 corresponds to the T1 ITT effect of being a potential indirect beneficiary in a T1 treatment ward, β_6 captures the T2 ITT effect of being a potential indirect beneficiary in a T2 treatment ward, and β_7 captures the T3 ITT effect of being a potential indirect beneficiary in a T3 treatment ward. We will test for whether the treatments have different indirect effects within wards using Wald tests for $\beta_5 = \beta_6$, $\beta_5 = \beta_7$, and $\beta_6 = \beta_7$.

5.2 Accounting for multiple inference

We have a rich dataset informing numerous hypotheses regarding behavioral change and improved welfare across several dimensions. Therefore, we follow the emerging standard in the program evaluation literature by accounting for multiple hypotheses in two ways. First, we construct one primary summary index for each dimension of welfare described in section **??**. Each summary index consolidates several individual tests into a single test. Second, because we still have multiple outcome dimensions, we report naive p-values and adjusted q-values that control for the false discovery rate (FDR). Specifically, we calculate q-values for multiple hypothesis tests across summary indices, but not across treatments, using the Benjamini and Hochberg (1995) step-up method outlined in Anderson (2008) and applied by Banerjee et al. (2015). To test treatment groups against each other we will conduct Wald tests as described in section 5. For these tests, we will report both naive p-values and q-values that control for FDR. As above, we calculate q-values for a specific hypothesis test across summary indices (and interaction terms when applicable), using the Benjamini and Hochberg (1995) step-up method. When estimating treatment effects on sub-indicators (rather than summary indices) we will report naive p-values. We test for the impact on sub-indicators primarily to identify the mechanism behind impact (or lack thereof) observed for the summary indices. We therefore consider this analysis exploratory, and take a less stringent approach to hypothesis testing.

We prefer controlling for FDR over controlling for the the family-wise error rate (FWER) because we are testing a large number of hypotheses (even after condensing them to summary indices), and FWER adjustments become increasingly severe as the number of tests grow (Anderson, 2008; Benjamini and Hochberg, 1995). Overall conclusions about the SLVC program effectiveness depend on many outcomes– the overall conclusion should not be that the intervention is ineffective because of one erroneously rejected null hypothesis– so it seems reasonable to be more tolerant of Type I error in exchange for greater power. The FDR formalizes this tradeoff between Type I and Type II error (see Benjamini and Hochberg

(1995) and Anderson (2008) for a more detailed discussion).

6 Results

6.1 Effects on Outcome Indices

Our priors with respect to the signs and magnitudes of the treatment effects are based on a combination of Heifer's theory of change, intuition, and economic theory. With few exceptions variables have been coded so that larger values are 'better', therefore postive regression coefficients represent improvements. We advance a general hypothesis that treatment improves outcomes, but stipulate that we do not expect not all outcomes to improve at the same rate, especially when subjected to different configurations of treatment. Welfare outcomes like assets, income, and physical health are more distal to treatment and require more time to come into statistical focus than outcomes like financial inclusion and empowerment, where welfare improvements are more proximate to treatment. Put another way, outcomes like financial inclusion and empowerment lay the groundwork for the ultimate intended program outcomes of improved income, assets holdings, and physical health.

Table 2 contains direct ITT effects on on eight of ten summary indices specified in our PAP. Column 1 contains control group means, columns 3-5 contain ITT estimates for the three disaggregated treatment groups (β_1, β_2, and β_3 in equation 4). Column 2 contains ITT estimates for the aggregated treatment groups (β_1 in equation 5).

We turn first to the average effect of being in any treatment group (column 2). We find significant impacts on financial inclusion (0.31 SD), and the Women's Empowerment in Agriculture Index (0.28 SD)[7]. The effects on both financial inclusion and empowerment are still significant to the 0.1 confidence level after controlling for the false discovery rate (FDR). When we look for disaggregated effects across treatment arms we see very little difference in the point estimates across treatments, and note that the standard errors are slightly larger

[7]following Alkire et al. (2012)

due to the smaller number of treatment units.

We do not observe changes in income, asset holdings or expenditures for beneficiaries of the program. In Section 6.3.1 we carefully consider livestock dynamics and note that the timing of livestock transfers has not yet allowed for goat sales. As such, it may be too early to observe income effects.

We do not observe statistically significant differences in outcomes across treatments. This suggests either that our analysis could not capture small differences between treatments or that the combination of activities is not critical for increased empowerment or financial inclusion in the short run. It is too early to say if the different program components are important for improving long-run economic outcomes.

Table 3 summarizes indirect IIT effects on summary indices. Here, the sample includes potential "pay it forward" beneficiaries who were not initially targeted by HI. As with the direct effects, we see significant increases in the financial inclusion index and the empowerment index. These results are impressive given the relatively short time horizon in which to observe indirect impacts; the vast majority if PIF beneficiaries had not yet received livestock. These results remain significant after controlling for FDR at the 0.1 confidence level. Notably, we do not observe these indirect impacts when the values-based trainings that encourage paying it forward are withheld.

These findings are consistent with a narrative that, a little more than one year after intervention, SLVC beneficiaries show improved levels of welfare outcomes that we might reasonably expect to respond to treatment in the short-term. Outcomes that we might classify as longer term demonstrate a null or noisy response. In the following subsection we carefully examine index components to develop a clearer picture of the driving forces behind the behavior of each index, and to get a sense of potential trajectories of long-term welfare outcomes.

6.2 Effects on Sub-indicator Outcomes

Tables summarizing intended-direct and intended-indirect treatment effects on outcome index components, which we term 'sub-indicators' outcomes, follow the tables of summary index treatment effects. As with the summary indices, column 1 contains control group means, column 2 contains ITT estimates for the aggregated treatment groups (β_1 in equation 5), columns 3-5 contain ITT estimates for the three disaggregated treatment groups (β_1, β_2, and β_3 in equation 4).

In this section we attempt to gain some insight into the behavior of the outcome indices by an exploratory analysis of the sub-index outcomes. Where significant treatment effects on indices exist, we look to the index components to isolate the driving forces behind the effect. Similarly, where a null effect exists or where large standard errors mute a potentially significant effect, we can examine components to glean a better understanding of the index behavior. As with direct effects on indices, our priors with respect to direct treatment effects on sub-index outcomes are informed by HI's expectations, intuition, and economic theory, and variables are typically coded so that positive regression coefficients represent desirable outcomes.

Assets

Although we see no significant effect on the assets summary index, we do observe noteworthy (yet in some cases statistically insignificant) dynamics with respect to individual assets as reported in tables 4 and 5. Among both direct and indirect beneficiaries livestock ownership increased by a statistically insignificant 0.208 tropical livestock units (TLUs) in the aggregated treatment. However, the point estimates for directly targeted beneficiaries under the FT and NVT treatments are 0.265 and 0.307 TLUs (respectively) and the point estimates for indirect beneficiaries under FT and NVT treatments are 0.440 and 0.200 TLUs. Among direct and indirect beneficiaries alike we observe point estimates near zero for the no-goat treatment effect on livestock holdings. A 0.2 increase in TLUs equals an addition of two goats. Therefore, these point estimates are broadly consistent with the magnitude of

the transfers received by FT and NSC beneficiaries, plus a reasonable herd growth rate.

Perhaps surprisingly, goat herds have grown at roughly the same rate among intended direct beneficiaries as they have among potential indirect ones. Given the size of the direct beneficiary group in each treatment ward (around 25 households), the total number of households within each treatment ward (roughly 100), and the gestational period of goats, we would expect no more than 25 percent of our sampled potential indirect beneficiary households to have received goats compared to 100 percent of our direct beneficiary households. Therefore we would expect our ITT estimates for direct beneficiaries to be four times as large as for potential indirect beneficiaries. We can think of a few reasons why this might be. First, not all women selected to be a direct beneficiary became one. Some ended up entering the Heifer program by receiving a "passed gift," perhaps because at first they were hesitant to join but after seeing the program at work decided to join a group (we will show some evidence of this in section 6.3.2). Likewise, some households from the potential indirect sample ended up joining the original group of beneficiaries, probably because of vacancies left by the aforementioned directly targeted women who decided to not join. As we stated earlier, the magnitude of herd size increases exceeds the two goats given by Heifer to original beneficiaries.

It is also possible that these households chose to invest in livestock after joining a self-help group. These households may have purchased livestock to leverage their new knowledge regarding animal husbandry and management, or to take advantage of their support for livestock shelters acquired as part of the program. However, we note that in the NG groups there was no increase in herd size, casting some doubt on this second explanation as women in this treatment group were also given technical training on animal husbandry and management and encouraged to build livestock shelters (indeed, they also show increased livestock expenditures, as we will show below).

We also see statistically significant increases in landholdings of about 0.05-0.1 hectares (compared to a control mean of 0.47 hectares) for direct beneficiaries belonging to the FT

and NVT treatment groups, and also for the combined direct treatment variable. We do not see an increase for the NG group. This may indicate that households are investing in land to grow fodder for their livestock; field staff report that this is the case. Field staff also report that households often purchase land with additional income generated from raising goats or other businesses. We note, however, that we do not observe increases in total income at this time.

Income

Tables 6 and 7 report sub-indicator outcomes for treatment effects on income. We find no significant increase in total income, and the point estimates are very small. We do, however, see some increases in different categories of income albeit without much consistency across treatments. Among directly targeted beneficiaries, we see a marked and significant decrease in crop income for the full treatment, and insignficant decreases of lesser magnitudes in the NG and NVT treatments. It's possible that beneficiaries may have removed some land from cash crop cultivation in favor of growing fodder for their goats.

While we see no overall increase in income for indirectly targeted households, we see large and highly significant increases in livestock income, a surprising result given the lack of a treatment effect on livestock income among direct beneficiaries. Control indirect households have incomes roughly 60 percent as large as directly targeted households in the control group, and these increases are such that in the treatment groups, direct and indirect households have similar levels of livestock income ex ante of intervention.

We also see a substantial gain in business income for the NG group for both direct and PIF beneficiaries that is not observed in the other treatment groups. At present we have no explanation for this finding.

Non-food Consumption

Table 8 reports sub-index outcomes for non-food consumption (expenditures) for direct beneficiaries. For the group receiving no goats, we note a negative and marginally signifi-

cant decrease in miscellaneous expenditures, which include some necessities (transport, rent, housing materials) and some that might be considered luxuries (festivals, ceremonies, jewelry, kitchen equipment). The effect is insignificant for the other two treatments or for households receiving any treatment, but has the same negative sign and only a slightly smaller magnitude.

We observe a similar effect for PIF beneficiaries in table 9. Just as we saw for directly targeted households, potential indirect beneficiary households in the NG treatment group spent significantly (and substantially) less on miscellaneous goods.

Financial Inclusion

Table 10 shows across the board large (0.31 SD) and significant increases in the financial inclusion index. Part of this increase is somewhat tautological, as being in a savings group is part of the Heifer program and also part of the index. However, we see a 75 percent increase in the amount households saved in the past month for households in any treatment. This effect is larger for those receiving the full treatment (74 percent) or no goats (106 percent) than it is for those receiving no values-based training (an insignificant 36 percent). Because saving ("income management") is stressed in the values-based training, it is logical that those receiving this training save more than those who do not.

While the results for borrowing are not significant, we note that the sign on the amount owed to formal lenders is positive and the sign on the amount owned to informal lenders is negative. We find a negative and significant decrease of 2.3 percent in the discount rate of beneficiaries receiving the full treatment. The sign on discount rate also is negative for the other treatments and for treatment households as a whole, but not significantly so.

Treatment effects presented in table 11 on the financial inclusion summary index for potential indirect households (0.21 SD) are significant and slightly smaller than they are for directly targeted households. Turning to the sub-indicators we see some important differences across treatments. First, while savings group membership went up substantially in all treatment groups for direct beneficiaries, increases in savings group membership rates are

confined to the FT and NG treatments for indirect beneficiaries. Second, potential indirect beneficiaries display a significant increase in the amount saved only in the NG treatment, whereas for direct households this effect was present overall and for the FT and NG groups. For indirect households, we see no increase in amount saved for NVT (the point estimate is actually strongly negative but insignificant). The values-based training appears to be very important for savings group formation and functionality for second generation beneficiaries.

Physical Health

We see no overall effect on the physical health summary index for direct beneficiaries. When we break it down into its components (table 12) we do see a significant effect on respondents' subjective opinion of their children's health (around 0.4 points on a ten point scale) for the NG treatment, the NVT treatment, and the aggregate of treated households. We interpret this result with caution given that those receiving benefits may tend to respond more favorably to highly subjective questions like this than those not receiving benefits. We find nothing noteworthy with respect of the physical health of PIF beneficiaries, either for the summary index or for subindicators (table 13).

Mental Health

We also see no increase in the mental health index, and point estimates are close to zero. We do find some significant effects on individual components of the index (table 14). Most notably, life satisfaction increases by 0.3 points on a ten point scale for households in any treatment. The point estimate for this effect is positive for the distinct treatments, but only significant (and larger) for those receiving the NG package. We find an increase in self-esteem for the aggregated treatments, FT, and NVT. We also find an increase in the worry score (which indicates less worrying) for the full treatment group only. Given the treatment and the effects we find in other areas, we are surprised at the lack of overall effect on mental health. Not only is there a general lack of precision, but the point estimates are small.

As was the case for directly targeted households, we see no significant change in the

mental health summary index for potential indirect households overall or for the disaggregated treatments. We do observe point estimates of treatment effects on the worries score that are slightly stronger than those for direct beneficiaries.

Aspirations

One purported important aspect of Heifer's programs is that they give poor women higher aspirations and the hope of a better future. We observe no change in the aspirations index[8] for the aggregate treatment group or for the disaggregated treatments (table 16). However, we do observe an increase in income aspirations in the full and NG treatment, and a similarly sized coefficient in the estimate for NVB treatment group. This effect is dampened for PIF households. In the NG group we see a marginally significant increase in income aspirations, and point estimates are positive for the other treatment groups.

Empowerment

We see an increase of 0.25-0.30 SD in empowerment as measured by the Women's Empowerment in Agriculture Index (Alkire et al. (2012)). To measure empowerment, we employed the Five Domains of Empowerment (5DE) subindex of the Women's Empowerment in Agriculture Index (WEAI) modified to the local context. This index aggregates an empowerment score across production, resources, income, leadership and time. Looking across the five domains of the WEAI index, we find a statistically significant impact in three sub-indicators. Compared to women in the control group, women who participated in the program are 4.3 percentage points more likely to own productive assets, 4.6 percentage points more likely to have some control over use of income and 15.6 percentage points more likely to belong to a group. Belonging to a group is easily explained through the program design, as beneficiaries join self-help groups. The size of the group membership coefficient can be explained by the fact that the majority of women (65 percent) are already in some kind of group (for example, a mother's group or savings group). However, the increases in asset ownership and control over income are not a mechanical result of receiving livestock.

[8] measured using Bernard and Taffesse (2014)

Beneficiaries in one of the NG treatment arms did not receive livestock and yet we observe a significant impact.

These results are fairly consistent across treatments. One exception is that empowerment in production decisions does not increase as much (or significantly) in the NG treatment group, whereas control over income only increases significantly in that treatment group, and by a larger amount. Importantly, asset ownership does not only increase due to women in treatment groups being given goats; the effect is just as large for the NG treatment as it is for the treatments that include goats.

The Heifer program increases empowerment among PIF beneficiaries in nearly the exact same manner as it does for directly targeted beneficiaries. If anything, the results are larger and statistically stronger.

6.3 Discussion

6.3.1 Livestock Dynamics

In several instances we've asserted that we fail to observe statistically significant improvements in long-range welfare outcomes like income and asset holdings because at the time of midline data collection the program had not been in place long enough for beneficiaries to sell any goats. Conversely, the outcome dimensions that exhibit improved welfare outcomes do not rely on proceeds from the sale of livestock, and might be expected to change as a result of the trainings, self-help group formation, and receipt of the transfer itself. In this subsection we validate this argument by presenting a simple model of asset accumulation that uses goat gestation length, kid growth rates, and market prices to predict the timing and magnitude of income and asset increases. We see that our observed increases in goat herd size are broadly consistent with the predicition, and confirm our intuition that increases in income and financial assets are unlikely to be observable at midline.

Figure 3 illustrates expected herd dynamics for direct (OG) and indirect (POG) bene-

ficiaries in the full treatment and no-values-based-training treatment arms. Calendar years 2015 through 2018 are arranged horizontally across the page, and the expected four generations of goats are represented vertically (heavy black lines delineate each generation; G1 is the original, donated goat). Pink bars represent female goats (does) and blue bars represent male goats (bucks). We assume that a doe can reasonably expect to be impregnated within any given four month window, a five month gestation period, and that offspring reach sexual maturity at around seven months (females) or an optimally marketable size at around 10 months (males).

Direct beneficiaries received livestock between March and June 2015. Depending on breeding cycles and the availability of an improved buck, most participants might have expected to impregnate their does between June and October of 2015, which would imply that the members of a second generation of program goats were on average born near the end of that year and the beginning of 2016. Recalling that the program requires beneficiaries to donate their firstborn female offspring to another beneficiary through the pay-it-forward mechanism, and taking note of the fact that goats normally experience single births (although multiples aren't uncommon), the earliest that the typical beneficiary might have expected to make a sale would have been the fall of 2016, or several months after midline. Other aspects of HI's programming might result in enhanced livestock income earlier on: improved access to veterinary care including deworming and antibiotics might result in lower mortality rates, knowledge of superior nutrition and access to better fodder might improve growth rates and animal size and quality, and improved shelters might lead to better health and larger herd sizes.

In addition to confirming that increases in income and financial assets probably should not have been clearly observable, this model also demonstrates that our point estimates of treatment effects on tropical livestock units fall within the expected range. For direct beneficiary estimates, the expected herd size would be as small as two goats larger than the counterfactual herd (assuming they've received a gift of two female goats and no mortality

or off-take), and no larger than six goats bigger than the counterfactual herd (assuming a gift of two female goats, both of whom had twin births, and no off-take).

6.3.2 Recruitment

Across treatments, direct participation was similarly high at around 75 percent. In the NG and NVT groups participation rates were slightly lower, although not significantly so. Indirect beneficiary recruitment rates display a greater degree of heterogeneity across treatment types. We find that in the full treatment wards, 66 percent of potential indirect households claimed to be actual beneficiaries. This indicates that the pay-it-forward aspect of Heifer's program is very successful under the full program. In the NG treatment arm, recruitment of indirect beneficiary households was slightly lower at 54 percent. This is surprisingly high given that there is no promise of receiving livestock. In the goats only treatment arm, indirect beneficiary recruitment is much lower at only 14 percent. This difference suggests that the values-based training is vital to the self-propagating nature of Heifer' intervention. This could be because it makes membership as an indirect beneficiary much more enticing, or because it incites direct members to recruit indirect members.

6.3.3 Cost-benefit analysis

Cost per beneficiary in our sample varies by treatment arm as well as direct/indirect status. Some program costs are common or shared across all treatments. All three treatment types receive the same human capital and technical trainings, for instance, and much of the NGO overhead and administrative expenses are spread evenly across all SHGs regardless of treatment status. Other costs are not incurred at all in certain treatment arms: the NG treatment arm incurs no costs for livestock, while NVT incurs no costs for values-based trainings. In addition to differential costs across treatment arms, recall that direct beneficiaries take on the responsiblity of passing on the gift of livestock and knowledge to indirect beneficiaries. While HI does conduct some trainings directly and does provide a limited amount of

backstopping to ensure the quality and completeness of the passed gifts, direct beneficiaries shoulder much of the costs associated with indirect beneficiaries. Therefore, any analysis of the costs of the treatment effects presented here must take into account both dimensions for heterogeneous costs.

We collected detailed cost data on all program activities in each treatment arm, and these amounts can be accurately attributed to direct and indirect beneficiaries with a few reasonable assumptions. We present costs per beneficiary broken down by treatment arm and direct/indirect status in table 6.[9] The table is organized into three panels. We report pooled costs associated with direct and indirect beneficiaries in the top panel. We tabulate costs associated exclusively with direct beneficiaries in the middle panel. The bottom panel includes HI's costs for provision of benefits to PIF beneficiaries; these costs would not have been occurred but-for the existence of purposefully programmed indirect beneficiaries. The right-most column aggregates across treatment arms.

Before any disaggregation, we calculate that Heifer spent approximately $137 USD per beneficiary on average. Without considering the distinction between direct and indirect beneficiaries, the data indicate that FT beneficiaries cost about $118 USD each, NG beneficiaries cost $87, and NVT beneficiaries cost $550. NVT beneficiaries cost so much because the direct NVT members do not pay it forward, and therefore do not take advantage of the economies of scale acquired in the exponential recruiting model. These economies of scale come into sharper focus in the bottom two panels. Costs per direct beneficiary are high relative to the aggregate, ranging from $270 to $465 depending on treatment type. With direct beneficiaries in place, however, treating additional beneficiaries implies decreasing marginal costs: the average indirect beneficiary in our sample costs only $65 USD.

These cost data enter a discussion of effectiveness in two ways, the first (type I cost effectiveness) relates to internal progam design and the second (type II) relates to external comparisons with related interventions. First, because we find neligible variation in treat-

[9]The appendix section ?? presents full cost data and details the assumptions and methodology used to arrive at the summarized figures presented in table **??**.

ment effects across treatment arms, it stands to reason that HI would be most interested in providing beneficiaries with the lowest-cost combination of benefits that achieves desired outcomes. Second, with the lowest-cost combination of benefits determined, we are interested in how the costs and benefits of the intervention compare to those of comparable interventions, especially other productive asset transfers and cash transfers.

With respect to type I cost effectiveness, the only conclusion that we might make at this point is also the most obvious: values based training helps recruit, retain, and transfer assets and skills to indirect beneficiaries. At this point we refrain from making any judgement as relative cost effectiveness of the remaining treatment arms (FT and NG), even though per beneficiary cost is lower under the NG regime and the recruitment data suggest that the promise of livestock does not increase uptake.

The two indices where we see improvement are not in the money-metric, so it's difficult to interpret a cost/benefit ratio. As we argue in section 6.3.1, endline may show differential effects for indices like income, assets, and expenditures which are in the money metric.

7 Conclusions

In this study we evaluate the short term impacts of Heifer International's (HI) livestock transfer and training program in Nepal using a randomized controlled trial. We find that in just over one year women beneficiaries are more empowered and connected to financial markets. Our findings suggest that women who participate in a multifaceted social protection program that combines trainings with an asset transfer have immediate effects. While we do not observe statistically significant changes to the longer-run outcomes of income, assets, and expenditures, the timing of livestock transfers has not yet allowed for goat sales. As such, it may be too early to observe these effects. It is also too early to say if the different program components are important for improving long-run economic outcomes. In future work we will measure the strength and persistence of these heterogeneous effects and the cost effectiveness of the program. Measuring the strength and persistence of effects is crucial to understand the full program impacts.

We also observe short-term impacts not only among households who received livestock and training directly from the program, but also for those brought into the program through encouragement to "pay it forward," where other women in the same village are recruited, trained and eventually given livestock by initial Heifer beneficiaries. These findings demonstrate how encouragement to "pay it forward" can help achieve a broader impact at lower cost.

References

Alkire, S., R. Meinzen-Dick, A. Peterman, A. Quisumbing, and G. Seymour. 2012. "A. Vaz. 2012."The Women's Empowerment in Agriculture Index," Poverty, Health & Nutrition Division, International Food Policy Research Institute." Working paper, IFPRI Discussion Paper.

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*, pp. .

Bandiera, O., R. Burgess, N. Das, S. Gulesci, I. Rasul, and M. Sulaiman. 2013. "Can basic entrepreneurship transform the economic lives of the poor?", pp. .

Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Parienté, J. Shapiro, B. Thuysbaert, and C. Udry. 2015. "A multifaceted program causes lasting progress for the very poor: Evidence from six countries." *Science* 348:1260799.

Benjamini, Y., and Y. Hochberg. 1995. "Controlling the false discovery rate: a practical and powerful approach to multiple testing." *Journal of the royal statistical society. Series B (Methodological)*, pp. 289–300.

Bernard, T., and A.S. Taffesse. 2014. "Aspirations: An approach to measurement with validation using Ethiopian data." *Journal of African Economies* 23:189–224.

Darrouzet-Nardi, A.F., L.C. Miller, N. Joshi, S. Mahato, M. Lohani, and B.L. Rogers. 2016. "Child dietary quality in rural Nepal: Effectiveness of a community-level development intervention." *Food Policy* 61:185–197.

Haushofer, J., and J. Shapiro. 2013. "Welfare Effects of Unconditional Cash Transfers: Preanalysis Plan." Unpublished.

Jodlowski, M., A. Winter-Nelson, K. Baylis, and P.D. Goldsmith. 2016. "Milk in the data: food security impacts from a livestock field experiment in Zambia." *World Development* 77:99–114.

Kafle, K., A. Winter-Nelson, and P. Goldsmith. 2016. "Does 25 cents more per day make a difference? The impact of livestock transfer and development in rural Zambia." *Food Policy* 63:62–72.

Miller, L.C., N. Joshi, M. Lohani, B. Rogers, M. Loraditch, R. Houser, P. Singh, and S. Mahato. 2014. "Community development and livestock promotion in rural Nepal: Effects on child growth and health." *Food and nutrition bulletin* 35:312–326.

Rawlins, R., S. Pimkina, C.B. Barrett, S. Pedersen, and B. Wydick. 2014. "Got milk? The impact of Heifer International's livestock donation programs in Rwanda on nutritional outcomes." *Food Policy* 44:202–213.

Roy, S., J. Ara, N. Das, and A.R. Quisumbing. 2015. ""Flypaper effects" in transfers targeted to women: Evidence from BRAC's "Targeting the Ultra Poor" program in Bangladesh." *Journal of Development Economics* 117:1–19.

Table 1: Elements of treatment by treatment arm

Treatment 1(FT)	**Treatment 2(NG)**	**Treatment 3(NSC)**
Livestock placement		Livestock placement
Support for home garden	Support for home garden	Support for home garden
Support for fodder/forage	Support for fodder/forage	Support for fodder/forage
Support for shed improvement		Support for shed improvement
Trainings:	Trainings:	Trainings:
•Cornerstones	•Cornerstones	
•SHG management	•SHG management	
•Exposure visit	•Exposure visit	
•Gender justice	•Gender justice	
•Improved Animal Management	•Improved Animal Management	•Improved Animal Management
•Fodder/forage development	•Fodder/forage development	•Fodder/forage development
•Nutrition	•Nutrition	•Nutrition
•Home gardening	•Home gardening	•Home gardening
•CAVE	•CAVE	•CAVE

Figure 1: VDC and ward sampling structure

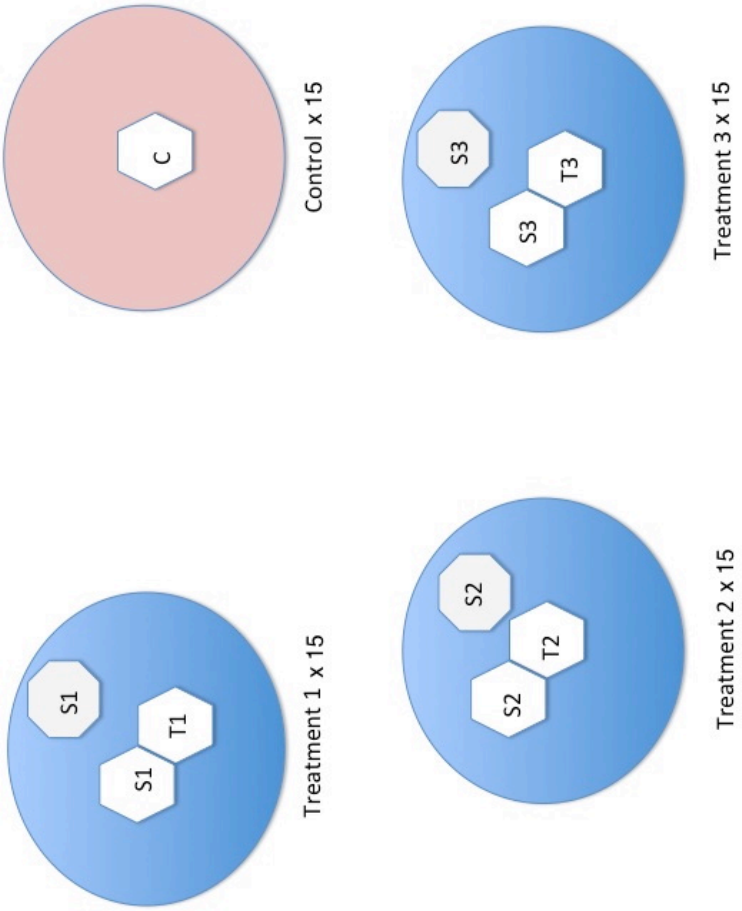

Figure 2: Household sampling structure

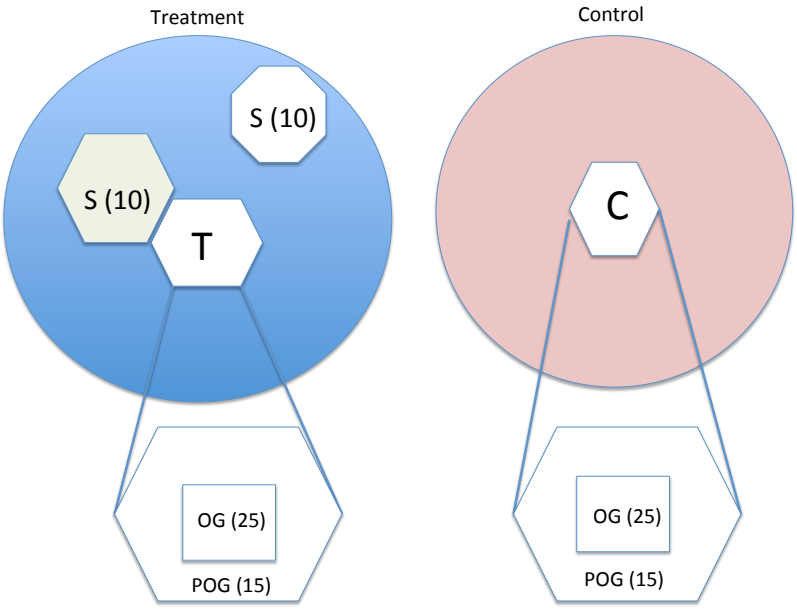

Table 2: ITT effects on summary indices for direct beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Asset index	0.080	0.006	-0.013	0.009	0.017	-0.022	-0.030	-0.008	1,031
	(1.032)	(0.109)	(0.125)	(0.152)	(0.116)	(0.874)	(0.779)	(0.956)	
Income (Rs.)	11.502	-0.030	-0.139	0.100	-0.058	-0.239*	-0.081	0.158	1,030
	(1.227)	(0.138)	(0.172)	(0.143)	(0.162)	(0.059)	(0.597)	(0.206)	
Non-food consumption	7.994	-0.154	-0.303	-0.277	0.069	-0.025	-0.372	-0.346	1,031
	(2.422)	(0.283)	(0.335)	(0.354)	(0.318)	(0.945)	(0.212)	(0.290)	
Finance	0.071	0.308***	0.307**	0.312**	0.304**	-0.006	0.003	0.008	1,031
	(1.006)	(0.113)†	(0.128)	(0.146)	(0.128)	(0.969)	(0.979)	(0.949)	
Physical health	0.041	0.048	-0.007	0.057	0.083	-0.064	-0.090	-0.026	1,030
	(0.813)	(0.069)	(0.075)	(0.103)	(0.088)	(0.544)	(0.325)	(0.816)	
Mental health	0.009	0.065	0.175	-0.065	0.093	0.240**	0.082	-0.158	1,031
	(0.969)	(0.080)	(0.107)	(0.104)	(0.090)	(0.048)	(0.432)	(0.131)	
Aspirations	-0.854	1.536	0.450	3.021	1.047	-2.572	-0.598	1.974	1,030
	(12.597)	(1.610)	(1.527)	(2.417)	(1.886)	(0.249)	(0.695)	(0.406)	
Empowerment	0.777	0.045**	0.043**	0.044*	0.048**	-0.000	-0.005	-0.004	1,020
	(0.189)	(0.018)†	(0.020)	(0.023)	(0.022)	(0.989)	(0.813)	(0.836)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. A † superscript affixed to the standard error of a regression coefficient indicates an FDR-adjusted q<.01. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 3: ITT effects on summary indices for indirect beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Asset index	-0.043	0.024	0.130	-0.060	0.002	0.190**	0.128	-0.062	797
	(1.073)	(0.097)	(0.123)	(0.099)	(0.108)	(0.037)	(0.209)	(0.422)	
Income (Rs.)	11.492	-0.116	-0.150	-0.186	-0.029	0.036	-0.121	-0.157	798
	(1.135)	(0.138)	(0.187)	(0.153)	(0.151)	(0.821)	(0.490)	(0.250)	
Non-food consumption	7.883	0.205	0.245	0.023	0.323	0.222	-0.079	-0.300	797
	(2.449)	(0.276)	(0.313)	(0.368)	(0.270)	(0.480)	(0.725)	(0.300)	
Finance	-0.067	0.213**	0.223**	0.339***	0.099	-0.116	0.124	0.240**	797
	(1.037)	(0.082)†	(0.097)	(0.103)	(0.110)	(0.283)	(0.271)	(0.029)	
Physical health	0.064	-0.023	-0.148	0.163	-0.066	-0.312***	-0.083	0.229**	796
	(0.932)	(0.082)	(0.107)	(0.098)	(0.094)	(0.007)	(0.471)	(0.019)	
Mental health	0.025	0.035	0.054	-0.165	0.186	0.219***	-0.131*	-0.351***	797
	(0.905)	(0.099)	(0.094)	(0.115)	(0.112)	(0.008)	(0.087)	(0.001)	
Aspirations	-1.286	0.319	-1.224	1.540	0.632	-2.764	-1.855	0.909	794
	(15.822)	(1.408)	(2.188)	(1.523)	(1.588)	(0.197)	(0.420)	(0.560)	
Empowerment	0.749	0.057***	0.065***	0.079***	0.031	-0.014	0.034	0.048**	787
	(0.195)	(0.017)†	(0.023)	(0.021)	(0.021)	(0.545)	(0.167)	(0.028)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. A † superscript affixed to the standard error of a regression coefficient indicates an FDR-adjusted q<.01. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 4: ITT effects on assets for direct beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Asset index	0.080	0.006	-0.013	0.009	0.017	-0.022	-0.030	-0.008	1,031
	(1.032)	(0.109)	(0.125)	(0.152)	(0.116)	(0.874)	(0.779)	(0.956)	
Productive asset index	0.134	-0.001	0.001	0.008	-0.010	-0.007	0.011	0.018*	1,028
	(0.068)	(0.008)	(0.008)	(0.009)	(0.010)	(0.367)	(0.211)	(0.072)	
Non-productive asset index	0.475	-0.009	-0.019	-0.004	-0.007	-0.015	-0.012	0.003	1,029
	(0.083)	(0.013)	(0.013)	(0.015)	(0.013)	(0.189)	(0.194)	(0.778)	
Livestock (TLU)	2.576	0.208	0.265	0.042	0.307	0.223	-0.042	-0.265	1,031
	(2.353)	(0.165)	(0.169)	(0.232)	(0.222)	(0.311)	(0.849)	(0.308)	
Land (hectares)	0.470	0.057*	0.054*	0.009	0.100**	0.045	-0.045	-0.091**	1,028
	(0.538)	(0.029)	(0.030)	(0.038)	(0.038)	(0.238)	(0.238)	(0.042)	
Housing index	2.531	-0.016	0.053	-0.110	0.016	0.163	0.037	-0.125	1,031
	(0.823)	(0.091)	(0.131)	(0.109)	(0.097)	(0.194)	(0.740)	(0.229)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 5: ITT effects on assets for indirect beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Asset index	-0.043	0.024	0.130	-0.060	0.002	0.190**	0.128	-0.062	797
	(1.073)	(0.097)	(0.123)	(0.099)	(0.108)	(0.037)	(0.209)	(0.422)	
Productive asset index	0.135	0.006	0.005	0.014	0.000	-0.009	0.004	0.014	794
	(0.069)	(0.007)	(0.007)	(0.010)	(0.008)	(0.258)	(0.512)	(0.157)	
Non-productive asset index	0.472	-0.013	-0.010	-0.017	-0.011	0.007	0.000	-0.007	794
	(0.082)	(0.010)	(0.011)	(0.012)	(0.011)	(0.416)	(0.960)	(0.471)	
Livestock (TLU)	2.359	0.208	0.440**	-0.021	0.200	0.462**	0.241	-0.221	797
	(2.053)	(0.135)	(0.188)	(0.163)	(0.164)	(0.020)	(0.229)	(0.245)	
Land (hectares)	0.392	0.031	0.040	0.024	0.029	0.016	0.010	-0.006	794
	(0.474)	(0.034)	(0.039)	(0.040)	(0.040)	(0.640)	(0.771)	(0.874)	
Housing index	2.423	0.053	0.166	-0.076	0.062	0.242**	0.104	-0.138*	797
	(0.937)	(0.109)	(0.150)	(0.104)	(0.118)	(0.039)	(0.409)	(0.086)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 6: ITT effects on income for direct beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Total income	11.502	-0.030	-0.139	0.100	-0.058	-0.239*	-0.081	0.158	1,030
	(1.227)	(0.138)	(0.172)	(0.143)	(0.162)	(0.059)	(0.597)	(0.206)	
Livestock income	4.908	0.360	-0.116	0.564	0.549	-0.680	-0.664	0.015	1,030
	(5.100)	(0.272)	(0.369)	(0.399)	(0.333)	(0.126)	(0.121)	(0.972)	
Crop income	3.274	-0.559	-0.974**	-0.469	-0.312	-0.506	-0.662	-0.157	1,030
	(4.902)	(0.410)	(0.465)	(0.461)	(0.500)	(0.267)	(0.182)	(0.712)	
Permanent income	2.317	0.067	-0.476	0.631	-0.001	-1.107	-0.475	0.631	1,030
	(4.699)	(0.434)	(0.495)	(0.655)	(0.551)	(0.103)	(0.405)	(0.379)	
Business income	2.750	0.525	0.213	1.146*	0.230	-0.933	-0.017	0.916*	1,030
	(4.633)	(0.497)	(0.677)	(0.583)	(0.511)	(0.110)	(0.976)	(0.056)	
Cash income	2.827	-0.150	-0.415	0.112	-0.172	-0.527	-0.243	0.284	1,030
	(5.012)	(0.687)	(0.712)	(0.803)	(0.783)	(0.353)	(0.684)	(0.674)	
Other income	2.241	0.570	0.418	0.935	0.371	-0.517	0.046	0.564	1,030
	(4.505)	(0.495)	(0.618)	(0.612)	(0.517)	(0.351)	(0.929)	(0.274)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 7: ITT effects on income for indirect beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Total income	11.492	-0.116	-0.150	-0.186	-0.029	0.036	-0.121	-0.157	798
	(1.135)	(0.138)	(0.187)	(0.153)	(0.151)	(0.821)	(0.490)	(0.250)	
Livestock income	4.382	1.514***	2.054***	1.276**	1.241***	0.778	0.813	0.035	798
	(4.977)	(0.448)	(0.617)	(0.603)	(0.462)	(0.280)	(0.148)	(0.949)	
Crop income	2.852	0.275	0.605	0.315	-0.052	0.290	0.657	0.367	798
	(4.631)	(0.425)	(0.560)	(0.374)	(0.673)	(0.562)	(0.374)	(0.565)	
Permanent income	2.608	-0.289	-0.471	0.166	-0.508	-0.636	0.037	0.674	798
	(4.904)	(0.476)	(0.497)	(0.646)	(0.588)	(0.264)	(0.941)	(0.299)	
Business income	2.722	0.533	0.064	1.239**	0.359	-1.174*	-0.295	0.880**	798
	(4.647)	(0.544)	(0.800)	(0.569)	(0.541)	(0.092)	(0.659)	(0.037)	
Cash income	2.478	-0.096	-0.584	-0.273	0.483	-0.310	-1.066*	-0.756	798
	(4.795)	(0.567)	(0.625)	(0.705)	(0.623)	(0.598)	(0.055)	(0.210)	
Other income	2.720	0.407	0.207	0.645	0.386	-0.438	-0.179	0.259	798
	(4.838)	(0.534)	(0.660)	(0.614)	(0.577)	(0.449)	(0.740)	(0.590)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 8: ITT effects on non-food expenditures for direct beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Total expenditure	7.994	-0.154	-0.303	-0.277	0.069	-0.025	-0.372	-0.346	1,031
	(2.422)	(0.283)	(0.335)	(0.354)	(0.318)	(0.945)	(0.212)	(0.290)	
Medical expenditures	6.713	-0.236	-0.216	0.130	-0.569	-0.346	0.353	0.699*	1,031
	(3.578)	(0.293)	(0.381)	(0.388)	(0.339)	(0.415)	(0.359)	(0.073)	
Clothing expenditures	7.358	0.044	-0.242	0.163	0.160	-0.405**	-0.401**	0.004	1,031
	(1.805)	(0.189)	(0.240)	(0.199)	(0.197)	(0.047)	(0.032)	(0.979)	
Misc. expenditures	6.084	-0.672*	-0.691	-0.854*	-0.495	0.162	-0.196	-0.358	1,031
	(3.468)	(0.383)	(0.470)	(0.451)	(0.455)	(0.732)	(0.658)	(0.405)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 9: ITT effects on non-food expenditures for indirect beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Total expenditure	7.883	0.205	0.245	0.023	0.323	0.222	-0.079	-0.300	797
	(2.449)	(0.276)	(0.313)	(0.368)	(0.270)	(0.480)	(0.725)	(0.300)	
Medical expenditures	6.633	-0.216	0.201	-0.386	-0.436	0.586**	0.637**	0.051	797
	(3.784)	(0.317)	(0.319)	(0.345)	(0.383)	(0.032)	(0.050)	(0.874)	
Clothing expenditures	7.446	-0.076	-0.119	-0.042	-0.065	-0.077	-0.054	0.023	797
	(1.743)	(0.143)	(0.176)	(0.184)	(0.147)	(0.669)	(0.710)	(0.877)	
Misc. expenditures	6.150	-0.786*	-0.267	-1.184**	-0.887*	0.917*	0.620	-0.297	797
	(3.126)	(0.430)	(0.572)	(0.466)	(0.513)	(0.085)	(0.275)	(0.520)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 10: ITT effects on financial inclusion for direct beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Financial index	0.071	0.308***	0.307**	0.312**	0.304**	-0.006	0.003	0.008	1,031
	(1.006)	(0.113)	(0.128)	(0.146)	(0.128)	(0.969)	(0.979)	(0.949)	
Amount saved	3.824	0.704**	0.735**	1.062***	0.364	-0.327	0.371	0.699**	1,031
	(2.930)	(0.269)	(0.323)	(0.313)	(0.337)	(0.307)	(0.243)	(0.034)	
Savings group	0.539	0.170***	0.195***	0.184***	0.137**	0.012	0.058	0.046	1,025
	(0.499)	(0.051)	(0.061)	(0.065)	(0.063)	(0.864)	(0.337)	(0.472)	
Owe formal lender	3.479	0.128	-0.011	0.100	0.257	-0.111	-0.267	-0.156	1,031
	(5.275)	(0.605)	(0.644)	(0.847)	(0.697)	(0.887)	(0.664)	(0.846)	
Owe informal lender	2.962	-0.650	-0.563	-0.975*	-0.435	0.412	-0.129	-0.540	1,031
	(4.915)	(0.397)	(0.481)	(0.577)	(0.439)	(0.498)	(0.776)	(0.311)	
Discount rate	0.053	-0.012	-0.023**	-0.003	-0.012	-0.020	-0.012	0.009	746
	(0.079)	(0.011)	(0.011)	(0.014)	(0.013)	(0.108)	(0.246)	(0.489)	
Planning horizon	1.829	0.069	0.059	-0.067	0.195	0.126	-0.136	-0.262*	1,028
	(0.943)	(0.122)	(0.144)	(0.154)	(0.143)	(0.427)	(0.327)	(0.092)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 11: ITT effects on financial inclusion for indirect beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Financial index	-0.067	0.213**	0.223**	0.339***	0.099	-0.116	0.124	0.240**	797
	(1.037)	(0.082)	(0.097)	(0.103)	(0.110)	(0.283)	(0.271)	(0.029)	
Amount saved	3.529	0.174	0.094	0.682**	-0.176	-0.588*	0.270	0.858**	797
	(2.977)	(0.257)	(0.304)	(0.310)	(0.343)	(0.082)	(0.477)	(0.020)	
Savings group	0.492	0.101**	0.131**	0.185***	0.005	-0.053	0.126*	0.180***	795
	(0.501)	(0.047)	(0.062)	(0.057)	(0.061)	(0.396)	(0.070)	(0.003)	
Owe formal lender	2.508	0.379	0.320	0.772	0.105	-0.453	0.214	0.667	797
	(4.729)	(0.502)	(0.581)	(0.630)	(0.669)	(0.483)	(0.748)	(0.342)	
Owe informal lender	2.929	-0.285	-0.476	-0.579	0.125	0.103	-0.601	-0.704	797
	(4.974)	(0.392)	(0.422)	(0.472)	(0.476)	(0.780)	(0.148)	(0.111)	
Discount rate	0.043	0.001	-0.013	0.010	0.004	-0.023**	-0.017	0.006	584
	(0.064)	(0.010)	(0.009)	(0.012)	(0.013)	(0.044)	(0.132)	(0.663)	
Planning horizon	1.755	0.157	0.147	0.082	0.227	0.065	-0.080	-0.145	794
	(0.961)	(0.106)	(0.110)	(0.127)	(0.143)	(0.614)	(0.527)	(0.292)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 12: ITT effects on physical health for direct beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Physical health index	0.041	0.048	-0.007	0.057	0.083	-0.064	-0.090	-0.026	1,030
	(0.813)	(0.069)	(0.075)	(0.103)	(0.088)	(0.544)	(0.325)	(0.816)	
Days work missed	1.554	0.198	0.119	0.384	0.098	-0.265	0.021	0.286	1,029
	(2.985)	(0.260)	(0.372)	(0.391)	(0.327)	(0.584)	(0.961)	(0.513)	
Subjective own health	6.468	0.070	-0.087	0.222	0.057	-0.308	-0.144	0.164	1,029
	(1.886)	(0.129)	(0.168)	(0.193)	(0.147)	(0.167)	(0.424)	(0.424)	
Subjective child health	7.048	0.364**	0.194	0.474**	0.403**	-0.280	-0.209	0.071	675
	(1.600)	(0.158)	(0.182)	(0.216)	(0.191)	(0.226)	(0.294)	(0.754)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 13: ITT effects on physical health for indirect beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Physical health index	0.064	-0.023	-0.148	0.163	-0.066	-0.312***	-0.083	0.229**	796
	(0.932)	(0.082)	(0.107)	(0.098)	(0.094)	(0.007)	(0.471)	(0.019)	
Days work missed	1.582	0.381	0.583	0.221	0.338	0.362	0.245	-0.117	794
	(3.533)	(0.315)	(0.418)	(0.453)	(0.328)	(0.462)	(0.512)	(0.771)	
Subjective own health	6.500	0.020	-0.166	0.510**	-0.228	-0.676***	0.062	0.738***	794
	(1.858)	(0.177)	(0.245)	(0.206)	(0.192)	(0.007)	(0.804)	(0.001)	
Subjective child health	7.134	0.190	-0.032	0.374	0.208	-0.405	-0.240	0.166	517
	(1.797)	(0.269)	(0.340)	(0.304)	(0.293)	(0.201)	(0.438)	(0.477)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 14: ITT effects on mental health for direct beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Mental health index	0.009	0.065	0.175	-0.065	0.093	0.240**	0.082	-0.158	1,031
	(0.969)	(0.080)	(0.107)	(0.104)	(0.090)	(0.048)	(0.432)	(0.131)	
Depression score	6.541	0.103	0.271	-0.101	0.150	0.372*	0.121	-0.251	1,031
	(1.906)	(0.204)	(0.240)	(0.227)	(0.231)	(0.066)	(0.572)	(0.188)	
Locus of control	2.915	-0.062	-0.200	0.054	-0.056	-0.254	-0.144	0.110	1,031
	(1.482)	(0.132)	(0.159)	(0.162)	(0.169)	(0.128)	(0.399)	(0.554)	
Optimism	6.404	-0.001	0.086	-0.047	-0.026	0.133	0.112	-0.022	1,031
	(1.184)	(0.124)	(0.161)	(0.175)	(0.138)	(0.490)	(0.457)	(0.892)	
Life Satisfaction	6.361	0.300*	0.242	0.485**	0.180	-0.243	0.062	0.305	1,029
	(1.904)	(0.178)	(0.227)	(0.223)	(0.196)	(0.291)	(0.773)	(0.141)	
Self-esteem	9.604	0.251**	0.330**	0.197	0.237*	0.133	0.094	-0.040	1,031
	(1.655)	(0.107)	(0.160)	(0.140)	(0.134)	(0.458)	(0.585)	(0.789)	
Happiness	2.070	-0.045	-0.012	-0.087	-0.034	0.075	0.022	-0.053	1,031
	(0.530)	(0.054)	(0.065)	(0.060)	(0.064)	(0.186)	(0.732)	(0.356)	
Worry score	9.075	0.224	0.515**	-0.046	0.242	0.561***	0.273	-0.288	1,016
	(2.091)	(0.195)	(0.200)	(0.230)	(0.225)	(0.003)	(0.144)	(0.170)	
Trust score	1.528	-0.018	0.081	-0.262	0.118	0.343*	-0.037	-0.380*	1,028
	(1.334)	(0.160)	(0.169)	(0.224)	(0.181)	(0.089)	(0.832)	(0.078)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 15: ITT effects on mental health for indirect beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Mental health index	0.025	0.035	0.054	-0.165	0.186	0.219***	-0.131*	-0.351***	797
	(0.905)	(0.099)	(0.094)	(0.115)	(0.112)	(0.008)	(0.087)	(0.001)	
Depression score	6.594	0.020	0.261	-0.176	-0.031	0.437*	0.291	-0.145	797
	(1.897)	(0.178)	(0.185)	(0.265)	(0.205)	(0.094)	(0.142)	(0.577)	
Locus of control	2.985	-0.113	-0.078	-0.222	-0.051	0.144	-0.027	-0.170	797
	(1.390)	(0.116)	(0.122)	(0.149)	(0.164)	(0.279)	(0.860)	(0.339)	
Optimism	6.345	0.051	0.040	-0.061	0.152	0.102	-0.112	-0.213	797
	(1.234)	(0.155)	(0.162)	(0.164)	(0.186)	(0.460)	(0.400)	(0.148)	
Life Satisfaction	6.408	0.101	-0.053	0.262	0.103	-0.315	-0.156	0.159	794
	(1.960)	(0.182)	(0.226)	(0.206)	(0.188)	(0.115)	(0.427)	(0.348)	
Self-esteem	9.812	-0.019	-0.037	-0.036	0.010	-0.001	-0.046	-0.045	797
	(1.876)	(0.162)	(0.197)	(0.204)	(0.236)	(0.996)	(0.859)	(0.859)	
Happiness	2.061	-0.031	-0.063	-0.067	0.026	0.005	-0.089*	-0.094*	797
	(0.541)	(0.056)	(0.059)	(0.068)	(0.060)	(0.926)	(0.052)	(0.090)	
Worry score	8.902	0.366*	0.550**	0.088	0.436*	0.462	0.114	-0.348	787
	(1.957)	(0.215)	(0.229)	(0.364)	(0.240)	(0.204)	(0.620)	(0.323)	
Trust score	1.520	0.114	0.359*	-0.265	0.215	0.623**	0.144	-0.479*	794
	(1.353)	(0.197)	(0.208)	(0.265)	(0.242)	(0.015)	(0.531)	(0.070)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 16: ITT effects on aspirations for direct beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Aspirations index	-0.854	1.536	0.450	3.021	1.047	-2.572	-0.598	1.974	1,030
	(12.597)	(1.610)	(1.527)	(2.417)	(1.886)	(0.249)	(0.695)	(0.406)	
Income aspirations	11.043	0.562**	0.434*	0.809**	0.449	-0.375	-0.015	0.360	1,030
	(3.038)	(0.241)	(0.240)	(0.347)	(0.276)	(0.228)	(0.949)	(0.282)	
Asset aspirations	13.794	0.035	-0.147	0.334	-0.090	-0.480	-0.057	0.423	1,030
	(2.754)	(0.257)	(0.271)	(0.392)	(0.357)	(0.233)	(0.875)	(0.356)	
Children's education aspirations	14.885	-0.033	-0.592	0.154	0.231	-0.746	-0.823*	-0.077	1,030
	(3.581)	(0.470)	(0.513)	(0.623)	(0.548)	(0.219)	(0.067)	(0.900)	
Daughters' education aspirations	14.185	-0.239	-1.004	0.123	0.033	-1.127	-1.037*	0.090	1,030
	(4.086)	(0.551)	(0.606)	(0.754)	(0.613)	(0.119)	(0.054)	(0.900)	
Sons' education aspirations	14.581	-0.070	-0.397	0.243	-0.094	-0.639	-0.302	0.337	1,030
	(3.810)	(0.417)	(0.447)	(0.593)	(0.547)	(0.288)	(0.546)	(0.612)	
Status aspirations	15.567	2.686	3.800	4.353	0.369	-0.554	3.431	3.985	1,030
	(19.168)	(2.162)	(2.501)	(2.914)	(2.496)	(0.848)	(0.158)	(0.157)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 17: ITT effects on aspirations for indirect beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Aspirations index	-1.286	0.319	-1.224	1.540	0.632	-2.764	-1.855	0.909	794
	(15.822)	(1.408)	(2.188)	(1.523)	(1.588)	(0.197)	(0.420)	(0.560)	
Income aspirations	10.989	0.281	0.072	0.663*	0.149	-0.591*	-0.077	0.515	795
	(3.221)	(0.330)	(0.422)	(0.361)	(0.361)	(0.092)	(0.842)	(0.107)	
Asset aspirations	13.744	-0.113	-0.382	0.179	-0.128	-0.561	-0.254	0.308	795
	(2.662)	(0.244)	(0.383)	(0.242)	(0.295)	(0.133)	(0.533)	(0.262)	
Children's education aspirations	14.617	0.153	0.017	-0.287	0.650	0.305	-0.633	-0.937**	795
	(3.753)	(0.376)	(0.496)	(0.418)	(0.423)	(0.478)	(0.200)	(0.021)	
Daughters' education aspirations	13.556	0.058	-0.186	-0.156	0.462	-0.030	-0.648	-0.618	794
	(4.427)	(0.399)	(0.489)	(0.494)	(0.465)	(0.952)	(0.190)	(0.201)	
Sons' education aspirations	14.500	-0.121	-0.351	-0.299	0.236	-0.052	-0.586	-0.535	794
	(3.649)	(0.315)	(0.468)	(0.341)	(0.370)	(0.907)	(0.234)	(0.161)	
Status aspirations	15.561	0.975	-0.125	2.037	1.053	-2.162	-1.179	0.983	795
	(20.606)	(2.093)	(2.507)	(2.537)	(2.454)	(0.388)	(0.625)	(0.664)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 18: ITT effects on empowerment for direct beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Empowerment Index	0.777	0.045**	0.043**	0.044*	0.048**	-0.000	-0.005	-0.004	1,020
	(0.189)	(0.018)	(0.020)	(0.023)	(0.022)	(0.989)	(0.813)	(0.836)	
Production decisions	0.903	0.030	0.040*	0.006	0.041	0.034	-0.001	-0.035	1,030
	(0.296)	(0.023)	(0.021)	(0.032)	(0.025)	(0.163)	(0.955)	(0.236)	
Asset owernship	0.926	0.043**	0.043**	0.046**	0.041**	-0.003	0.002	0.005	1,030
	(0.262)	(0.016)	(0.017)	(0.019)	(0.020)	(0.835)	(0.917)	(0.787)	
Access to and control over credit	0.416	-0.027	-0.043	-0.022	-0.018	-0.022	-0.025	-0.003	1,025
	(0.494)	(0.044)	(0.051)	(0.065)	(0.055)	(0.751)	(0.668)	(0.963)	
Control over income	0.892	0.046**	0.021	0.078***	0.038	-0.057**	-0.018	0.039	1,026
	(0.311)	(0.022)	(0.025)	(0.025)	(0.029)	(0.026)	(0.546)	(0.148)	
Group membership	0.651	0.156***	0.139***	0.167***	0.158***	-0.028	-0.019	0.009	1,027
	(0.478)	(0.035)	(0.051)	(0.045)	(0.040)	(0.616)	(0.699)	(0.836)	
Works $\leq$ 10.5 hours per day	0.685	-0.008	0.020	-0.055	0.012	0.074	0.007	-0.067	1,031
	(0.465)	(0.027)	(0.039)	(0.036)	(0.041)	(0.117)	(0.887)	(0.166)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 19: ITT effects on empowerment for indirect beneficiaries

	Control mean	Any	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Empowerment Index	0.749	0.057***	0.065***	0.079***	0.031	-0.014	0.034	0.048**	787
	(0.195)	(0.017)	(0.023)	(0.021)	(0.021)	(0.545)	(0.167)	(0.028)	
Production decisions	0.843	0.107***	0.118***	0.104***	0.099***	0.015	0.020	0.005	796
	(0.365)	(0.029)	(0.033)	(0.030)	(0.032)	(0.465)	(0.454)	(0.822)	
Asset owernship	0.918	0.051**	0.057***	0.061***	0.037*	-0.004	0.020	0.024	795
	(0.275)	(0.019)	(0.021)	(0.022)	(0.020)	(0.825)	(0.168)	(0.113)	
Access to and control over credit	0.325	0.058	0.019	0.099*	0.058	-0.080	-0.039	0.041	793
	(0.470)	(0.039)	(0.044)	(0.052)	(0.050)	(0.129)	(0.438)	(0.454)	
Control over income	0.893	0.054**	0.030	0.086***	0.049*	-0.056**	-0.019	0.037*	794
	(0.310)	(0.026)	(0.031)	(0.027)	(0.026)	(0.017)	(0.431)	(0.077)	
Group membership	0.577	0.113**	0.130**	0.191***	0.034	-0.061	0.095	0.156**	794
	(0.495)	(0.048)	(0.056)	(0.063)	(0.060)	(0.350)	(0.139)	(0.019)	
Works $\leq$ 10.5 hours per day	0.706	-0.034	0.005	-0.060	-0.046	0.065	0.051	-0.014	797
	(0.457)	(0.034)	(0.033)	(0.055)	(0.050)	(0.226)	(0.295)	(0.834)	

OLS regressions, clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. N.B. that FT=NG, FT=NSC, NG=NSC represent Wald tests of a null of equal treatment effects, p-values are reported as the sub-statistic. Control variables include baseline dependent variable, stratification bin dummies, and imbalanced variables at baseline. FT: full treatment, NG: no-goats treatment, NSC: no-values-based-training treatment.

Table 20: Treatment compliance by arm and direct/indirect

	Control mean (SD)	FT	NG	NSC	N
Original Group (OG)	0.113	0.666	0.651	0.622	1,088
	(0.318)	(0.069)	(0.066)	(0.063)	
Passing of the Gift (POG)	0.060	0.671	0.544	0.138	887
	(0.237)	(0.052)	(0.078)	(0.063)	

Table 21: Costs by treatment arm and direct/indirect

		Treatment Arm			
	OG+POG	*T1*	*T2*	*T3*	*All T*
Operations	Livestock	7,926,795	18,000	7,109,931	15,054,725
	Horticulture	2,388,577	2,310,583	1,187,457	5,886,617
	Equipment & Supply	3,409,281	3,772,544	1,975,263	9,157,088
	Trainings	4,639,482	5,167,277	2,043,355	11,850,115
Administrative	Tech Services & Eval	2,083,117	2,148,510	2,095,112	6,326,738
	Personnel	5,056,910	5,056,910	5,056,903	15,170,723
	Office Expenses	1,156,125	1,129,859	1,154,781	3,440,765
	Total	26,660,286	19,603,683	20,622,802	66,886,771
	Per Beneficiary	11,849	8,713	54,994	13,720
	OG				
Operations	Livestock	7,926,795	18,000	7,109,931	15,054,725
	Horticulture	2,388,577	2,310,583	1,187,457	5,886,617
	Equipment & Supply	3,409,281	3,772,544	1,975,263	9,157,088
	Trainings	2,088,512	2,356,270	910,377	5,355,160
Administrative	Tech Services & Eval	416,623	429,702	419,022	1,265,348
	Personnel	1,011,382	1,011,382	1,011,381	3,034,145
	Office Expenses	231,225	225,972	230,956	688,153
	Total	17,472,395	10,124,453	12,844,387	40,441,235
	Per Beneficiary	46,593	26,999	34,252	35,948
	POG				
Operations	Livestock	0	0	0	0
	Horticulture	0	0	0	0
	Equipment & Supply	0	0	0	0
	Trainings	2,020,692	2,020,692	0	4,041,385
Administrative	Tech Services & Eval	2,504,538	2,556,852	0	5,061,390
	Personnel	6,068,289	6,068,289	0	12,136,578
	Office Expenses	1,386,812	1,365,800	0	2,752,612
	Total	11,980,332	12,011,634	0	23,991,965
	Per Beneficiary	6,390	6,406	0	6,398

Nepali rupees, exchange rate roughly 100 NPR per USD.

Figure 3: Goat cycle timeline