Pay it forward: Impacts of a rural livelihoods program with built-in spillovers

Sarah Janzen[1], Nicholas Magnan[2], Sudhindra Sharma[3], and William M. Thompson[4]

[1]University of Illinois, sjanzen@illinois.edu
[2]University of Georgia, nmagnan@uga.edu
[3]Nepa School of Social Sciences and Humanities and IDA Associates, sudhindra@ida.com.np
[4]IDinsight, will.thompson@idinsight.org

June 14, 2021

Abstract

We evaluate 2.5-3.5 year impacts of a livelihoods program using an RCT in Nepal. The program targets women and employs self-help groups, livestock transfers, and trainings. We assigned three variations of the program: full benefits, no livestock, and no values-based training, which includes encouragement to "pay it forward" (PIF) by training and giving livestock to others. With this encouragement in mind, the study is uniquely designed to evaluate the program's ability to generate intentional spillover (PIF) effects. We observe impacts for direct beneficiaries in line with outcomes directly targeted by the intervention - beneficiaries have bigger herds, improved livestock practices, more goat sales and higher profit from goat production. In addition, women are more empowered and have greater financial inclusion. The asset transfer increases program costs, but there is weak evidence suggestive of an independent marginal impact. We observe strong PIF spillover effects, with important implications for cost effectiveness.

Acknowledgements: The USAID Feed the Future BASIS Assets and Market Access Innovation Lab was the primary funder of this research. The work was also supported by the CGIAR Research Program on Agriculture for Nutrition and Health (A4NH) Gender and Assets in Agriculture Project Phase 2 (GAAP2) led by the International Food Policy Research Institute (IFPRI) and funded by the Bill & Melinda Gates Foundation (BMGF), A4NH, and USAID. Heifer International funded all programmatic aspects of the study. We thank Savannah Noray, Giovanni Righi, and Shruti Sharma for research assistance, and Hiranya Baral, Rajendra Pradhan, Samyam Shrestha, Ashika Sharma, and Pranaya Sthapit for their roles in data collection. We also thank Neena Joshi, Kishor Kafle, Rienzzie Kern, Hilary Haddigan, Mahendra Lohani, Shubh Mahato, Dipesh Pokharel and Ben Wood at Heifer International and Heifer Nepal for their roles in the formulation and execution of this research. All errors are our own.

1 Introduction

The rural poor are often assumed to lack access to the productive assets and human and social capital required to be successful entrepreneurs. Productive asset transfer programs are one way non-governmental organizations (NGOs) and governments try to relax these critical constraints. The goal is to facilitate permanent transitions out of persistent poverty through a productive asset transfer combined with extensive context-specific livelihoods training. This naturally raises the question, do asset transfer programs successfully transition poor households into successful entrepreneurs?

In this paper, we evaluate the long-term (2.5 and 3.5 year) impacts of Heifer International's (Heifer) Smallholders in Livestock Value Chain Program (SLVC) in rural Nepal using a randomized controlled trial (RCT). The SLVC program targets women in poor rural communities, and provides a package of benefits that includes a livestock transfer (two doe goats and a shared breeding buck), technical training on improved animal management and entrepreneurship, self-help group (SHG) formation, and values-based training. Rather unique to the Heifer program, the values-based training encourages beneficiaries to "pay it forward" (PIF) by sharing newly acquired technical skills and giving the first-born female offspring of their received goats to another individual in their community. Our study is designed to answer the following questions: What are the impacts of the program, and do they persist over time? Does the PIF mechanism spread impacts? Are all program components necessary? Is the program cost effective?

This paper represents the first large-scale RCT of a Heifer asset transfer program. Prior studies of Heifer programs have employed cross sectional observational data in Rwanda (Rawlins et al., 2014), differences-in-differences in Zambia (Jodlowski et al., 2016; Kafle, Winter-Nelson, and Goldsmith, 2016; Phadera et al., 2019), and a much smaller matched pair RCT in Nepal (Miller et al., 2014; Darrouzet-Nardi et al., 2016). Each of these studies reveal positive impacts. However, they also rely on small sample sizes which could lead to inflated effect sizes or an inflated false discovery rate, raising concerns that inference may be invalid. Only the Zambia studies are equipped to estimate potentially important spillover effects.

Rigorous impact evaluations of bundled asset transfer and training programs, particularly evaluations designed to measure spillover effects, are a recent development. The SLVC is similar to the the six BRAC Graduation Program impact evaluations studied by Banerjee et al. (2015) and the BRAC Targeting the Ultra-poor Program (TUP) considered in Bandiera et al. (2017). Graduation programs take a holistic livelihoods approach to tackling the interrelated challenges faced by the poorest of the poor, bundling a productive asset grant with

technical skills training, access to financial services, intensive monitoring, and short-term cash stipends to support consumption. Evidence suggests that after three to four years, graduation programs increase financial inclusion, psychological well-being, assets, income, total expenditures, food security, and political awareness (Bandiera et al., 2017; Banerjee et al., 2015).

This analysis will contribute to the literature in four important ways. First, we add to a small but growing body of empirical evidence on the welfare impacts of productive asset transfer programs, especially those bundled with extensive training components. Many of the graduation programs previously studied allow beneficiaries to choose a productive asset and associated livelihood. The intervention we study focuses solely on goat livelihoods. This accommodates careful attention to unique goat livelihood mechanisms. This focus is particularly important to our evaluation since the anticipated spillover effects associated with paying benefits forward can redistribute some of the welfare impacts for direct beneficiaries. A focus on mechanisms can help enlighten whether this is indeed happening.

Second, our evaluation is carefully designed to estimate built-in spillover effects in targeted communities. The SLVC employs a distinct targeting and recruitment model characterized by a community-level intervention and encouragement to pay forward the benefits received. Rather than targeting the poorest of the poor (the approach taken by most graduation programs), Heifer recruits all households residing in a targeted (and usually central) neighborhood regardless of relative wealth or poverty. Through values-based training sessions, these directly targeted beneficiaries are then encouraged to establish second generation SHGs among others in their community, with the goal of rapidly scaling benefits. These built-in spillovers embodied in the encouragement to pay benefits forward are an essential feature of all livestock transfer and training programs implemented globally by Heifer. To our knowledge, no study has evaluated the impact of a program with this type of pay it forward (PIF) model. Measuring the strength and persistence of this element of the program design is crucial to understanding overall program impacts.

Third, our evaluation includes three treatments designed to test the impacts of different program components. To our knowledge, previous studies in this area have not attempted to disaggregate the impacts of a similarly bundled treatment.[1] In the first treatment arm of our study, beneficiaries received a full treatment (FT) package that included a livestock transfer, skills-based technical training and values-based non-technical training. In the second treatment arm, beneficiaries received skills-based technical training and values-based non-technical training, but *not* goats (NG). In the third treatment arm beneficiaries received

[1]We are aware of a contemporaneous study by Banerjee et al. (2018) seeking to unpack components of a graduation program in Ghana.

a livestock transfer and skills-based technical training, but *not* values-based non-technical training (NVT). Because encouragement to pay forward benefits is the primary element of the values-based training, the third treatment arm allows analysis of the PIF mechanism.

Finally, we evaluate the persistence of impacts. In Janzen et al. (2018b), we reported the results of a short-term impact analysis after 1.5 years. At 1.5 years we reported significant increases in financial inclusion and women's empowerment, with similar impacts observed for both direct and PIF beneficiaries. These early findings provided preliminary evidence the PIF program component effectively spreads impacts. Here we expand on that analysis with a richer analysis incorporating data 2.5-3.5 years after the beginning of the program. The difference between the two time frames presented within this paper are not inconsequential. Shortly after 2.5 years, the program was considered complete and Heifer no longer interacted with beneficiaries or monitored behavior. Thus, the 3.5 year impacts signal persistent effects beyond the life of the program.

We observe impacts for direct beneficiaries in line with outcomes directly targeted by the intervention – beneficiaries have bigger herds, improved livestock practices, more goat sales, and higher profit from goat production. In addition, women beneficiaries have greater financial inclusion and are more empowered, especially over their goat enterprises. The asset transfer increases program costs, but there is weak evidence to suggest the goats are necessary for achieving impact. We observe no evidence of improvement in total household income or beneficiary psychological well-being. We observe strong spillover effects to PIF beneficiaries, with important implications for cost effectiveness.

The rest of the paper is structured as follows. In the following section, we describe Hl's livelihoods program in Nepal. Section 3 describes our experimental design. Section 4 describes the sample and data, and includes a description of our outcome variables, some descriptive statistics, and balance checks. The details of our empirical approach are presented in section 5, followed by our results in sections 6-8 . We first present the impacts of the program on livestock livelihood impacts, followed by household welfare impacts and heterogenous treatment effects. For ease of comparison, we will concurrently present direct and PIF impacts, as well as differential treatment effects, for each outcome. After presenting our main results, cost-benefit analysis is presented in section 9. Section 10 concludes.

2 Heifer's livelihoods program

Non-governmental organizations and even governments have utilized productive asset transfers for a long time. Heifer, BRAC, Save the Children, World Vision, Oxfam, and the

Government of Rwanda are a few examples. Heifer first employed their now well-known livestock-based approach in 1944, when the organization shipped 17 cows from Arkansas to Puerto Rico. Since then Heifer has expanded its reach to over 125 countries. Although the asset transfer is typically a key component, most asset transfer programs are multifaceted and include aspects of human, social and financial capital development. The goal is to unlock a pathway out of poverty using a new asset-based livelihood. As a result, food and non-food expenditures are expected to increase, with potential impacts on food security and nutrition.

The intervention we evaluate replicates Heifer's flagship program in Nepal, the SLVC. The SLVC intervention targets women and provides a package of benefits including a productive asset transfer (in this case goats), the formation of women's self-help and savings groups, technical trainings on improved animal management and entrepreneurship, and values-based trainings during which beneficiaries are taught to value and practice paying it forward. The process is as follows: After identifying a location for project implementation, Heifer recruits an original group of direct beneficiaries. Direct beneficiary groups typically consist of most or all of the households in a given neighborhood (*tole*). As a rule, Heifer considers all the households in a targeted area to be objectively poor and therefore eligible for the SLVC program, allowing for the possibility that a considerable range of relative wealth and poverty exist within a group. Once selected, direct beneficiaries are organized into a self-help group (SHG). Over a period of several 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) Heifer'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, as well as a shared buck of improved stock for the SHG to facilitate a breeding program.

A unique component of Heifer's approach globally is that it encourages members to pay benefits 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 technical knowledge that was gained through participation in the program. Heifer facilitates values-based empowerment training for both direct and PIF beneficiaries (albeit separately and at different points in time), while all other paid forward trainings are implemented by direct beneficiaries with minimal support from Heifer. As such, what might typically be thought of as a spillover effect is in fact an essential component of the overall program design (thus, "built-in" spillover effect). In Nepal, the SLVC follows an innovation to the basic Heifer PIF

model, in which each direct beneficiary SHG is tasked with recruiting up to five PIF SHGs, with the goal of full saturation and complete adoption of improved practices and technologies within a community over a short time frame.

The Heifer program can improve welfare and facilitate an exit from poverty in many ways. Increasing income generation and associated increases in consumption (food and otherwise) and asset accumulation are primary goals, but the program also addresses others aspects of welfare such as psychological well-being, women's empowerment, and financial inclusion.

Poverty can lead to anxiety, depression, and a loss of self-agency. These factors can in turn lead to behavior that is not proactive or future oriented, making it more difficult to escape poverty Haushofer and Fehr (2014). Anti-poverty programs such as the values-based training of the SLVC increasingly include activities designed to improve psychological well-being in order to incite behavioral change that can lead to better outcomes. And, improved economic circumstances can lead to better psychological well-being, a desirable outcome on its own.

Financial inclusion is also thought to be an important aspect of transitioning out of poverty. Poor people often lack access to formal savings and lending, leading them to borrow at very high interest rates, forgo potentially profitable investments, and engage in costly consumption reduction in the face of shocks. An extensive literature shows access to credit can spur business creation and increase investment, food consumption, income, subjective well-being, and empowerment, but it by no means accomplishes all goals or works in all contexts (see Cull, Ehrbeck, and Holle (2014) for a brief review and and Banerjee (2013) for a more expansive one). Access to saving mechanisms is less researched, but most studies show it does improve outcomes in a number of ways including increasing business investment, agricultural production, managing shocks, and consumption (see Cull, Ehrbeck, and Holle (2014) for a brief review and Karlan, Ratan, and Zinman (2014) for a more expansive one). Financial inclusion can also facilitate investment in education, and help smooth consumption in the wake of negative shocks Ksoll et al. (2016); Karlan et al. (2017). The program aims to improve household finance, and women's access to financial services in particular, through the creation of village credit and savings groups, and improving money management is part of Heifer's values-based training. Furthermore, households can increase their savings through increased income generation.

Financial inclusion for women is believed to be especially important, as they generally have less control over household economic resources and are less likely to hold a formal financial account than men (Halloway, Ziazi, and Rouse, 2017). Interventions to improve women's financial inclusion have also been shown to increase women's empowerment (Ashraf,

Karlan, and Yin, 2010).[2] In addition to being a desirable outcome in and of itself (as evidenced by its inclusion as a Millennium Development Goal), women's empowerment is thought to be an essential component to achieving economic efficiency and improving a variety of household welfare outcomes. Women serve a prominent role in food production, acquisition, and preparation. Women's empowerment therefore can improve agricultural production, food security, and children's nutrition (Alkire et al., 2013; Malapit et al., 2019; Food and Agriculture Organization of the United Nations, 2011). Women's empowerment can also lead to better educational outcomes for children, particularly girls, leading to a more productive workforce and further increases in women's empowerment (World Bank, 2011).

Prior research has established positive welfare impacts of Heifer programs in a number of different contexts. The Heifer program in Zambia has been the subject of several such studies. In Zambia, the transfer includes a dairy cow, two draft cattle, or eight goats. Valued at approximately 2,000 USD, this transfer is five times as expensive as the SLVC in Nepal per direct beneficiary. Using difference-in-differences, Jodlowski et al. (2016) find evidence of increased consumption (expenditures) and dietary diversity. The analysis treats the intervention as a household-level one (rather than community-level), and includes PIF beneficiaries in the control group. Thus, they are unable to estimate impacts on PIF households and may find (likely downward) biased results for direct beneficiaries. Kafle, Winter-Nelson, and Goldsmith (2016) uses the same data and similar methodology to find the program increases livestock revenue and food expenditures and decreases subjective feelings of poverty. Phadera et al. (2019) leverages the panel nature of the Zambia data to investigate program impacts on resilience; direct beneficiaries were less likely to fall into poverty and more likely to be food secure. These latter two studies treat PIF households as if they were treated rather than in the control group (composed of "eligible" households in two other communities), and find some attenuated effects on PIF households. However, the analysis is conducted as if treatment was assigned at the individual level, raising some concern about statistical inference. It should also be noted that all three Zambia studies are based on 105 treated households receiving various livestock transfers (31 received a dairy cow, 20 received draft animals, and 54 received goats).

In Rwanda, Rawlins et al. (2014) used cross sectional observational data to investigate the impact of a Heifer program on child anthropometrics and consumption of animal source foods. They evaluate both a dairy cow transfer (valued at 3000 USD) and a meat goat transfer (of unspecified value). Using regression controlling on observables including program eligibility, they find the dairy cow transfer increased milk consumption and the goat transfer increased meat consumption. With both regression and propensity score matching, they find

[2]While we see improvements in women's empowerment as a result of this study, we cannot estimate how much of this improvement can be attributed to increased financial inclusion.

evidence of impacts of both transfer types on child anthropometrics. This study was also based on a relatively small sample, with 155 households receiving livestock (78 received a cow and 77 received a goat) and a similarly sized comparison group.

The SLVC in Nepal has been previously studied as well. Miller et al. (2014) and Darrouzet-Nardi et al. (2016) use a matched pair randomized control trial comparing three communities assigned to the SLVC program to three controls. Miller et al. (2014) finds significant improvement in terms of height for age but not weight for age or days of school missed. Darrouzet-Nardi et al. (2016) find no overall effect on household dietary diversity or consumption of animal sourced foods, but do find strong effects in the Middle Hills region (which consisted of one of the three matched pairs). In both SLVC studies, the authors did not account for the clustered level of the intervention, casting doubt on statistical inference. These two studies were based on 201 and 181 households receiving (the same) treatment, respectively.

Compared to the above studies, this paper is the most rigorous and comprehensive evaluation of a Heifer livelihoods program to date. The program evaluated here also shares several similarities with the graduation programs featured in the more rigorous and wide-ranging evaluations conducted by Banerjee et al. (2015) and Bandiera et al. (2017), but is distinct in several important ways. Like graduation programs, the SLVC bundles a one-time productive asset transfer with technical skills training on management of the transferred asset, some basic health and/or life skills training, and access to financial services. However, overall graduation program costs are significantly larger: the per-beneficiary cost across the six interventions in Banerjee et al. (2015) is up to thirteen times larger than the standard SLVC program cost per *direct* beneficiary. The value of the transferred assets are much lower in the SLVC, which explains much of the difference in cost. In addition, the SLVC does not conduct frequent follow-up home visits from program officers. Instead, Heifer relies on self-help groups (SHGs) to ensure beneficiaries internalize and implement the skills acquired through the intervention. While beneficiaries do receive small amounts of cash for specific purposes like building improved goat shelters and fodder/forage production, they do not receive a regular cash stipend for consumption support (graduation program recipients receive cash transfers for approximately one year). Finally, graduation program beneficiaries are not expected to pass on benefits, and existing impact studies indicate no evidence of spillover effects.

3 Experimental design

To establish a causal relationship between the program and changes in outcomes, this study uses a cluster RCT. The RCT and related pre-analysis plan are registered with the American Economic Association (Janzen et al., 2018a). A cluster design was employed for two reasons. First, by design Heifer's intervention targets groups rather than individuals. Thus, a non-cluster design would be biased downward if groups successfully facilitate development. Second, local programatic spillover effects are an integral part of the intervention though the PIF mechanism. As described below, we will seek to estimate both direct and PIF effects.

Nepal comprises 75 districts. When this study began, districts were subdivided into Village Development Committees (VDCs). As of March 10, 2017, the VDC was dissolved and replaced by the *gaunpalika*, or rural municipality. In this paper we will continue to use the VDC, as this was the geographical unit our design uses. A VDC can be thought of as groupings of villages within a district. Every VDC is split into nine wards, and each ward contains multiple *toles*, or neighborhoods. A typical *tole* in the study area has approximately twenty to thirty households; a typical ward has roughly 150 households.

Nepal-based Heifer 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, Heifer also identified a central ward and targeted *tole* within that 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 beneficiary. Through this process, Heifer pre-identified all targeted direct beneficiaries (but not necessarily actual beneficiaries). Following treatment assignment, Heifer put these targeted direct beneficiaries into SHGs 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.

Direct beneficiaries are encouraged to pay benefits forward to individuals residing outside the *tole* but within the targeted ward. Although indirect PIF effects are anticipated within each central ward, and may spillover beyond ward-level administrative boundaries, we do not anticipate contamination of control VDCs. 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 sixty percent of our sample), where communities are isolated by low population density, rugged terrain, poor roads, and inferior cellular connectivity. 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. Apart from naturally occurring geographic and social barriers to contamination, we also buffered treated wards from each other and from control VDCs by selecting

a 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 across VDCs 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 regional clusters containing 15, 15, 10, and 20 VDCs. Using administrative data, we ordered VDCs by caste/ethnic compositions to further subdivide regional clusters into strata bins of four VDC when possible.[3] We then randomly assigned treatment within bins.[4] Within each stratification bin, we randomly assigned the 60 VDCs to one of three treatment arms or pure control.

All three treatments share some common features. First, Heifer facilitates the formation of women's SHGs, so all beneficiaries are expected to acquire some level of social capital through group membership and participation. Group members are then encouraged to contribute to group savings accounts with a goal of increasing financial inclusion. All beneficiaries are trained on a variety of technical topics including nutrition, home gardening, fodder and forage production, and improved animal management. In addition, all beneficiaries are provided a small amount of cash support for home gardens (approximately 5 USD), fodder/forage production (approximately 10 USD), and goat shed improvement (approximately 40 USD). Finally, all treatment VDCs receive access to a community animal health worker. We refer to these common features as the basic intervention.

In order to unpack the benefits of various program components, two additional programmatic elements vary across treatment arms: a productive asset transfer and additional values-based trainings. The productive asset transfer included two doe goats to each individual beneficiary, as well as a shared buck of improved breeding stock for the self-help group. The values-based trainings cover accountability; sharing and caring; sustainability and self-reliance; improving the environment; income; genuine need and justice; gender and focus on the family[5]; full participation; training, education, and communication; and spirituality.[6] Perhaps most importantly, the values-based training encourages beneficiaries to pay benefits forward by providing technical training and giving the first two female offspring

[3]We first calculated the proportion of residents in each VDC from each of 38 caste/ethnic groups. Then, within each regional cluster 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. Finally, we ordered VDCs within regional clusters based on rank prevalence of caste/ethnicity.

[4]Eleven bins had four VDCs, two bins had three VDCs and one had two VDCs. To ensure 15 VDCs in each treatment arm and the control arm we randomly re-allocated treatment in two VDCs.

[5]Notably, both men and women are encouraged to participate in gender and justice training.

[6]These trainings combined with the improved animal management and nutrition trainings (which are part of the basic intervention in our study) cover all Heifer's signature "Cornerstones" described in DeVries (2011).

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 livestock.
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. Table 1 summarizes the elements of each treatment arm.

Figure 1 illustrates a timeline of relevant programatic activities and events. Project implementation began in mid to late 2014 (depending on location). All direct beneficiaries first formed SHGs (shortly after the baseline survey, as described below) and were encouraged to begin saving at this time. They also began training and built improved livestock shelters. Approximately six months later, between March and June 2015, these same direct beneficiaries received livestock if they were assigned to either the FT or NVT treatments. In late 2015 the second generation of beneficiaries, recruited by direct beneficiaries in their area, entered the program through the PIF mechanism, began to form groups, and participated in the various trainings. Notice that while we know when program activities for these beneficiaries began, it is difficult to know exactly when second generation PIF beneficiaries received livestock transfers because such transfers depend on livestock fertility, which is inherently random. In fact, the program is designed in such a way that PIF livestock transfers will be staggered, with some receiving livestock transfers within six months of enrolling in the program, while others will wait years before receiving a livestock transfer.[7]

For establishing hypotheses regarding mechanisms and the anticipated timing of impacts, we must carefully consider livestock fertility cycles. We assume a doe can reasonably 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 ten months (males). Depending on breeding cycles and the availability of an improved buck, most direct beneficiaries would have been expected to impregnate their does between June and October of 2015, implying the members of a second generation of program goats were typically born near the end of that year and the beginning of 2016. Goats normally experience single births (although multiples are not uncommon), and the gender of the kid has important implications for impact. The program requires beneficiaries to donate their first two female offspring (once they have reached sexual maturity) to another

[7]SHGs figure out among themselves who will receive goats when among PIF beneficiaries based on their own criteria.

beneficiary through the PIF mechanism. Male kids are sold for meat and not passed on to PIF beneficiaries.

Taken together these facts imply three noteworthy features of this study, all shown in Figure 1: (1) the earliest PIF beneficiaries could possibly have received goats was in mid 2016, (2) the earliest possible goat sales (of transferred goat kids) for direct beneficiaries would have taken place after in late 2016 (after the data analyzed in Janzen et al. (2018b), and (3) the earliest possible goat sales (of transferred goat kids) for second generation PIF beneficiaries would have taken place in early 2018, shortly before the 3.5 year data was collected. The intervention concluded in mid 2017 after the 2.5 year data was collected. At this time, official Heifer program activities and monitoring ceased. These features are important for understanding mechanisms and impacts.

4 Sample and data

4.1 Sample description

We collected baseline data from rural women eligible to participate in the program across three regions of rural Nepal in June-September 2014. Midline data was collected in June-July 2016, approximately 1.5 years after initial enrollment in the program and analyzed in Janzen et al. (2018b). This paper evaluates program impacts 2.5-3.5 years after the program began, utilizing data collected in mid-2017 (which we refer to as "endline 1") and mid-2018 ("endline 2").

There are two types of respondents in the sample: targeted direct beneficiaries (in the central ward), and prospective PIF beneficiaries (also in the central ward). 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 originally selected a random sample of 15 potential PIF beneficiaries from a complete roster of all households in the same ward. Because of the aggressive nature of the PIF model, we expect that many (if not most) of these households had the opportunity to enter the program. 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* (that were subsequently not treated), and 15 individuals from a complete roster of all other households in the central ward.

Our total baseline sample used in the impact analysis is 2,376 women, including 1,286 targeted for direct treatment, and 1,089 households from the central ward likely to enter the program through the PIF mechanism (i.e. PIF households). Shortly after Heifer delivered

training and livestock to the direct beneficiaries of the project, a devastating earthquake struck Nepal. The April 2015 Gorkha earthquake greatly affected the 10 VDCs belonging to the Middle Hills stratification pool and were therefore spread evenly across treatment groups and control. At that time, we made the decision to drop these VDCs from the RCT so that Heifer could provide earthquake relief in whatever manner they deemed appropriate. Following additional attrition not explicitly related to the earthquake, the remaining midline sample consisted of 50 VDCs and 1,828 households, including 1,031 from targeted *toles* and 797 PIF households from the central ward more broadly.

4.2 Defining outcomes

We consider two categories of outcome variables: livestock livelihoods and household welfare. Each category is described in the subsection below.

4.2.1 Livestock livelihood outcomes

Since we are evaluating a livestock-based livelihoods intervention, our first category of outcomes relates to livestock livelihoods. This set of outcomes includes goat herd size and herd dynamics (e.g. births and deaths), and income, investment and profit from goat production activities. To assess whether the intervention affects the ways in which livestock are managed, we employ binary indicator variables for a variety of improved goat production practices: using an improved animal shelter, goat manure removal once per week, goat manure removal at least once per month, goat manure used for fertilizer, used medicine, used vaccinations, harvested home grown fodder, used mineral block, received an animal health worker visit at home, and access to community animal health worker in their area. Finally, the intervention targets women, so we also consider women's empowerment over livestock production. Specifically, we consider women's ownership of goats, and control over decisions regarding the care and maintenance of goats, sales and renting of goats, and decisions over livestock income.

We will employ summary indices for women's empowerment in livestock and improved goat production practices. Summary indices offer several benefits. First, they allow us to draw tractable conclusions about the program reaching broad objectives. Second, their use can also increase power by aggregating a number of impacts that are not statistically significant but move in the same direction, resulting in a statistically significant impact on index. Third, they can also reduce dimensionality to mitigate problems arising from multiple hypothesis testing. The summary indices we employ are calculated as standardized inverse covariance weighted (ICW) averages of subindicators following Anderson (2008).

4.2.2 Household welfare indicators

Our second category of outcomes relates to household welfare. We consider five primary welfare outcomes of interest: total household income, assets, women's empowerment, financial inclusion, and psychological well-being.[8] Total income is straightforward to calculate. For the remaining welfare indicators we follow the same approach described above, and calculate a standardized ICW average of subindicators following Anderson (2008). Each of these dimensions consists of multiple subindicators described and summarized below.

Although our data allows calculating the value of productive and non-productive assets, we do not have value of land owned or livestock. Because land and livestock are important productive assets in this context, our summary index uses the value of productive assets,[9] the value of non-productive assets[10], land, and livestock. Livestock represents owned goats, cattle, water buffalo, swine and chickens aggregated into tropical livestock units (TLUs)[11]. Land is measured by total hectares owned.

[8]We originally proposed to evaluate the impact on food security and nutrition using a summary index composed of a dummy variable equal to one if the respondent indicates all household members get enough to eat every day, a dummy variable equal to one if the household cut back on meals following a shock, the number of days the household consumed meat, fish or eggs, and a food consumption score (FCS). More than 95% of respondents are classified as food secure using the first two indicators. With such limited variation, we are unable to detect an effect. Our registered pre-analysis plan (Janzen et al., 2018a) suggest those subindicators should not be included in the analysis. Since meat, fish and egg consumption is already component of the FCS, it seems inappropriate to construct a summary of index combining the two remaining proposed subindicators (doing so simply alters the weights). We are thus left with FCS, however, we observe concerning levels of imbalance (all treatment arms have lower FCS than control households at baseline). Even after controlling for the outcome at baseline, such imbalance could lead to bias under some circumstances. Given these challenges, and lacking the added robustness offered by using the summary index in correcting for multiple hypotheses, we choose to omit this welfare indicator from our results.

[9]Productive assets include draft animals (excluding water buffalo), plows, computers, printers, grinders, looms, sewing inventories, mechanical tools, hand tools, bicycles, motorcycles/scooters, solar panels, batteries, inverters, and improved livestock pens. Although pre-specified, we omit cars and other vehicles, tractors and threshers which are uncommon and valuable, thereby potentially creating outliers that lead to bias. We don't expect the intervention to increase any of these assets, so while their omission makes the asset list incomplete, it should not bias the impact estimates for the subset of variables that remain in the productive asset list. The baseline survey does not accommodate calculating the value of productive or non-productive assets. The baseline survey also varies in the types of assets. In lieu of value, we sum dummy variables indicating ownership of productive assets (including plows, grinders, threshers, sewing inventories, mechanical tools, tractors, motorbikes, bicycles, cars, computers and improved livestock pens). We note that principle components analysis led to unintuitive signs for weights.

[10]Non-productive assets include phones, radios, cassette recorders, DVD players, televisions, satellite dishes, cameras, camcorders, electric fans, heaters, refrigerator/freezers, gas stoves, cupboards, jewelry, watches, tables, chairs, sofas, and mattresses. The baseline survey does not accommodate calculating the value of non-productive assets and varies in the types of assets. Instead of using value, the non-productive asset index is a principle components index of dummy variables indicating ownership of non-productive assets (including radios, televisions, mobile phones, heater/pressure lamps, electric fans, camera/camcorders, furniture, irons, jewelry and watches).

[11]We follow the FAO's guide (FAO, 2005) to calculating TLUs in Nepal: cattle = 0.5, buffalo = 0.5, sheep & goats = 0.1, pigs = 0.2, chicken = 0.01.

Our summary index of women's empowerment utilizes subindicators (modified to the local context) from the Five Domains of Empowerment (5DE) of the Abbreviated Women's Empowerment in Agriculture Index (A-WEAI) to calculate an empowerment score for all women in the sample (Alkire et al., 2013; Malapit et al., 2015). The A-WEAI was developed based on pilot surveys conducted in three countries through extensive collaboration between the United States Agency for International Development, the International Food Policy Research Institute and Oxford Poverty and Human Development Initiative. The A-WEAI aggregates an empowerment score across decisions about production, ownership of productive assets, access to and decisions on credit, control over income, group membership & leadership, and workload. Each binary subindicator equals one if the respondent achieves adequate empowerment in that area, and zero otherwise. Although we employ subindicators based on the A-WEAI, in a deviation from the A-WEAI, we calculate the ICW average following Anderson (2008) rather than using standard weights defined by the A-WEAI. We do this for consistency across indices and to better leverage components of the index where there is greater variation among respondents.

Definitions of adequacy were pre-specified and adjusted to the local context. Using the A-WEAI as our guide, we set out to use the following definitions of adequacy: A respondent is adequately empowered in production decisions if she has at least some input into at least one production decision (types of production decisions include staple grain farming, high value crop farming, large livestock raising, small livestock raising, poultry raising, and non-farm economic activities). Adequate ownership means the household owns at least one productive asset (as defined above), livestock or land, and the respondent (individually) has at least some ownership of an asset, livestock or land. A respondent is adequately empowered in access to and control over credit if the household has at least some credit and the respondent participated to any extent in the decision to borrow. Adequacy in control over income means the respondent participates in at least one decision regarding non-food expenditures. A respondent is adequately empowered in group membership if she is a member of any group. A respondent is adequately empowered in leadership if she holds a leadership position in any group. A respondent is adequately empowered for time use if she worked 10.5 hours or less (not including domestic work or cooking) in the previous 24 hours.

Using these original definitions of adequacy, we observe greater than 95% of control women are adequately empowered in four of the seven subindicators - which suggests surprisingly high levels of empowerment using the A-WEAI subindicators. Using the pre-analysis plan as our guide (Janzen et al., 2018a), we omit those subindicators for which at least 95% of control observations have the same value, and construct the summary index using only

three inputs: access to and decisions on credit, control over income, group membership & leadership. As a robustness check, we will consider alternative definitions of adequacy in the appendix.

To measure program impacts on financial inclusion our summary index employs the following outcomes: membership in a village credit and savings group (Heifer or otherwise), whether and how much the household saved in the past month, whether the household has any savings, total amount the household has saved, whether the respondent personally has any savings, total amount she has saved, whether the household has an outstanding formal loan, whether the household has an outstanding informal loan, and the percent of current debt that is from a formal loan. We consider banks, credit and savings groups, cooperatives, and microfinance organizations to be formal lenders and finance corporations (essentially payday lenders in this context), friends and family, village money lenders, and local retailers offering goods on credit to be informal lenders. Table 12 shows the percent of households with outstanding debt overall, owed to formal and informal lenders, and to different lender types at baseline. One-third of households have debt to any lender, and the vast majority of them have debt to only one lender. Eleven percent have debt to any formal lender and 23 percent have debt to any informal lender. Most households with debt to a formal lender borrowed from a village credit and savings group. Loans from banks are less common, but much larger on average. Most households with debt to an informal lender borrowed from friends and family or a local shop for consumption credit. Formal and informal debts were of roughly equal size, conditional on having that kind of loan.

Financial inclusion reflects access to useful and affordable financial products and services. Analyzing financial inclusion is complicated because what constitutes "useful and affordable" products is difficult to define – in other words, debt can be both good and bad. The ICW summary index increases efficiency by ensuring highly correlated outcomes receive less weight than outcomes that are uncorrelated, but it nonetheless requires a researcher to assign value to whether an indicator should enter the index positively or negatively. In our construction of the financial inclusion index, we assume all of the savings variables enter the index positively. Having formal loans enters the index positively, as does the percent of debt from formal lenders, whereas having informal loans enters negatively. If a household has no debt, we consider the percent of debt from formal lenders to be zero.[12]

To measure program impacts on psychological well-being we consider the following

[12]Here we deviate from the pre-specified index by replacing the amount of formal debt and the amount of informal debt with the percent of debt from formal sources. This is because while access to formal credit is a good thing, too much formal debt could certainly be bad. Given some amount of debt, the percentage of that debt to formal lenders is more unambiguously a positive indicator of financial inclusion. We also deviate by including variables for the respondent's personal savings in addition to household savings. The treatment effects on the financial inclusion index are not very sensitive to these deviations.

outcomes: depression, worrying, self-esteem, life satisfaction, optimism, and locus of control (the extent to which one has control over one's outcomes). Depression (inverse) is based on nine questions from an abbreviated version of the CES-D scale Radloff (1977) with a high value indicating *low* levels of depression. Worry (inverse) employs nine questions from the Penn State worries questionnaire, and a high value indicates *less* worried. Self-esteem is based on eight questions from Rosenberg (1965). Life satisfaction is based on one question from the World Values Survey (worldvaluessurvey.org). Optimism is based on six questions from Scheier, Carver, and Bridges (1994). Locus of control is an abbreviated Rotter (1966) scale based on 19 questions where a high value indicates a stronger internal locus of control.[13] With these variables we construct a summary index where depression and worrying enter negatively and all other variables enter positively.

4.3 Descriptive statistics and balance

Tables 2 and 3 contains descriptive statistics for both direct (top panel) and PIF (bottom panel) beneficiaries in the sample at baseline, where column 1 contains mean values for all treatment groups in a given subsample. The average direct beneficiary is 40 years old with 2.6 years of education. The average direct beneficiary husband is 44 years old with 5.1 years of education. The most educated member of the household is typically a son or daughter, with 9.0 years of education. Average household size is 5.8, and 12 percent of beneficiaries are unmarried. Potential PIF households are very similar along these dimensions, although slightly less educated.

Although not targeted for being poor, global standards suggest most households in the sample are extremely poor. Mean reported annual household income among both direct and potential PIF beneficiaries is 99,537 NPR or 995 USD.[14] Note that this includes remittances, which represents 44% of income for the average household in our rural Nepal sample at endline 1.[15] On average, direct beneficiaries had four goats, 2.5 total livestock units (TLU, which includes goats) and 0.46 hectares of land. Potential PIF beneficiaries had 2.4 TLU,

[13]The baseline psychological well-being index is a weighted standardized average of the same subindicators, but many of the subindicators are calculated differently due to survey changes between baseline and endline. At baseline, depression (inverse) was calculated using four questions from an abbreviated version of the CESD scale Radloff (1977) with a high value indicating *low* levels of depression. Worry (inverse) was calculated using four questions from the Penn State worries questionnaire, and a high value indicates *less* worried. Self-esteem was calculated using six questions from Rosenberg (1965). Optimism was calculated using four questions from Scheier, Carver, and Bridges (1994). Locus of control was calculated using an abbreviated Rotter (1966) scale based on six questions where a high value indicates a stronger internal locus of control. Life satisfaction was calculated in the same way at baseline as at endline 1 and endline 2.

[14]For convenience, we use the simple rule of thumb exchange rate of 0.001, which tracks the average exchange rate throughout the life of the study with reasonable accuracy.

[15]Data on remittances at baseline cannot be disaggregated from other sources of income.

3.9 goats, and 0.46 hectares of land. So while these are poor households, most (91 percent) had some livestock and 71 percent had some goats before entering the Heifer program.

We are confident treatment was randomly assigned, as there was no possibility for Heifer to re-assign treatment after our randomization. However, imbalance by chance is a distinct possibility. To test for balance across treatments we regress $y_{hv}^{t=0}$, a demographic characteristic or outcome for household h residing in VDC v as measured at baseline ($t = 0$), onto treatment status. Specifically, we estimate the regression below separately for the subsamples of direct and PIF beneficiaries:

$$y_{hv}^{t=0} = \beta_0 + \beta_1 T_{hv}^{FT} + \beta_2 T_{hv}^{NG} + \beta_3 T_{hv}^{NVT} + \varepsilon_{hv}. \quad (1)$$

In equation T^{FT}, T^{NG}, and T^{NVT} are dummy variables for a household assigned to receive the "full treatment" package, the "no goats" package, and "no values-based training" package, respectively, and ε_{hv} is an idiosyncratic error term clustered at the VDC level.

Tables 2 and 3 report the results of the balance tests for direct and PIF beneficiaries, respectively. In both tables, columns 2-4 contain baseline differences between the three treatment groups and the control, and columns 5-6 contain baseline differences between treatment groups.

4.4 Heifer self-help group membership

For the purposes of estimating participation rates, we use each respondent's stated, self-reported membership in a Heifer group. We accept self-reported membership for several reasons, even though Heifer rosters do exist that permit crosschecks.[16] First, SLVC is a community level intervention, intended to be inclusive and ultimately aimed at enrolling all members of a community. It is plausible that unlisted individuals attend meetings, interact with groups, consider themselves Heifer members, and most importantly derive benefits from their association or from social learning. Second, the Heifer lists themselves aren't perfectly suited to an authoritative cross-reference. It is possible that the lists do not reflect the full and most updated membership at the VDC level. Further, these lists are originally written in Devanagari script and then transliterated into the English alphabet, requiring imperfect fuzzy matching between two sets of records that may fail to detect true matches. Third, it seems unlikely that completely unaffiliated people self-identify as members. If this were the case we would see higher rates of self-reported membership in control areas than the near-zero levels we actually observe. Finally, since we report intent-to-treat (ITT) effects,

[16]We report the results of crosschecks between the evaluation sample and Heifer rosters in appendix Appendix A.

our definition of the participation rate has no bearing on estimated treatment effects.

Figure 2 presents the proportion of each treatment type claiming membership in a Heifer group at the first endline, and figure 3 is the analogous figure for the second endline. We also report stated membership rates in control villages, which we assume are reported incorrectly (Heifer does not operate any projects in control areas and clusters are separated by expansive geography, but there are many other NGO-operated groups one could belong to). Thus recruitment rates in treatment villages can be calculated as the recruitment rate in a treatment group less the corresponding recruitment rate (direct or PIF, endline 1 or 2) in the control group.

Recall that direct beneficiaries were targeted by Heifer based on geography (*tole*) within the ward, and potential PIF beneficiaries are other randomly selected individuals in the ward. Among targeted direct beneficiaries, 93 percent of households offered the FT reported belonging to a Heifer group at endline 1, and the vast majority of these identified themselves as direct . Among directly targeted NG beneficiaries, 85 percent reported being members at endline 1 and 83 percent being members at endline 2. Membership rates were similar in the NVT treatment group, with 85 percent reporting being members at endline 1 and 80 percent at endline 2 (this drop is not statistically significant). In sum, recruitment and retention among direct beneficiaries is very high, particularly in the full treatment group (significantly higher than both other treatment groups in both years).

Compared to direct beneficiaries, membership rates were approximately 12-14 percentage points lower among potential PIF beneficiaries in the FT and NG treatment groups. Given that potential PIF beneficiaries are a random sample of all households in a treatment ward (excluding the direct beneficiaries), the PIF mechanism is very effective at building new SHGs and recruiting and retaining members.

In the NVT group, direct beneficiaries were not instructed to create new SHGs and recruit new members. Consequently, membership rates are much lower. At endline 1, 40 percent of potential PIF households claimed to be Heifer SHG members and at endline 2, 32 percent did. However, among these over 80 percent reported being direct beneficiaries. We conclude these potential PIF members managed to enter the original SHGs, either to replace intended beneficiaries who did not join, or managing to otherwise integrate themselves into the group.

Assuming that wards contain an average of 125 households, and that 25 households in the central ward are direct beneficiaries, on average 83 additional households from the central ward are brought into the program through the PIF mechanism in the FT group, 67 additional households in the NG group, and 40 in the NV group. We can add this to the number of direct beneficiary households in each treatment (23 in the FT group and 20

in the NG and NV groups) to reach 85 percent membership in the FT group, 70 percent membership in the NG group, and 48 percent in the NVT group. While goats are clearly important for recruitment, the values-based training, which includes the requirement to pay benefits forward, are what drives the spread of the program within a community.

5 Impact estimation approach

To analyze the welfare impacts of a productive asset transfer and training program, we estimate the intent to treat (ITT) effects for each of the three treatment groups relative to a common control. ITT will be conservative compared to average treatment effect on the treated (ATT) since take-up rates were not 100 percent. To analyze whether treatment effects reach subsequent generations of beneficiaries, we estimate these effects separately for two subsamples: direct and PIF beneficiaries. In the analysis of direct treatment effects, the sample consists of those pre-selected for direct benefits (including those in control areas). In the analysis of PIF treatment effects, the sample consists of all other individuals in the central ward (also including those in control areas). We note that PIF effects could arise through technical training conducted by direct beneficiaries, values-based training conducted by Heifer staff, asset transfers from direct beneficiaries, or through observation and replication. If households observe and replicate the behavior of direct beneficiaries, they may benefit indirectly from trainings, even if they do not identify as a second generation program beneficiary.[17] By estimating ITT effects, we can capture benefits coming through all of these channels.

We estimate the following equation at time t separately for the direct and PIF beneficiary subpopulations:

$$y_{hv}^{t} = \beta_0 + \beta_1 T_{hv}^{FT} + \beta_2 T_{hv}^{NG} + \beta_3 T_{hv}^{NVT} + \delta y_{hv}^{t=0} + \boldsymbol{X_{hv}'}\boldsymbol{\gamma} + \boldsymbol{S_b'}\boldsymbol{\rho} + \varepsilon_{hv} \tag{2}$$

In equation 2, y_{hv}^{t} is the outcome of interest for household h in ward v, measured at time t. T_{hv}^{FT}, T_{hv}^{NG}, and T_{hv}^{NVT} are indicator variables for a household being in a VDC selected to receive the corresponding treatment. Control variables include the outcome of interest measured at baseline ($y_{hv}^{t=0}$), a vector of normalized covariates measured at baseline ($\boldsymbol{X_{hv}}$), and strata bin dummies ($\boldsymbol{S_b}$). For each subsample (direct or PIF) used in the estimation, β_1 represents the ITT effect on households in VDCs where direct beneficiaries were selected to receive the full treatment package (FT), β_2 identifies the same for the no-goats package (NG), and β_3 identifies the same for the no-values-based training treatment package (NVT).

[17] This is why we do not estimate local average treatment effects (LATE).

The counterfactual is targeted (direct or PIF) beneficiaries located in control VDCs.

To maximize power we employ machine learning to select the covariates to be included in $\boldsymbol{X_{hv}}$ for each regression. Specifically, we employ the post-double-selection lasso (Least Absolute Shrinkage and Selection Operator) estimator developed by Belloni, Chernozhukov, and Hansen (2014) and implemented using the PDSLASSO command in Stata (Ahrens, Hansen, and Schaffer, 2018). All candidate control variables (described below) are normalized, and missing values are replaced with group (direct or PIF) medians. Dummy variables for missing control variable values are also included as candidate control variables. As indicated in equation 2, the outcome at baseline (when available), and strata bin dummies will always be controlled for.

Candidate demographic controls include household size; respondent and spouse age, literacy, and years of schooling; a dummy for no spouse; and the maximum years of schooling of the most educated household member. We also include all seven welfare indices and a subset of pre-selected subindices and mechanism variables measured at baseline as candidate control variables. The 13 subindices are land area, a productive assets index, total livestock herd size (TLU), total livestock income, total livestock investment, goat herd size, goat revenue, goat production practice index, decision-making over goats index (see section 4.2 for defining the latter two indices), income aspirations, social status aspirations, an ordered categorical "patience" variable, and an ordered categorical "planning horizon" variable. Household size is also always included in the amelioration set because we do not calculate per capita values.

6 Goat livelihood impacts

Our first research question is whether the program impacts livestock livelihoods. Are households building their goat herds and generating income? Do direct beneficiaries give away goats? Are PIF beneficiaries receiving goats? Do livestock management practices improve? Do women have more control over livestock-related business decisions? In this section we answer these questions.

6.1 Goat herd dynamics

We first look at goat herd dynamics. The results are presented in table 4. The first four columns present 2.5 year impacts alongside the control mean. The latter four columns do the same for 3.5 year impacts. The top panel shows impacts for direct beneficiaries, while the latter presents PIF impacts. All subsequent results tables follow the same format.

The results are strong: the intervention increased direct beneficiary herds by 2.7 goats at endline 1, with similar impacts observed at endline 2. Respondents were given 2 goats through the intervention,[18] but the increase in livestock herds is even more notable given the fact that we also observe direct beneficiary households giving away on average 1.3 goats in the year prior to endline 1 and 0.3 goats in the year prior to endline 2. In other words, households were given 2 goats, gave away 1.6 goats over two years, and yet still have herds increase by more than what they were originally given. As herds grow, we naturally see more goat deaths. However, growth in herds – beyond the size of the asset transfer net paying it forward – is achieved through 1.7 more births at endline 1, an effect that appears to weaken by endline 2 (1.2 births) but is still statistically significant. In order for this to be meaningful as a livelihood, it's also essential that goats are being sold. Our results suggest FT direct beneficiaries sell on average 0.5 more goats than the pure control at endline 1, and nearly double that (0.9) at endline 2. We will explore the implications of these increased goat sales for income in subsection 6.2. We also observe small statistically significant increases in goats slaughtered for home consumption (0.1), which could have downstream effects on food security and nutrition, but it's unlikely the magnitude of the effect is large enough to be observed downstream.

Comparing across treatment arms, herd size increases are similar in both treatment arms that received goats. However, this comparison also clouds the fact that FT households gave away an average of 1.6 goats over two years, whereas NVT households did not. The difference is not met by greater sales, and if anything there seems to be higher mortality in the NVT group. If FT households had not redistributed some of their benefits, then conceivably the impacts would have been even larger than the NVT impact, which may imply there is impact from the values training beyond the PIF component, perhaps through greater social cohesion and group dynamics. Herd size impacts are about half as large (1.4) but still statistically significant in the NG group. Although the NG group didn't receive the asset transfer, they also were not expected to pay forward an asset transfer (instead they were expected to pay forward all knowledge accumulated and other benefits of the program). Births and sales increase at both time periods for NG beneficiaries, but with the exception of endline 1 sales, the impacts are smaller than the impacts for FT direct beneficiaries.

We next turn to PIF beneficiaries. As noted, direct beneficiaries report giving away 1.3 goats at endline 1 and another 0.3 goats at endline 2. Do PIF households confirm receipt of the asset transfer? Yes. Recall that the ratio of potential PIF beneficiaries to direct beneficiaries is approximately 4:1. The average number of goats received among potential PIF beneficiaries in the FT group (the only ones that should have received a passed goat) is one-

[18]This is not evident in the estimate of goat gifts received from Heifer because the question only asks about the year prior, and the original asset transfer was received approximately two years prior to the survey.

fifth the number gifted by FT group direct beneficiaries, indicating transfers are occurring as intended. But that's only one small part of the impact on herd sizes, which increase by 1.1 goats at endline 1 and persists into endline 2. Most increases in herd size among potential PIFs arise from increased births, which also explains why impacts are similar between the FT and NG treatment arms. Sales also increase in these two treatment arms among potential PIF beneficiaries. Perhaps most surprising is the observed PIF impact on goat herd size for the NVT treatment, where paying it forward was not explicitly encouraged. We offer two potential explanations for this surprising impact: the impact could be picking up unintentional spillovers from neighboring direct beneficiaries, or more likely it's related to the compliance issues identified in section 4.4 where some individuals in our sample of PIF beneficiaries may actually have joined the program as a direct beneficiary. This finding also suggests using caution in interpreting PIF impacts in other treatments, which may similarly be upward biased due to compliance issues.

6.2 Goat production practices

The next set of results highlight the ways in which the intervention affects goat production practices. The first indicator presented in table 5 shows the impact on our summary index in this area, and suggests a very large 0.77 standard deviation impact for FT beneficiaries, a 0.46 standard deviation impact for NG beneficiaries and 0.66 standard deviation impact for NVT at endline 1. Smaller effects in the latter treatment arms could be a reflection of lower recruitment rates, a less effective program, or noise. In all treatments, results are consistent between endline 1 and 2, with only slight declines.

We can unpack what this impact means by looking at the impact for each production practice used to calculate the summary index. Relative to the control at endline 1, FT direct beneficiaries are 51 percentage points more likely to have an improved pen, 37 percentage points more likely to remove manure at least weekly, 32 percentage points more likely to remove manure at least monthly, 20 percentage points more likely to use manure as fertilizer, 21 percentage points more likely to use medicine when livestock are sick, 20 percentage points more likely to vaccinate goats, 14 percentage points more likely to use home fodder, 37 percentage points more likely to have a community animal health worker visit their home, and 36 percentage points more likely to know a community animal health worker in the village. We observe no evidence the intervention increases use of mineral blocks as a nutrient-dense animal feed supplement. For the most part, these changes persist at endline 2 (if anything, the frequency of manure removal seems to decline). Looking across treatments, we observe somewhat smaller but still strong impacts for both NG and NVT direct impacts.

The values-based training encourages direct beneficiaries to not only pay forward physical benefits, but it also encourages teaching new PIF recruits about improved production practices that have been learned through the technical trainings of the program. In this way, pay it forward can have an effect on goat production practices. The results presented in the second panel of table 5 suggest that it does. The summary index increases by 0.43 standard deviations for FT beneficiaries and 0.38 standard deviations for NG beneficiaries. For each practice we observe smaller, but mostly still significant effects. These smaller effects should be interpreted with the recognition that these are ITT impacts, and recruitment rates are lower in PIF relative to direct for all treatments, but also lower for NG relative to FT (for both direct and PIF). No PIF impact is observed for NVT at endline 1 or 2, but this is not surprising since the cost of training one's neighbors, in the absence of encouragement and social pressure, is likely quite high.

6.3 Goat profit, gross revenue and investment

Table 6 demonstrates the program impact on profit, gross revenue and investment related to goat production activities. The results suggest positive gross revenue increases (around 3000 NPR, or 30 USD). The market price of a goat is approximately 7500 NPR (75 USD), so the revenue impact roughly reflects the value of selling 0.4 goats on average (the observed direct FT impact for goat sales presented in table 4. The revenue impact for FT direct beneficiaries doubles by endline 2. The NG gross revenue effect is twice as large as the FT effect at endline 1, which seems to reflect their ability to reap benefits more quickly, rather than redistributing economic gains in the short run (NG beneficiaries reported selling twice as many goats at both endline 1 and endline 2). Revenue impacts for NVT direct beneficiaries are similar to FT direct beneficiaries at endline 1, but unlike FT beneficiaries who see impacts strengthened over time, NVT revenue impacts are consistent across endline 1 and 2.

Although revenue impacts are encouraging, section 6.2 revealed important impacts on production practices, many of which entail costs. Not surprisingly, we find the treatment does indeed increase the amount invested in fodder, veterinary services, and breeding.[19] Investment increases most in the FT, followed by NVT, followed by NG, a pattern that persists at endline 2. As a result, the net revenue effect is zero for FT direct beneficiaries at endline 1, but positive and significant (around 3800 NPR) by endline 2. NG beneficiaries reap the benefits sooner and do not increase investment as dramatically, so they see profit

[19]Goat shed improvement is also encouraged through the program, but improved goat sheds constitute a large capital investment primarily subsidized by Heifer, and thus are omitted here and included as part of program costs discussed in section 9.

impacts of approximately 3600 NPR at endline 1, increasing to around 5500 NPR by endline 2. These findings suggest that by endline 2, direct beneficiaries in FT and NG treatment arms increase their net revenue by 70-100% (the control mean at endline 2 is 5300 NPR). Although we continue to observe increased gross income and investment for NVT direct beneficiaries, it does not translate into a profit effect, providing another signal that the values-training may do more than simple encouragement to pay it forward.

Do profit, gross revenue and investment impacts carry over to PIF beneficiaries? At endline 1, FT PIF beneficiaries demonstrate small increases in income (not significant), and increases in investment that are approximately half as large as direct FT beneficiaries. Income effects swell by endline 2 leading to positive and statistically significant profit impacts (3200 NPR) by endline 2. NG PIF impacts are similar to NG direct impacts - impacts are observed more quickly than those observed by FT PIF beneficiaries. For direct beneficiaries differential early impacts between FT and NG can be explained by a differences in the required redistribution that stems from paying forward, but when comparing PIF direct and NG beneficiaries this no longer holds, and we are left with a puzzle as to why the impacts are observed more quickly for NG than FT. One statistical explanation is imbalance at baseline – NG households, both direct and PIF, have higher goat incomes at baseline. Although we control for the outcome at baseline, any imbalance could still bias our results if incomes are growing at differential rates (parallel trends does not hold), or if there is treatment heterogeneity related to the imbalanced variable (in section 8 we find little evidence of meaningful heterogeneity). Last, at endline 2 here again we observe a small (2,400 NPR) statistically significant but surprising profit impact for NVT PIF beneficiaries. Once again, this impact could be picking up unintentional spillovers from neighboring direct beneficiaries, although this seems unlikely since we do not see evidence of increased investment (which re-confirms the findings presented in section 6.2 where we observed no change in production practices). It could also be related to the compliance issues identified in section 4.4 where some individuals in our sample of PIF beneficiaries may actually have joined the program as a direct beneficiary.

6.4 Women's empowerment over goat production

In our last set of results regarding the program impact on goat livelihoods, table 7 presents our analysis of program impacts on women's empowerment over livestock production. The first row of table 7 presents standard deviation impacts of our women's empowerment in goat production summary index. At endline 1, empowerment in this area increases by 0.31-0.47 across all treatments, and the estimates persist at endline 2. Relative to the control at endline 1, FT direct beneficiaries were 23 percentage points more likely to say they were the

owners (or jointly owners) of the household goats, 22 percentage points more likely to have control over decisions regarding the care and maintenance of goats, 24 percentage points more likely have control over decisions regarding livestock sales, and 14 percentage points more likely to control the income earned from livestock.

Direct impacts are strongest for FT and NVT beneficiaries who received an asset transfer; in both of these treatment arms women were more likely than NG beneficiaries to say they were the owners (or jointly owners) of the household goats, and had control over decisions regarding goat care, maintenance, sales and livestock income. This result reverses when comparing FT and NG PIF beneficiaries at endline 1 (a finding that is a bit puzzling), but adjusts to reflect equal impacts (0.26-0.27 standard deviations) between FT and NG PIF beneficiaries at endline 2. We observe no impacts at the summary index level for NVT beneficiaries at either endline 1 or endline 2 and attribute the 1 out of 8 statistically significant subindicator impacts for NVT PIF beneficiaries to the multiple comparisons problem. In the absence of encouragement to pay it forward, we do not observe PIF impacts on women's empowerment over goat production.

7 Household welfare impacts

Taken together, the findings presented in section 6 provide strong evidence the program impacts livestock livelihoods. In this section we ask whether the observed goat livelihood impacts lead to measurable improvements in household welfare. Specifically, does total income increase as a result of the program? Does the household asset base expand? Are women more empowered? Do beneficiaries have greater financial inclusion? Does mental health improve? This section sets out to answer each of these questions.

7.1 Total income

We do not find robust evidence the Heifer program had a significant effect on total household income (Table 9). We do see a significant positive effect on total income for direct beneficiaries under the NG treatment at endline 1, a marginal significant negative effect on total income for PIF beneficiaries under the FT treatment at endline 1, and a marginal significant positive effect on PIF beneficiaries under the NVT at endline 2, but we believe these are from chance, and not due to the program. The biggest drivers of these associations between treatment and total household income are driven by large differences in income sources that should not be affected by the Heifer program (salary, remittances, and off-farm labor). Looking at the income source most directly tied to the Heifer program, we do not see statistically

robust evidence of an impact on livestock income. We also do not find statistically robust impacts on crop income or small business formation, which could be potentially impacted by Heifer's training.

In light of the observed impacts on goat revenue and profit, why do we not observe an impact on total household income? To better understand these null findings it's helpful to take a closer look at the income portfolio of sample households. Table 10 shows the percent of household earning money from an activity (column 1) and the average net revenue from that activity (column 2) among all households at endline 1 (the endline income data is more comprehensive than the baseline income data). Household income is 285,000 NPR, or approximately 2,850 USD. The average household size is 6.44, making this 1.20 USD per person, per day, which is considerably below the global poverty line. Nearly 60 percent of households receive migrant remittances (domestic or international), consistent with migration data from Nepal, and migration remittances constitute approximately 45 percent of household income. Ninety percent of farmers grow crops (about half of which sell them), and 76 percent produce livestock (almost all of who sell some). However, agriculture composes a very small portion of the income portfolio on average. Households produce 36,000 NPR worth of crops and 11,200 NPR worth of livestock. Livestock accounts for only approximately 4% of income. Goats are the most common source of livestock income, with 54 percent of households selling goats. Average goat revenue in the sample (which is three-quarters treatment and one-quarter control) is 6,200 NPR. That is approximately 62 USD, or two percent of household income. To cross check, we tested for program impacts on household expenditure and found no positive impact of the program (results available upon request).

While the Heifer program is effective at increasing income earned from goat enterprises, livestock income is simply not a substantial source of income for sample households. Given this, it is not surprising we do not see a statistically significant impact on household or even livestock income 2.5 and 3.5 years after the start of the program, especially considering the slow reproduction cycle of goats.

7.2 Household assets

We observe a positive 0.12 standard deviation impact on assets for FT beneficiaries at endline 1, and the impact doubles to 0.21 by endline 2. We observe no impact on assets for NG or NVT beneficiaries. The FT impact stems from increased herds, where the magnitude of the effect reflects the observed increase in goat herds rather than investments in other livestock. FT households also amass twice as much land as control households at endline 2 (the impact is half as large but not significant at endline 1). This may indicate that beneficiaries who

received goats are investing in land to grow fodder for their livestock.[20] We do not observe evidence of impacts on productive or non productive assets.

We do not detect statistically significant effects on the asset summary index for PIF beneficiaries. Herd size increases are not significant but the magnitude of the coefficient is similar in size to the estimated effects on PIF goat herds. We do observe some statistically significant negative coefficients for productive assets across multiple treatment arms and time periods, a result which is either puzzling or the result of a multiple comparisons problem.

7.3 Women's empowerment

We find positive impacts on women's empowerment at endline 1 for direct beneficiaries in all treatments. The impacts are strongest for FT direct beneficiaries (0.33 standard deviations) and weakest but still statistically significant for NG direct beneficiaries (0.20). Impacts persist into endline 2 only for FT beneficiaries. Impacts are driven by higher levels of group participation and leadership. 25% of control women were not participating in groups; the FT treatment effect more than halves this (a 16 percentage point increase). Relative to the control, 50% more FT women take on leadership positions. We observe no evidence of the intervention improving women's access to and control over credit.

PIF beneficiaries are also more likely to participate in groups, but the summary index results are considerably weaker (marginally significant 0.17 standard deviation impact for NG and not significant 0.14 standard deviation impact for FT PIF beneficiaries.) There are some imbalance issues that could be affecting access to and control over credit for PIF beneficiaries (all treatments appear worse off relative to the control), which could be driving the surprising negative credit control impacts for PIF beneficiaries.

We had originally planned on using a more comprehensive summary index using all seven subindicators of the women's empowerment in agriculture index. As noted in section 4.2.2, several of the proposed indicators demonstrated extremely high empowerment with limited variation. To more fully exploit this data, we "raise the bar" for empowerment as much as we reasonably can with the data we have.[21] The results are presented in appendix table A1. The summary index results are weaker, but consistent. We observe limited evidence of

[20] As early as our midline study, field staff had reported that this was the case. They also reported that households often purchase land with additional income generated from raising goats for other businesses. We note, however, that we do not observe impacts on total income.

[21] Specifically we make the following four adjustments: (1) we do not count staple grain farming or poultry raising for control over production decisions. (2) we do not include hand tools or simple cell phones as a productive asset (these two items are the most commonly owned of all productive assets) for asset ownership. (3) we calculate the percent of household expenditures over which the woman has control as our indicator of a woman's control over income (dropping common expenditures dwindles the list to something quite limited). (4) we include domestic activities and cooking as work hours for adequacy in workload.

impacts in these other dimensions of empowerment.

7.4 Household finance

Among direct beneficiaries, we find positive impacts of all three treatments on the household finance index at endline 1, and of the FT and NG treatments at endline 2. The impacts on the index are largely driven by strong impacts on savings group membership and whether the respondent personally has any savings. The only robust impacts we see on amount saved are for the personal savings of the respondent at endline 2 under the FT and NG treatments, but not the NVT treatment. Point estimates suggest a shift from informal to formal credit, although the results are not statistically significant. Among PIF beneficiaries we see positive impacts on the index at endline 1 of the FT and NG treatments, but not in the NVT treatment, as expected. These impacts are largely driven by strong impacts on savings group membership and whether the household saved anything last month or has any savings. We do not see robust impacts on amount saved or in savings. As for direct beneficiaries, we see some evidence of a shift from informal to formal credit. At endline 2, the impacts on the index dissipate for PIF beneficiaries. In the NVT treatment we see a marginally significant negative impact, although this is likely due to imbalance at baseline that we cannot fully control for. We also see marginally negative impacts (also likely due to imbalance) on household saving. However, we again see a positive impact on personal saving by the respondent that is significant under the FT but not the NG treatment. We observe a decrease in both informal and formal debt.

7.5 Psychological well-being

We see no impact on the psychological well-being index among direct or PIF beneficiaries at endline 1 or endline 2. Effects on subindicators can be seen sporadically across treatments and time periods. We see some evidence of decreased depression and worrying and increased self-esteem. Surprisingly, we also observe weak evidence of a negative impact on locus of control. On aggregate, these results indicate the intervention had no discernible effect on overall psychological well-being.

8 Impact heterogeneity

In this section we attempt to identify observable characteristics that predict which households are most likely to benefit from the program. Without strong priors about what character-

istics might drive impact heterogeneity, it is difficult to estimate heterogenous treatment effects without confronting problems associated with multiple hypothesis testing. Novel machine learning techniques can be useful in such instances because they allow us to test for statistically robust sources of heterogeneity among many possible variables.

We employ the method of Chernozhukov et al. (2018) to test for heterogeneous impact on our key goat livelihoods (goat herd size, goat income, improved goat production practice index, and women's empowerment over goat production index) and household welfare indicators (income, assets, womens empowerment, financial inclusion, and psychological well-being). We include the following variables as candidate dimensions of heterogeneity: respondent age, respondent education, education level of most educated household member, livestock (TLU) at baseline, goat herd size at baseline, land at baseline, and all household welfare index values at baseline. We test for heterogeneous effects separately for original and PIF beneficiaries, and separately at endline 1 and endline 2. We initially test for heterogeneous impact of the full Heifer treatment. If we detect heterogeneity here, we then test for heterogenous effects of the NG and NVT treatments.

We find virtually no statistically robust evidence of impact heterogeneity.[22] The only outcome for which we do observe heterogeneity is women's empowerment over goat production, and we only see heterogeneity of the full treatment and the NVT treatment. Under these treatments, classification analysis (CLAN) reveals that those most impacted at both endline 1 and endline 2 were younger, exhibited lower women's empowerment at baseline, had fewer livestock at baseline, and had fewer goats in particular at baseline. Given that we find no such heterogeneity in the NG treatment, we conclude the goat asset transfer is most impactful for women who have fewer goats and are less empowered at baseline, at least in terms of empowering them in goat production activities.

While this may seem like a specific instance of heterogeneity among many tested, it is likely an important one. Overall treatment effects fail to demonstrate the necessity of a goat asset transfer for achieving impact across most outcomes. In contrast, the heterogeneity analysis suggests the goat asset transfer may still be important for increasing empowerment over livestock production for women who are not empowered to begin with, at least in part because they own fewer livestock. Noting different targeting principles, this finding stands in contrast to the Banerjee et al. (2015) Graduation program findings, which observed the highest impacts for the wealthiest among the targeted ultra-poor.

[22] As a check, we use the less conservative honest causal trees method of Athey and Imbens (2016) and also find very little evidence of heterogeneous treatment effects.

9 Cost-benefit analysis

An important question for any anti-poverty program is whether it is cost-effective. Table 10 presents our cost-benefit analysis (the appendix provides the same analysis using alternative discount rates). We calculate benefits and costs separately for each treatment and beneficiary type. However, from a policy prescription standpoint, the most relevant numbers are a weighted average of direct and PIF beneficiaries compared across treatments.

We calculate benefits as the combined value of increases in goat herd size and goat profits.[23] Using only benefit estimates that are statistically significant at the 0.05 confidence level, we sum the total value of additional goats owned (at 6,000 NPR per goat)[24] at endline 2, the value of additional goats sales at endline 1 and endline 2, and a discounted continued stream of additional goat sales equal to the impact on goat sales at endline 2. We do not assume increasing sales growth beyond endline 2 since we do not observe significant increases in the impacts on herd size between endline 1 and endline 2. In this way, our estimates are conservative. Because we do not observe statistical differences in benefits across FT and NG, we use the average across both treatments.

The discount rate chosen has major implications for the value of the continued stream of additional goat sales, and thus on total program benefits. We assume a 10% discount rate, noting that 10% is the most conservative discount rate used by Banerjee et al. (2015). We discount from endline 2, and therefore inflate the value of livestock sold at endline 1 by one year. Using a 10% discount rate these benefits are 66,516 NPR for FT and NG direct beneficiaries. In the NVT treatment direct benefits are estimated to be 8,838 NPR. PIF benefits are estimated as 37,310 NPR for both FT and NG, and 14,412 NPR for NVT.

We then use detailed cost data from Heifer and its implementing field partners. Costs can be broken down into operations and administration costs. Operations costs include livestock, vegetable garden inputs, training, and equipment and supplies for goat shelter construction and community animal health workers. Administrative costs include technical services, personnel (most notably a project manager) and office expenses.[25] The average cost per direct beneficiary in the FT treatment is 68,676 NPR accounting for both operating

[23] Although we observe some increased savings, we exclude savings from our measure of benefits – observed increases in savings may have come from decreases in expenditures that we cannot statistically identify, or from livestock sales in which case including them would be double counting.

[24] Heifer purchases doe goats for approximately 6,000 NPR. From another project in Nepal we have data on goat prices at time of sale that range from 4,000 for a small female goat to 15,000 for a large male goat. We believe this makes 6,000 a conservative price to use for cost-benefit analysis.

[25] Operations expenses can be linked to a treatment and in some cases to direct or indirect beneficiaries. Administrative costs are allocated 30 percent to direct beneficiaries and 70 percent to PIF beneficiaries. Ultimately BCRs are most interesting for the aggregate of original and PIFs in a given treatment, so this distinction is unimportant.

and administrative cost. In the NG treatment the cost is 36,342 NPR. In the NVT group the cost per is 45,836 NPR. The cost per PIF is much lower. We assume a ratio of five PIF beneficiaries to one direct beneficiary. In the FT group the cost per PIF beneficiary is 6,290 NPR. In the NG treatment group it is 5,915 NPR and in the NVT groups it is 5,533 NPR. Taking a weighted average across both direct and PIF beneficiaries, the cost per beneficiary is 17,932 NPR in the FT treatment, 12,176 NPR in the NG treatment, and 12,502 NPR in the NVT treatment. In the combined FT and NG treatment group, the cost per beneficiary is 14,928 NPR. To accommodate the research design, Heifer incurred higher than typical administrative costs. Typically, Heifer intervenes in all nine wards in a VDC, whereas here they only intervened in the central ward. At scale, the administrative portion of the costs per beneficiary would be much lower than they are here, whereas the operations cost per beneficiary would be unchanged. We do not conduct the analysis here as if the program was done at the typical scale, again making our estimates conservative.

Aggregating direct and PIF costs and benefits, the BCR using the assumed 10% discount rate is 2.5 for the FT treatment, 3.7 for the NG treatment, and 1.4 for the NVT treatment. These high BCRs are due to low cost for PIF beneficiaries. Recall that a PIF beneficiary is any household in a Heifer ward that is not designated as a direct beneficiary. Thus, even in the NVT treatment there are PIF beneficiaries, but they typically benefit much less from the program than their counterparts in the FT and NG treatments. Because there are so many of them (approximately 140 per ward), the cost for each is low. The benefits, however, compare favorably with those of direct beneficiaries. The BCR for FT direct beneficiaries is 0.7, whereas the BCR for FT PIF beneficiaries is 7.0. Of course, PIF beneficiaries do not exist without direct beneficiaries. In this way, the relevant number is the weighted average across direct and PIF beneficiaries.

The CBRs described here are generally high, even under conservative assumptions. However, this is in large part due to the extremely low cost per beneficiary of the program, at least relative to other similarly designed programs. We also calculate the per beneficiary net present value (NPV) of the program. The average NPV for FT beneficiaries is 36,050 NPR. That is approximately 10% of total household income for a single year. Thus while the return on investment from the program is very high, the ultimate monetary impact for the household is rather modest.

10 Conclusion

This paper presents RCT results from a goat livelihoods program in rural Nepal. The program aims to empower poor rural women and enable the achievement of a sustainable exit

from poverty. In the main treatment arm, direct beneficiaries receive a livestock transfer, technical training on improved animal management and entrepreneurship, values-based training, and join SHGs. We present evidence that these households have bigger herds, improved livestock practices, more goat sales and higher profit from goat production than control households. Women are more empowered - particularly in goat production decisions and through greater participation in groups - and have greater financial inclusion.

We do not find robust evidence the Heifer program had a significant effect on total household income (Table 9). While the Heifer program expands a household's productive asset stocks (goats) and therefore increases goat income, livestock income is simply not a substantial source of income for households in the context being studied. This finding raises a number of questions. Would the program be more effective in a context where livestock income matters more *ex ante*? Can livestock production be a main livelihood and pathway out of poverty or are there limitations? Remittances represent 45% of income on average; in a context where external sources of revenue are more limited, would the findings be different? Future research could explore answers to these questions.

A unique aspect of the intervention is values-based training that includes a PIF mechanism. Under the PIF model, initial beneficiaries subsequently recruit, train, and give a non-negligible asset (valued at approximately 150 USD) to others in their community. Our study is uniquely designed to estimate the impact for both direct and PIF beneficiaries. We observe positive welfare impacts not only among households who received livestock and training directly from the program, but also for those brought into the program through the PIF mechanism. These results are corroborated by evidence of much weaker PIF impacts in a second treatment group that didn't participate in the values-based training modules. Our observed PIF impacts, combined with cost analysis, demonstrate how encouragement to pay benefits forward can help achieve a broader impact at much lower cost than a program without it.

Although paying it forward is a well-known concept – particularly popular during the holiday season in developed countries with widely publicized examples of paying for a stranger's coffee or leaving an unfathomably large tip at a restaurant – it is rarely incorporated into the design of anti-poverty programs.[26] Yet it could be, and our analysis suggests this unique program component could be an important and cost-effective tool for achieving program goals. Although we observe positive effects of the program, one is left to wonder if PIF can be as effective when the costs of giving are higher. For example, if a larger asset had been given and the requirements for giving were raised accordingly, would beneficiaries still PIF or would they be more likely to evade this program requirement? What are the

[26]PIF is a component of all Heifer programs globally.

implications for cost effectiveness?

We also test the importance of the productive asset transfer in achieving impacts by comparing a third treatment arm without a livestock transfer. In most cases, we cannot detect a statistically significant difference in impact between those whom received the goat and those whom did not. When beneficiaries pay it forward, they essentially postpose their own ability to internalize economic benefits from the newly acquired asset. Further research could explore this critical question of how important the productive asset transfer really is, particularly given the added cost. A related question is whether a larger asset would be more beneficial, and whether the CBR would be strengthened or weakened.

Does the program provide a sustainable pathway out of poverty? Section 9 presents evidence of cost effectiveness. Even with the least conservative discount rate of 5%, the total monetary impact on goat production is approximately one-third of total household income for a single year. Yet, the positive program impacts do not seem large enough to be considered path-breaking, at least not compared to relevant alternatives such as the prospect of sending a household member away as a migrant, as is common in this context. Does this suggest the existence of additional market-level or behavioral constraints – beyond access to productive assets and human, social and financial capital – simultaneously working against the poor's transition into becoming a successful entrepreneurs? We expect this question will be of critical importance for policy makers and practitioners seeking impactful, sustainable and cost effective solutions for ending poverty in the coming years.

References

2005. *Livestock sector brief: Nepal*. Livestock information, sector analysis and policy branch: Food and Agriculture Organization of the United Nations, July.

Ahrens, A., C.B. Hansen, and M.E. Schaffer. 2018. "PDSLASSO: Stata module for post-selection and post-regularization OLS or IV estimation and inference." Working paper, Boston College Department of Economics.

Alkire, S., R. Meinzen-Dick, A. Peterman, A. Quisumbing, and G. Seymour. 2013. "The Women's Empowerment in Agriculture Index." *World Development* 52:71–91.

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* 103:1481–1495.

Ashraf, N., D. Karlan, and W. Yin. 2010. "Female empowerment: Impact of a commitment savings product in the Philippines." *World Development* 38:333–344.

Athey, S., and G. Imbens. 2016. "Recursive partitioning for heterogeneous causal effects." *Proceedings of the National Academy of Sciences* 113:7353–7360.

Bandiera, O., R. Burgess, N. Das, S. Gulesci, I. Rasul, and M. Sulaiman. 2017. "Labor markets and poverty in village economies." *The Quarterly Journal of Economics* 132:811–870.

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.

Banerjee, A., D. Karlan, R.D. Osei, H. Trachtman, and C. Udry. 2018. "Unpacking a Multi-Faceted Program to Build Sustainable Income for the Very Poor." Working paper, National Bureau of Economic Research.

Banerjee, A.V. 2013. "Microcredit Under the Microscope: What Have We Learned in the Past Two Decades, and What Do We Need to Know?" *Annual Review of Economics* 5:487–519.

Belloni, A., V. Chernozhukov, and C. Hansen. 2014. "Inference on treatment effects after selection among high-dimensional controls." *Review of Economic Studies* 81:608–650.

Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernandez-Val. 2018. "Generic machine learning inference on heterogenous treatment effects in randomized experiments." Working paper, National Bureau of Economic Research.

Cull, R., T. Ehrbeck, and N. Holle. 2014. "Financial inclusion and development: Recent impact evidence." CGAP Focus Note No. 92, Consultative Group to Assist the Poor.

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.

DeVries, J. 2011. "Heifer's Approach to Sustainable Community Development." *Journal of Agricultural and Food Information* 12:40–48.

Food and Agriculture Organization of the United Nations. 2011. "The state of food and agriculture: Women in agriculture: Closing the gender gap for development." Working paper.

Halloway, K., Z. Ziazi, and R. Rouse. 2017. "Women's economic empowerment through financial inclusion: A review of existing evidence and remaining knowledge gaps." Working paper, Innovations for Poverty Action, March.

Haushofer, J., and E. Fehr. 2014. "On the psychology of poverty." *Science* 344:862–867.

Janzen, S., N. Magnan, S. Sharma, and W. Thompson. 2018a. "Evaluation of the Welfare Impacts of a Livestock Transfer Program in Nepal."

Janzen, S.A., N.P. Magnan, S. Sharma, and W.M. Thompson. 2018b. "Short-term impacts of a pay-it-forward livestock transfer and training program in Nepal." *AEA Papers and Proceedings* 108:422–425.

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.

Karlan, D., A.L. Ratan, and J. Zinman. 2014. "Savings by and for the Poor: A Research Review and Agenda." *Review of Income and Wealth* 60:36–78.

Karlan, D., B. Savonitto, B. Thuysbaert, and C. Udry. 2017. "Impact of savings groups on the lives of the poor." *Proceedings of the National Academy of Sciences* 114:3079–3084.

Ksoll, C., H.B. Lilleør, J.H. Lønborg, and O.D. Rasmussen. 2016. "Impact of Village Savings and Loan Associations: Evidence from a cluster randomized trial." *Journal of Development Economics* 120:70–85.

Malapit, H., C. Kovarik, K. Sproule, R. Meinzen-Dick, and A. Quisumbing. 2015. "Instructional Guide on the Abbreviated Women's Empowerment in Agriculture Index (A-WEAI)." Unpublished.

Malapit, H.J.L., E. Sraboni, A.R. Quisumbing, and A.U. Ahmed. 2019. "Intrahousehold empowerment gaps in agriculture and children's well-being in Bangladesh." *Development Policy Review* 37:176–203.

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.

Phadera, L., H. Michaelson, A. Winter-Nelson, and P. Goldsmith. 2019. "Do Asset Transfers Build Household Resilience?" *Journal of Development Economics* 138:205–227.

Radloff, L.S. 1977. "The CES-D scale a self-report depression scale for research in the general population." *Applied Psychological Measurement* 1:385–401.

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.

Rosenberg, M. 1965. *Society and the adolescent self-image*, Princeton University Press Princeton, NJ.

Rotter, J.B. 1966. "Generalized expectancies for internal versus external control of reinforcement." *Psychological Monographs: General and Applied* 80:1.

Scheier, M.F., C.S. Carver, and M.W. Bridges. 1994. "Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test." *Journal of personality and social psychology* 67:1063.

World Bank. 2011. *World Development Report 2012: Gender equality and development*. World Bank Publications, Washington DC.

Figure 1: Study timeline

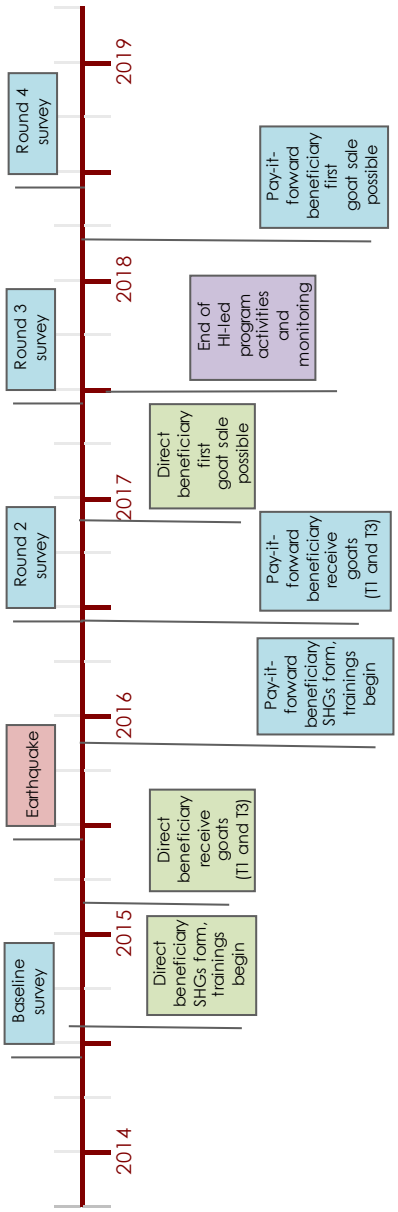

Figure 2: Heifer self-help group membership status at endline 1

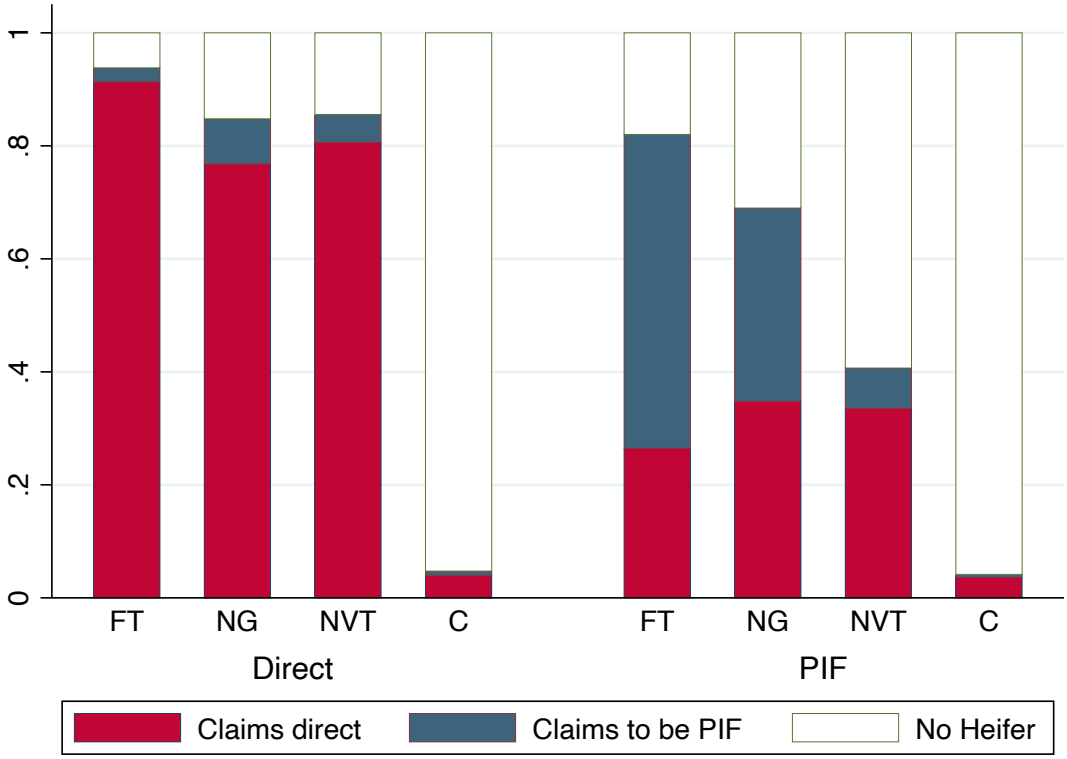

Figure 3: Heifer self-help group membership status at endline 2

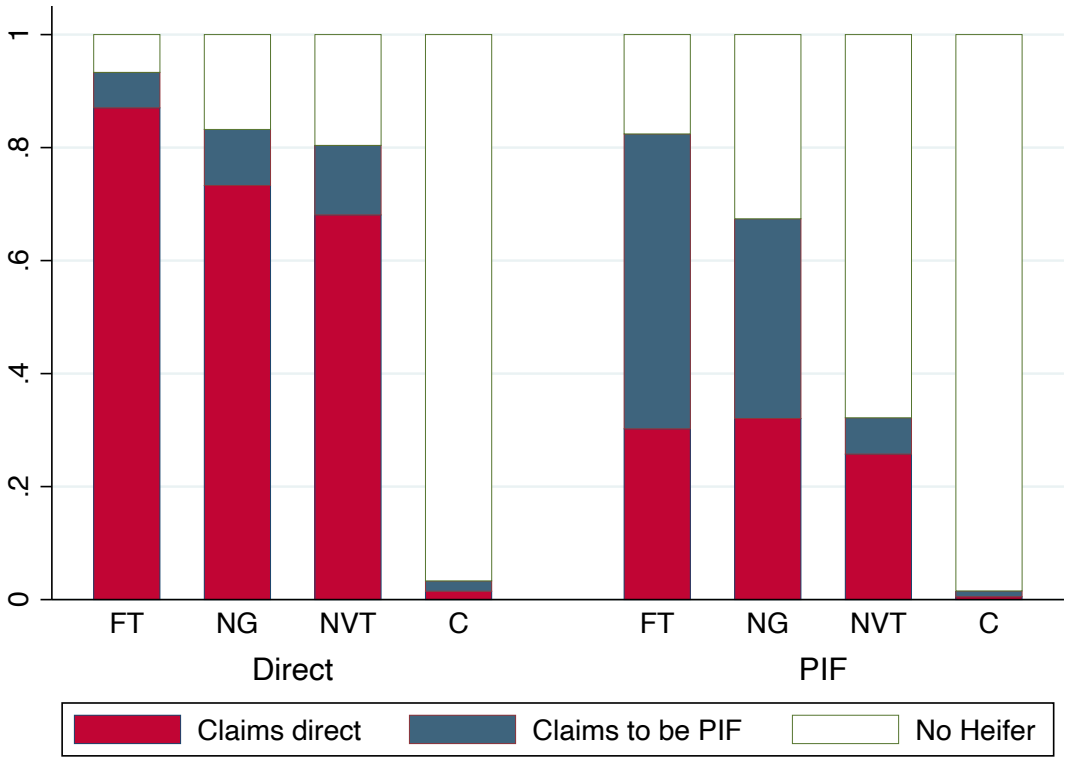

Table 1: Description of program components by treatment arm

Description of Program Components	**FT**	**NG**	**NVT**
Basic intervention			
SHG formation	✓	✓	✓
SHG savings encouragement	✓	✓	✓
training on nutrition	✓	✓	✓
training on improved animal management	✓	✓	✓
training and cash support ($5) for home gardening	✓	✓	✓
training and cash support ($10) for fodder & forage production	✓	✓	✓
cash support ($40) for goat shed improvement	✓	✓	✓
access to community animal health worker	✓	✓	✓
Productive asset transfer			
2 doe goats	✓		✓
1 shared buck of improved breeding stock (per SHG)	✓		✓
Values-based trainings			
encouragement to "pay-it-forward"	✓	✓	
training on SHG management	✓	✓	
training on gender and justice	✓	✓	
*training on remaining HI Cornerstones**	✓	✓	

*The remaining HI Cornerstones not noted elsewhere in this table include: accountability; sharing and caring; sustainability and self-reliance; improving the environment; income; full participation; training, education, and communication; and spirituality.

Table 2: Descriptive statistics and balance on demographic variables

Direct beneficiaries

	Control mean	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Land (hectares)	0.465	0.002	0.017	-0.053	-0.015	0.027	0.041	1,341
	(0.567)	(0.075)	(0.072)	(0.078)	(0.074)	(0.081)	(0.077)	
Productive asset index	0.786	0.115	0.253	0.010	-0.004	0.145	0.148	1,341
	(0.810)	(0.129)	(0.160)	(0.112)	(0.202)	(0.150)	(0.182)	
Tropical livestock units	2.523	0.076	0.155	0.357	0.131	-0.430	-0.560	1,341
	(2.130)	(0.376)	(0.362)	(0.361)	(0.404)	(0.413)	(0.386)	
Livestock income (gross, NPR)	12,930.970	2,394.458	2,407.404	-21.007	-1,249.469	-500.244	749.225	1,341
	(22,031.023)	(2,899.058)	(2,861.862)	(2,668.024)	(2,941.976)	(2,725.793)	(3,017.341)	
Goat herd size	4.036	0.723	0.627	0.497	0.151	-0.326	-0.478	1,341
	(4.370)	(0.691)	(0.961)	(0.701)	(1.048)	(0.640)	(1.087)	
Goat income (gross, NPR)	6,912.715	434.510	1,653.063	-510.918	-2,975.077	-969.304	2,005.774	1,341
	(12,431.064)	(1,866.838)	(2,011.495)	(1,569.859)	(1,984.822)	(1,161.967)	(1,950.858)	
Income aspirations (NPR)	136,677.594	-78,526.167	-121,109.181*	-61,815.616	26,019.669	-33,749.920	-59,769.589	1,341
	(582,219.010)	(70,978.388)	(66,657.634)	(78,342.789)	(41,067.544)	(61,302.190)	(53,530.962)	
Status aspirations (num. women)	14.988	-5.810	-2.290	1.531	-4.909	-8.251	-3.342	1,341
	(68.643)	(6.103)	(6.630)	(8.837)	(4.261)	(7.789)	(8.349)	
Discount rate	0.040	-0.008	-0.020	-0.016	0.005	0.003	-0.002	1,341
	(0.082)	(0.014)	(0.012)	(0.012)	(0.011)	(0.011)	(0.010)	
Planning horizon (weeks)	1.369	-0.176	-0.185	-0.016	-0.217	-0.166	0.051	1,341
	(1.618)	(0.246)	(0.244)	(0.257)	(0.237)	(0.246)	(0.228)	
Household size	5.790	-0.434*	0.005	-0.021	-0.367	-0.520**	-0.153	1,341
	(2.628)	(0.226)	(0.251)	(0.237)	(0.281)	(0.232)	(0.272)	
Respondent age	40.731	-0.498	-0.660	-2.091	0.939	2.060	1.120	1,341
	(13.831)	(1.344)	(1.675)	(1.557)	(1.901)	(1.697)	(2.076)	
Respondent education (years)	2.604	-0.456	0.416	-0.110	-0.992*	-0.296	0.696	1,341
	(3.918)	(0.561)	(0.641)	(0.598)	(0.500)	(0.425)	(0.546)	
Respondent litterate	0.535	-0.020	0.081	0.017	-0.095	-0.046	0.049	1,341
	(0.499)	(0.082)	(0.076)	(0.089)	(0.070)	(0.088)	(0.077)	
Spouse age	44.357	-0.466	-0.931	-1.970	0.701	1.629	0.928	1,341
	(13.052)	(1.350)	(1.560)	(1.530)	(1.817)	(1.618)	(1.946)	
Spouse education (years)	5.105	-0.279	0.603	0.367	-0.812	-0.649	0.162	1,341
	(4.020)	(0.732)	(0.779)	(0.776)	(0.626)	(0.534)	(0.662)	
Spouse litterate	0.791	0.081	0.116*	0.076	-0.014	0.007	0.020	1,341
	(0.407)	(0.072)	(0.066)	(0.074)	(0.059)	(0.066)	(0.062)	
No spouse	0.117	-0.027	0.033	-0.037*	-0.055	0.017	0.072*	1,341
	(0.322)	(0.023)	(0.031)	(0.022)	(0.038)	(0.026)	(0.036)	
Maximum education in household (years)	8.970	-0.451	0.097	-0.261	-0.674	-0.385	0.289	1,341
	(3.923)	(0.672)	(0.762)	(0.755)	(0.654)	(0.573)	(0.729)	

Potential PIF beneficiaries

	Control mean	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Land (hectares)	0.465	0.036	0.132*	0.119*	-0.056	-0.099	-0.044	1,028
	(0.607)	(0.062)	(0.071)	(0.069)	(0.088)	(0.086)	(0.091)	
Productive asset index	0.735	0.214*	0.329**	0.098	0.058	0.187	0.129	1,028
	(0.789)	(0.114)	(0.125)	(0.123)	(0.152)	(0.163)	(0.156)	
Tropical livestock units	2.387	0.313	0.523	0.217	-0.078	0.009	0.087	1,028
	(1.995)	(0.268)	(0.366)	(0.316)	(0.396)	(0.335)	(0.444)	
Livestock income (gross, NPR)	13,925.102	-765.266	4,248.954	-1,528.265	-2,105.212	833.105	2,938.317	1,028
	(29,086.324)	(1,995.886)	(3,581.797)	(2,747.261)	(2,954.052)	(2,307.226)	(3,092.641)	
Goat herd size	3.890	1.491**	0.941	0.965	0.479	0.022	-0.457	1,028
	(4.038)	(0.624)	(0.660)	(0.676)	(0.615)	(0.644)	(0.777)	
Goat income (gross, NPR)	6,827.650	394.443	2,986.557	-1,335.812	-1,774.015	655.393	2,429.408	1,028
	(13,305.308)	(1,387.886)	(1,993.653)	(1,371.447)	(1,514.782)	(936.314)	(1,651.862)	
Income aspirations (NPR)	171,313.663	-35,950.727	-57,901.388	206,911.018	31,668.729	-283,902.685	-315,571.415	1,028
	(1,673,165.306)	(47,397.326)	(50,402.801)	(190,691.045)	(23,237.553)	(214,931.805)	(214,672.843)	
Status aspirations (num. women)	13.177	-2.118	2.317	4.081	-7.358	-7.898	-0.540	1,028
	(56.393)	(3.310)	(5.084)	(5.551)	(5.511)	(5.940)	(7.477)	
Discount rate	0.040	-0.017	-0.018	-0.027**	-0.004	0.008	0.012	1,028
	(0.078)	(0.016)	(0.018)	(0.013)	(0.020)	(0.013)	(0.017)	
Planning horizon (weeks)	1.425	0.135	-0.036	-0.209	-0.059	0.244	0.303	1,028
	(1.697)	(0.287)	(0.231)	(0.221)	(0.342)	(0.326)	(0.279)	
Household size	5.835	-0.468	0.138	0.068	-0.462	-0.517**	-0.055	1,028
	(2.714)	(0.286)	(0.366)	(0.301)	(0.371)	(0.250)	(0.374)	
Respondent age	40.689	1.803	0.517	1.206	1.959	2.011	0.052	1,028
	(13.929)	(1.707)	(1.370)	(1.497)	(2.145)	(2.193)	(1.866)	
Respondent education (years)	2.430	-0.346	-0.040	-0.310	-0.169	-0.199	-0.029	1,028
	(3.822)	(0.492)	(0.465)	(0.536)	(0.492)	(0.564)	(0.528)	
Respondent litterate	0.476	-0.044	0.040	-0.051	-0.050	0.021	0.071	1,028
	(0.500)	(0.073)	(0.069)	(0.072)	(0.082)	(0.083)	(0.080)	
Spouse age	44.187	1.397	2.178	0.751	-0.548	1.110	1.658	1,028
	(13.144)	(1.443)	(1.578)	(1.346)	(1.970)	(1.684)	(1.800)	
Spouse education (years)	4.829	-0.290	0.021	-0.117	-0.240	-0.252	-0.012	1,028
	(3.975)	(0.489)	(0.516)	(0.589)	(0.459)	(0.578)	(0.573)	
Spouse litterate	0.757	-0.031	0.021	-0.012	-0.008	0.004	0.012	1,028
	(0.429)	(0.058)	(0.048)	(0.058)	(0.060)	(0.071)	(0.059)	
No spouse	0.110	0.002	0.017	0.016	0.001	0.009	0.008	1,028
	(0.313)	(0.037)	(0.033)	(0.034)	(0.040)	(0.040)	(0.035)	
Maximum education in household (years)	8.739	-0.563	0.077	-0.193	-0.616	-0.315	0.301	1,028
	(4.100)	(0.523)	(0.611)	(0.591)	(0.600)	(0.567)	(0.698)	

FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses. Columns 4-6: Wald tests of equal treatment effects with p-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Descriptive statistics and balance on key outcomes of interest

Direct beneficiaries

	Control mean	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Goat herdsize	4.036	0.723	0.627	0.497	0.151	-0.326	-0.478	1,341
	(4.370)	(0.691)	(0.961)	(0.701)	(1.048)	(0.640)	(1.087)	
Goat good practice index	-0.012	0.015	-0.031	-0.030	0.029	0.003	-0.027	1,341
	(0.983)	(0.128)	(0.145)	(0.157)	(0.146)	(0.151)	(0.171)	
Goat income (gross, NPR)	6,912.715	434.510	1,653.063	-510.918	-2,975.077	-969.304	2,005.774	1,341
	(12,431.064)	(1,866.838)	(2,011.495)	(1,569.859)	(1,984.822)	(1,161.967)	(1,950.858)	
Women's empowerment over goats	0.121	0.108	0.149	0.228	-0.015	-0.139	-0.124	1,341
	(0.962)	(0.158)	(0.149)	(0.145)	(0.168)	(0.164)	(0.144)	
Household income (NPR)	99,536.708	-4,095.668	4,732.059	-8,532.565	-14,079.038*	1,494.679	15,573.717**	1,341
	(72,828.669)	(6,974.698)	(7,937.233)	(6,393.766)	(7,996.208)	(5,508.928)	(7,594.899)	
Asset index	-0.136	-0.191	-0.119	-0.239	0.014	0.046	0.032	1,341
	(0.914)	(0.154)	(0.163)	(0.170)	(0.128)	(0.123)	(0.148)	
Women's empowerment index	-0.081	-0.171	-0.022	-0.139	-0.119	0.033	0.152	1,341
	(0.998)	(0.106)	(0.119)	(0.122)	(0.123)	(0.125)	(0.137)	
Financial inclusion index	-0.078	-0.141	0.055	-0.241*	-0.263	0.078	0.340*	1,341
	(0.945)	(0.119)	(0.173)	(0.127)	(0.162)	(0.098)	(0.171)	
Mental Health index	-0.072	-0.154	-0.030	-0.107	-0.221*	-0.087	0.134	1,341
	(1.015)	(0.128)	(0.120)	(0.149)	(0.114)	(0.151)	(0.143)	

Potential PIF beneficiaries

	Control mean	FT	NG	NVT	FT=NG	FT=NVT	NG=NVT	N
Goat herdsize	3.890	1.491**	0.941	0.965	0.479	0.022	-0.457	1,028
	(4.038)	(0.624)	(0.660)	(0.676)	(0.615)	(0.644)	(0.777)	
Goat good practice index	0.062	0.135	0.117	-0.009	0.065	0.135	0.070	1,028
	(0.942)	(0.150)	(0.140)	(0.151)	(0.124)	(0.140)	(0.116)	
Goat income (gross, NPR)	6,827.650	394.443	2,986.557	-1,335.812	-1,774.015	655.393	2,429.408	1,028
	(13,305.308)	(1,387.886)	(1,993.653)	(1,371.447)	(1,514.782)	(936.314)	(1,651.862)	
Women's empowerment over goats	0.170	0.276**	0.270*	0.122	0.054	0.229*	0.175	1,028
	(0.929)	(0.134)	(0.147)	(0.152)	(0.138)	(0.133)	(0.158)	
Household income (NPR)	108,211.882	-16,075.052	681.707	-20,021.135**	-19,842.329*	-1,283.774	18,558.556*	1,028
	(92,100.607)	(11,774.346)	(11,505.681)	(8,714.430)	(11,125.386)	(6,137.142)	(10,832.459)	
Asset index	0.030	0.024	0.131	-0.051	0.018	0.088	0.070	1,028
	(0.975)	(0.158)	(0.157)	(0.155)	(0.170)	(0.155)	(0.158)	
Women's empowerment index	0.032	-0.050	0.067	0.090	0.006	-0.021	-0.027	1,028
	(1.001)	(0.124)	(0.098)	(0.106)	(0.127)	(0.133)	(0.109)	
Financial inclusion index	-0.061	-0.067	-0.007	-0.168	0.003	0.136	0.132	1,028
	(0.845)	(0.125)	(0.133)	(0.120)	(0.116)	(0.106)	(0.110)	
Mental Health index	-0.094	-0.118	-0.046	-0.205	-0.090	0.010	0.099	1,028
	(0.982)	(0.154)	(0.147)	(0.153)	(0.138)	(0.149)	(0.154)	

FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses. Columns 4-6: Wald tests of equal treatment effects with p-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Treatment effects on goat herd dynamics

Direct beneficiaries

	Endline 1 (2.5 years post-treatment)					Endline 2 (3.5 years post treatment)				
	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Herd size	2.691***	1.359**	2.423***	4.193	1,053	2.626***	1.578**	1.473**	4.283	1,030
	(0.467)	(0.537)	(0.491)	(4.796)		(0.512)	(0.618)	(0.587)	(4.952)	
Purchased	0.040	0.077	-0.003	0.156	1,053	-0.026	-0.028	-0.021	0.180	1,030
	(0.053)	(0.061)	(0.058)	(0.548)		(0.060)	(0.053)	(0.067)	(0.729)	
Births	1.672***	1.087***	1.616***	1.869	1,053	1.170***	0.542	0.920***	2.268	1,030
	(0.259)	(0.262)	(0.276)	(2.699)		(0.332)	(0.338)	(0.343)	(4.104)	
Gifts received (non-Heifer)	-0.022*	0.009	-0.001	0.022	1,053	0.026	0.007	0.040	0.018	1,030
	(0.012)	(0.018)	(0.015)	(0.169)		(0.032)	(0.018)	(0.033)	(0.201)	
Gifts received (Heifer)	0.144***	0.019	0.216*	0.000	1,053	0.038*	0.000	0.121***	0.000	1,030
	(0.049)	(0.040)	(0.128)	(0.000)		(0.022)	(0.023)	(0.042)	(0.000)	
Gifts given	1.285***	0.062	-0.004	0.069	1,053	0.272***	-0.000	0.022	0.085	1,030
	(0.196)	(0.095)	(0.074)	(0.434)		(0.073)	(0.046)	(0.043)	(0.417)	
Deaths	0.557***	0.215	1.008***	0.502	1,053	0.638***	0.142	0.927***	0.511	1,030
	(0.206)	(0.215)	(0.286)	(1.160)		(0.128)	(0.129)	(0.204)	(1.274)	
Sold	0.486***	0.830***	0.519***	1.120	1,053	0.890***	0.694**	0.539**	1.353	1,030
	(0.183)	(0.270)	(0.161)	(1.869)		(0.271)	(0.340)	(0.217)	(3.835)	
Consumed	0.068**	0.087**	0.071**	0.062	1,053	-0.000	-0.040	0.013	0.121	1,030
	(0.031)	(0.044)	(0.029)	(0.241)		(0.042)	(0.042)	(0.046)	(0.327)	

Potential PIF beneficiaries

	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Herd size	1.160***	1.487***	0.923*	3.542	774	1.280***	1.980***	0.362	3.538	759
	(0.331)	(0.464)	(0.473)	(3.494)		(0.362)	(0.524)	(0.487)	(3.641)	
Purchased	-0.013	0.048	-0.033	0.234	774	-0.149*	-0.144*	-0.231***	0.338	759
	(0.080)	(0.085)	(0.090)	(0.718)		(0.080)	(0.074)	(0.076)	(0.973)	
Births	0.533**	0.863***	0.378*	1.609	773	0.320	0.866***	0.202	1.774	758
	(0.209)	(0.236)	(0.213)	(2.216)		(0.288)	(0.321)	(0.356)	(2.704)	
Gifts received (non-Heifer)	-0.027	0.004	0.045	0.052	774	-0.042	-0.018	-0.005	0.062	759
	(0.030)	(0.032)	(0.047)	(0.266)		(0.033)	(0.036)	(0.037)	(0.472)	
Gifts received (Heifer)	0.240***	0.017	0.055	0.000	774	0.049*	0.012	0.064*	0.005	759
	(0.047)	(0.020)	(0.037)	(0.000)		(0.025)	(0.019)	(0.033)	(0.072)	
Gifts given	0.091***	0.076*	0.016	0.026	774	0.060	0.001	0.012	0.062	759
	(0.035)	(0.046)	(0.030)	(0.190)		(0.043)	(0.043)	(0.049)	(0.401)	
Deaths	0.125	0.004	-0.108	0.646	774	-0.062	0.050	0.003	0.697	759
	(0.133)	(0.176)	(0.139)	(1.436)		(0.117)	(0.183)	(0.131)	(1.413)	
Sold	0.434**	0.950***	0.224	0.880	774	0.364*	0.780***	0.267	0.949	759
	(0.192)	(0.268)	(0.191)	(2.746)		(0.202)	(0.224)	(0.215)	(1.923)	
Consumed	0.008	0.007	-0.014	0.094	774	0.026	-0.007	0.016	0.092	759
	(0.033)	(0.030)	(0.024)	(0.292)		(0.024)	(0.025)	(0.026)	(0.290)	

FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Treatment effects on livestock management practices

Direct beneficiaries

	Endline 1 (2.5 years post-treatment)					Endline 2 (3.5 years post treatment)				
	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Goat good practice index	0.771***	0.459***	0.664***	-0.000	1,053	0.692***	0.430***	0.500***	-0.000	1,030
	(0.133)	(0.118)	(0.134)	(1.000)		(0.120)	(0.117)	(0.112)	(1.000)	
Use improved pen	0.508***	0.335***	0.357***	0.149	1,053	0.444***	0.330***	0.373***	0.143	1,030
	(0.062)	(0.061)	(0.047)	(0.357)		(0.066)	(0.071)	(0.064)	(0.351)	
Remove manure at least weekly	0.372***	0.280***	0.239***	0.415	1,053	0.294***	0.112*	0.143**	0.456	1,030
	(0.068)	(0.065)	(0.089)	(0.494)		(0.057)	(0.060)	(0.061)	(0.499)	
Remove manure at least monthly	0.317***	0.193***	0.224***	0.564	1,053	0.227***	0.137**	0.192***	0.592	1,030
	(0.055)	(0.052)	(0.060)	(0.497)		(0.054)	(0.058)	(0.054)	(0.492)	
Use manure as fertilizer	0.199***	0.129***	0.203***	0.669	1,053	0.177***	0.101**	0.172***	0.643	1,030
	(0.045)	(0.047)	(0.042)	(0.471)		(0.044)	(0.050)	(0.043)	(0.480)	
Use any goat medicine	0.209***	0.128**	0.220***	0.676	1,053	0.217***	0.120**	0.215***	0.647	1,030
	(0.051)	(0.054)	(0.049)	(0.469)		(0.051)	(0.057)	(0.047)	(0.479)	
Vaccinate any goats	0.203***	0.081	0.262***	0.571	1,053	0.244***	0.142***	0.174***	0.482	1,030
	(0.055)	(0.058)	(0.055)	(0.496)		(0.058)	(0.055)	(0.046)	(0.501)	
Produce fodder	0.138*	0.060	0.103	0.215	1,053	0.201***	0.118***	0.103**	0.114	1,030
	(0.082)	(0.061)	(0.064)	(0.411)		(0.056)	(0.032)	(0.047)	(0.318)	
Use mineral block	0.050	0.023	0.043	0.065	1,053	0.021	0.034	0.011	0.070	1,030
	(0.042)	(0.030)	(0.049)	(0.248)		(0.032)	(0.030)	(0.028)	(0.255)	
Cave has visited home	0.376***	0.240***	0.332***	0.498	1,053	0.410***	0.249***	0.323***	0.379	1,030
	(0.052)	(0.054)	(0.055)	(0.501)		(0.052)	(0.058)	(0.053)	(0.486)	
Cave has visited ward	0.358***	0.229***	0.342***	0.542	1,053	0.373***	0.247***	0.319***	0.456	1,030
	(0.052)	(0.053)	(0.050)	(0.499)		(0.055)	(0.060)	(0.054)	(0.499)	

Potential PIF beneficiaries

	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Goat good practice index	0.432***	0.382***	0.130	0.000	774	0.425***	0.557***	0.019	-0.009	759
	(0.142)	(0.126)	(0.150)	(1.003)		(0.071)	(0.093)	(0.082)	(0.995)	
Use improved pen	0.230***	0.210***	0.097**	0.151	774	0.200***	0.254***	0.128**	0.159	759
	(0.052)	(0.049)	(0.049)	(0.359)		(0.058)	(0.050)	(0.051)	(0.367)	
Remove manure at least weekly	0.143**	0.115*	0.012	0.495	774	0.138***	0.168***	-0.061	0.467	759
	(0.073)	(0.064)	(0.074)	(0.501)		(0.051)	(0.041)	(0.052)	(0.500)	
Remove manure at least monthly	0.199***	0.131**	0.055	0.562	774	0.151***	0.187***	-0.021	0.564	759
	(0.059)	(0.056)	(0.071)	(0.497)		(0.038)	(0.040)	(0.042)	(0.497)	
Use manure as fertilizer	0.106***	0.076*	0.035	0.667	774	0.069**	0.060**	-0.034	0.656	759
	(0.038)	(0.044)	(0.045)	(0.473)		(0.028)	(0.029)	(0.034)	(0.476)	
Use any goat medicine	0.085**	0.093**	0.017	0.682	774	0.101***	0.081**	-0.032	0.667	759
	(0.042)	(0.047)	(0.054)	(0.467)		(0.027)	(0.035)	(0.042)	(0.473)	
Vaccinate any goats	0.135*	0.120*	0.023	0.568	774	0.127**	0.139***	-0.012	0.503	759
	(0.070)	(0.067)	(0.069)	(0.497)		(0.054)	(0.052)	(0.049)	(0.501)	
Produce fodder	-0.017	0.005	-0.010	0.240	774	0.005	0.071**	0.012	0.149	759
	(0.057)	(0.058)	(0.065)	(0.428)		(0.037)	(0.036)	(0.045)	(0.357)	
Use mineral block	0.058	0.037	0.044	0.047	774	0.038*	0.085***	-0.010	0.046	759
	(0.057)	(0.031)	(0.037)	(0.212)		(0.020)	(0.029)	(0.019)	(0.210)	
Cave has visited home	0.293***	0.220***	0.101*	0.469	774	0.303***	0.140***	0.042	0.436	759
	(0.046)	(0.053)	(0.061)	(0.500)		(0.054)	(0.051)	(0.055)	(0.497)	
Cave has visited ward	0.279***	0.223***	0.120**	0.505	774	0.307***	0.234***	0.068	0.462	759
	(0.049)	(0.053)	(0.055)	(0.501)		(0.055)	(0.056)	(0.058)	(0.500)	

FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Treatment effects on income, investment and profit

Direct beneficiaries

	Endline 1 (2.5 years post-treatment)					Endline 2 (3.5 years post treatment)				
	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Goat net income (NPR)	475.245	3,630.699**	391.647	5,097.662	1,053	3,852.134**	5,522.264**	941.671	5,294.334	1,030
	(976.011)	(1,669.680)	(1,304.053)	(10,506.690)		(1,620.495)	(2,175.931)	(1,454.607)	(10,999.545)	
Goat gross income (NPR)	2,969.446**	5,102.077***	2,056.620*	7,448.364	1,053	6,595.433***	6,679.341***	2,690.892**	7,581.448	1,030
	(1,155.131)	(1,978.455)	(1,068.120)	(11,681.747)		(1,865.068)	(2,457.529)	(1,364.001)	(12,881.056)	
Total goat costs (NPR)	2,776.927***	1,430.504*	2,016.023**	2,350.702	1,053	2,669.446***	1,175.794*	1,677.369**	2,287.114	1,030
	(713.101)	(767.938)	(801.745)	(4,836.564)		(816.918)	(659.467)	(765.600)	(5,073.662)	
Fodder costs (NPR)	1,853.660***	1,108.871	1,319.447*	1,814.545	1,053	1,733.343**	792.460	958.849	1,787.592	1,030
	(619.225)	(714.262)	(729.609)	(4,107.525)		(689.022)	(571.688)	(675.282)	(4,483.351)	
Veterinary costs (NPR)	801.761***	211.665	612.810***	447.855	1,053	762.323***	157.737	578.134***	405.368	1,030
	(150.650)	(151.929)	(195.114)	(1,045.677)		(197.128)	(170.622)	(178.571)	(954.110)	
Breeding costs (NPR)	95.165***	94.540***	73.576***	83.782	1,053	165.726***	235.841	140.139**	89.375	1,030
	(19.375)	(30.575)	(23.363)	(142.009)		(59.210)	(170.878)	(58.005)	(198.866)	
Marketing costs (NPR)	7.779	2.705	3.922	4.520	1,053	-1.846	-0.877	-1.743	4.779	1,030
	(4.797)	(3.254)	(4.169)	(34.157)		(2.396)	(2.630)	(1.816)	(29.993)	

Potential PIF beneficiaries

	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Goat net income (NPR)	502.704	4,991.952***	88.221	3,621.562	774	3,206.232**	5,320.522***	2,401.885*	3,462.347	759
	(1,107.058)	(1,437.315)	(1,234.087)	(11,172.110)		(1,405.608)	(1,493.290)	(1,276.937)	(8,937.463)	
Goat gross income (NPR)	1,746.552	6,111.196***	827.964	5,440.625	774	4,208.684***	6,598.559***	2,535.513*	5,434.398	759
	(1,138.362)	(1,396.758)	(1,247.497)	(11,488.235)		(1,525.795)	(1,624.172)	(1,424.571)	(9,625.371)	
Total goat costs (NPR)	1,187.433**	1,163.538**	661.916	1,819.062	774	930.439*	1,267.073**	-10.750	1,972.051	759
	(541.962)	(526.796)	(603.186)	(3,382.815)		(533.374)	(607.453)	(557.396)	(4,604.238)	
Fodder costs (NPR)	764.724*	1,042.030**	476.012	1,371.875	774	564.530	1,069.710**	97.365	1,444.103	759
	(446.265)	(479.343)	(518.772)	(3,015.621)		(406.415)	(536.988)	(471.218)	(3,739.687)	
Veterinary costs (NPR)	458.187***	104.555	184.096*	379.115	774	251.189	9.117	-151.769	463.436	759
	(125.456)	(85.310)	(104.395)	(895.763)		(175.963)	(129.875)	(151.604)	(1,210.910)	
Breeding costs (NPR)	64.760**	100.514***	15.514	64.948	774	138.623***	219.354**	38.430	62.718	759
	(26.309)	(26.784)	(23.971)	(157.067)		(44.155)	(87.690)	(37.523)	(133.584)	
Marketing costs (NPR)	2.512	5.529	0.999	3.125	774	-0.831	-0.768	-0.909	1.795	759
	(4.306)	(5.671)	(3.288)	(32.207)		(1.588)	(1.548)	(1.501)	(17.859)	

FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Treatment effects on empowerment in goat production

Direct beneficiaries

	Endline 1 (2.5 years post-treatment)					Endline 2 (3.5 years post treatment)				
	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Women's empowerment over goat index	0.470***	0.308***	0.429***	-0.000	1,053	0.514***	0.327***	0.456***	0.000	1,030
	(0.109)	(0.109)	(0.093)	(1.000)		(0.109)	(0.117)	(0.109)	(1.000)	
Owns goats	0.234***	0.131***	0.195***	0.625	1,053	0.221***	0.115**	0.227***	0.588	1,030
	(0.048)	(0.047)	(0.042)	(0.485)		(0.056)	(0.051)	(0.048)	(0.493)	
Decides on goat care	0.221***	0.127***	0.209***	0.647	1,053	0.235***	0.115**	0.232***	0.603	1,030
	(0.046)	(0.045)	(0.040)	(0.479)		(0.050)	(0.052)	(0.047)	(0.490)	
Decides on goat sales	0.235***	0.099*	0.177***	0.607	1,053	0.225***	0.133**	0.216***	0.585	1,030
	(0.054)	(0.055)	(0.050)	(0.489)		(0.059)	(0.057)	(0.057)	(0.494)	
Controls goat income	0.135***	0.137**	0.150***	0.389	1,053	0.201***	0.188***	0.147**	0.386	1,030
	(0.051)	(0.059)	(0.047)	(0.488)		(0.058)	(0.065)	(0.062)	(0.488)	

Potential PIF beneficiaries

	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Women's empowerment over goat index	0.243***	0.410***	0.125	0.007	774	0.273***	0.255***	0.058	-0.001	759
	(0.080)	(0.096)	(0.108)	(0.998)		(0.059)	(0.075)	(0.099)	(1.002)	
Owns goats	0.089**	0.124***	0.047	0.625	774	0.111***	0.108***	-0.010	0.600	759
	(0.042)	(0.044)	(0.050)	(0.485)		(0.034)	(0.033)	(0.044)	(0.491)	
Decides on goat care	0.089**	0.082*	-0.007	0.667	774	0.075**	0.061*	-0.033	0.646	759
	(0.035)	(0.044)	(0.048)	(0.473)		(0.034)	(0.036)	(0.042)	(0.479)	
Decides on goat sales	0.109**	0.111**	0.036	0.609	774	0.115***	0.116***	0.005	0.585	759
	(0.051)	(0.053)	(0.063)	(0.489)		(0.028)	(0.037)	(0.039)	(0.494)	
Controls goat income	0.125***	0.307***	0.112***	0.286	774	0.144***	0.136***	0.089	0.333	759
	(0.036)	(0.042)	(0.041)	(0.453)		(0.048)	(0.052)	(0.060)	(0.473)	

FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Treatment effects on household assets

Direct beneficiaries

	Endline 1 (2.5 years post-treatment)					Endline 2 (3.5 years post treatment)				
	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Asset index	0.118*	-0.015	0.028	0.000	1,053	0.213**	-0.059	0.001	-0.000	1,030
	(0.071)	(0.080)	(0.095)	(1.000)		(0.093)	(0.074)	(0.082)	(1.000)	
Value of productive assets (NPR)	-8,413.267	-6,593.500	-7,876.230	57,899.582	1,053	-1,881.404	-8,588.923	-5,620.079	42,059.926	1,030
	(5,559.116)	(4,654.547)	(6,074.037)	(89,188.378)		(6,179.347)	(5,549.729)	(6,187.190)	(69,078.370)	
Value of non-productive assets (NPR)	-7,219.341	5,195.769	-5,523.035	179,682.491	1,053	-5,758.810	-406.836	-1,638.560	158,538.650	1,030
	(8,344.015)	(13,954.559)	(12,120.524)	(131,924.301)		(12,172.658)	(13,678.225)	(12,737.496)	(119,358.998)	
Land (hectares)	0.382	0.103	0.059	1.038	1,053	0.896**	0.067	0.105	0.820	1,030
	(0.344)	(0.353)	(0.382)	(2.943)		(0.393)	(0.211)	(0.236)	(2.503)	
Tropical livestock units	0.214**	-0.025	0.105	1.655	1,053	0.310***	-0.057	0.014	1.566	1,030
	(0.099)	(0.111)	(0.131)	(1.327)		(0.108)	(0.120)	(0.132)	(1.285)	

Potential PIF beneficiaries

	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Asset index	-0.057	-0.056	-0.107	-0.003	774	-0.087	-0.189*	-0.198*	-0.002	759
	(0.060)	(0.077)	(0.069)	(1.002)		(0.110)	(0.100)	(0.119)	(1.002)	
Value of productive assets (NPR)	-17,372.453*	-11,231.407	-10,340.696	58,551.458	774	-12,514.653*	-17,991.116***	-9,354.057	47,057.026	759
	(9,583.014)	(7,623.935)	(8,050.148)	(92,180.617)		(7,538.871)	(6,215.866)	(5,865.913)	(78,112.557)	
Value of non-productive assets (NPR)	-2,868.254	-7,572.944	-2,949.400	165,654.219	774	-2,768.822	-14,395.529	-10,469.954	149,517.567	759
	(11,123.400)	(9,486.818)	(10,101.177)	(133,872.960)		(10,472.617)	(9,426.299)	(9,856.840)	(119,606.060)	
Land (hectares)	-0.103	0.087	-0.173	0.904	773	-0.350	-0.527	-0.621	1.472	758
	(0.245)	(0.242)	(0.249)	(2.626)		(0.518)	(0.487)	(0.580)	(4.623)	
Tropical livestock units	0.112	-0.037	-0.049	1.542	774	0.086	0.126	-0.089	1.461	759
	(0.111)	(0.135)	(0.140)	(1.163)		(0.132)	(0.145)	(0.163)	(1.217)	

FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Treatment effects on household income

Direct beneficiaries

	Endline 1 (2.5 years post-treatment)					Endline 2 (3.5 years post treatment)				
	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Total household income	-4,466	57,279**	-27,465	301,133	1,047	-3,550	13,140	-7,344	324,774	1,014
	(23,269)	(28,237)	(22,087)	(302,309)		(29,653)	(29,769)	(22,590)	(339,304)	
Livestock income (net)	832	3,513	1,978	9,827	1,053	2,574	6,247**	299	12,592	1,030
	(1,675)	(2,429)	(2,130)	(18,667)		(2,444)	(2,706)	(2,174)	(21,532)	
Crop income (net)	179	-6,354	2,989	41,502	1,053	-171	-6,924	8,304	40,900	1,030
	(4,231)	(4,582)	(4,930)	(43,875)		(5,129)	(5,752)	(6,243)	(44,109)	
Business income	12,222**	8,069	-1,614	4,368	1,053	3,053	2,036	-1,584	8,235	1,030
	(5,691)	(5,128)	(3,802)	(64,264)		(3,627)	(5,392)	(3,816)	(46,490)	
Day labor income	5,480	3,182	7,143	18,438	1,047	12,751	5,945	5,530	29,450	1,014
	(4,311)	(4,744)	(5,970)	(47,523)		(9,004)	(9,585)	(10,143)	(86,722)	
Salary income	188	28,923	-17,822	52,328	1,047	12,507	-2,075	786	58,494	1,014
	(15,575)	(20,903)	(14,280)	(149,774)		(15,970)	(14,559)	(16,355)	(145,071)	
Other income	-169	2,015	6,458	32,058	1,047	115	-385	-4,560	33,912	1,014
	(10,419)	(9,311)	(10,222)	(90,392)		(10,054)	(7,385)	(10,504)	(84,283)	
Remittance income	-27,511	8,345	-29,303	142,624	1,047	-36,643**	1,862	-16,063	139,986	1,014
	(19,418)	(21,190)	(20,216)	(229,648)		(17,962)	(21,462)	(16,871)	(254,595)	

Potential PIF beneficiaries

	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Total household income	-40,319*	-14,961	-12,201	291,467	763	11,307	-26,547	28,022	274,288	744
	(24,110)	(24,501)	(23,397)	(345,415)		(23,901)	(23,992)	(21,440)	(251,902)	
Livestock income (net)	419	5,417***	-970	9,605	774	-1,251	1,092	-1,114	12,763	759
	(1,749)	(1,885)	(1,750)	(19,445)		(2,072)	(2,299)	(1,983)	(24,901)	
Crop income (net)	-5,065	-10,525***	-7,270*	43,504	774	2,584	-6,711**	1,086	40,416	759
	(4,320)	(3,814)	(3,879)	(50,564)		(3,691)	(2,936)	(4,403)	(44,504)	
Business income	7,766	11,975**	5,741	-2,819	774	-1,657	-1,265	-2,710	7,099	759
	(5,270)	(4,991)	(4,404)	(74,091)		(3,272)	(3,110)	(3,717)	(62,220)	
Day labor income	12,582*	12,031*	6,499	19,856	763	37,590***	5,006	22,856***	20,779	744
	(7,321)	(6,152)	(6,085)	(56,338)		(14,324)	(6,711)	(7,861)	(55,714)	
Salary income	-1,272	13,168	22,558*	38,738	763	-2,541	-9,442	2,180	48,550	744
	(11,012)	(15,408)	(13,523)	(101,328)		(11,071)	(13,509)	(12,407)	(140,878)	
Other income	-7,421	-8,794	-3,477	28,815	763	-2,668	-7,954	-561	28,636	744
	(5,933)	(6,348)	(6,941)	(87,720)		(7,471)	(6,599)	(7,329)	(81,627)	
Remittance income	-45,942**	-42,058**	-36,346	153,254	763	-19,670	-7,423	2,357	114,914	744
	(19,192)	(20,126)	(24,414)	(260,385)		(15,523)	(17,000)	(20,546)	(175,465)	

All values are in NPR. FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Sources of income at Endline 1

	Proportion earning	Average amount (net)
Total income	1.000	284,889
	(0.000)	(294,241)
Remittances	0.589	126,513
	(0.492)	(206,236)
Salary	0.216	50,876
	(0.412)	(142,800)
Crops	0.904	36,380
	(0.295)	(38,946)
Other	0.346	28,865
	(0.476)	(86,641)
Labor	0.308	24,262
	(0.462)	(57,930)
Business	0.177	6,449
	(0.382)	(61,478)
Livestock	0.760	11,198
	(0.427)	(19,706)
Goats	0.543	6,193
	(0.498)	(12,801)
Buffalo	0.294	4,068
	(0.456)	(13,446)
Cows	0.079	234
	(0.269)	(5,348)
Pigs	0.142	1,471
	(0.349)	(5,453)
Chickens	0.290	937
	(0.454)	(2,924)

Averages are in NRP and include zero values for those without that type of debt. Standard deviations in parentheses. Averages include original and PIF beneficiaries in all treatment groups at Endline 1.

Table 11: Treatment effects on women's empowerment

Direct beneficiaries

	Endline 1 (2.5 years post-treatment)					Endline 2 (3.5 years post treatment)				
	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Women's empowerment index	0.328***	0.195***	0.242***	-0.000	1,052	0.213**	0.019	0.052	-0.000	1,030
	(0.075)	(0.075)	(0.070)	(1.000)		(0.088)	(0.121)	(0.089)	(1.000)	
Access to and control over credit	0.033	-0.054	-0.015	0.622	1,052	-0.011	-0.146**	-0.066	0.592	1,030
	(0.044)	(0.046)	(0.040)	(0.486)		(0.037)	(0.066)	(0.041)	(0.492)	
Leadership position	0.070**	0.005	0.102***	0.156	1,051	0.009	-0.020	0.015	0.169	1,030
	(0.027)	(0.028)	(0.032)	(0.364)		(0.030)	(0.031)	(0.025)	(0.376)	
Group membership	0.158***	0.153***	0.109***	0.753	1,052	0.151***	0.118**	0.062	0.750	1,030
	(0.036)	(0.032)	(0.034)	(0.432)		(0.045)	(0.053)	(0.050)	(0.434)	

Potential PIF beneficiaries

	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Women's empowerment index	0.136	0.167*	-0.105	0.002	771	-0.048	-0.084	-0.298***	-0.002	759
	(0.091)	(0.092)	(0.102)	(1.002)		(0.066)	(0.064)	(0.090)	(1.002)	
Access to and control over credit	-0.074	-0.060	-0.150***	0.641	771	-0.093*	-0.088*	-0.151***	0.585	759
	(0.058)	(0.059)	(0.052)	(0.481)		(0.053)	(0.047)	(0.052)	(0.494)	
Leadership position	0.044	0.068**	0.004	0.141	771	-0.036	-0.045	-0.036	0.169	759
	(0.028)	(0.031)	(0.026)	(0.349)		(0.033)	(0.031)	(0.030)	(0.376)	
Group membership	0.113***	0.098**	0.027	0.714	771	0.105***	0.053	-0.071	0.769	759
	(0.043)	(0.045)	(0.052)	(0.453)		(0.037)	(0.032)	(0.043)	(0.422)	

FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 12: Sources of credit at baseline

	Proportion with	Average amount
Total	0.553	75,851
	(0.497)	(146,036)
Any formal	0.181	27,390
	(0.385)	(101,672)
Bank	0.049	18,432
	(0.217)	(147,724)
MFI	0.004	266
	(0.063)	(5,164)
Coop	0.041	6,450
	(0.198)	(54,669)
Village C and S	0.089	9,810
	(0.285)	(105,121)
Any informal	0.380	48,460
	(0.486)	(114,998)
Finance company	0.007	717
	(0.084)	(10,598)
Family and friends	0.168	20,896
	(0.374)	(94,143)
Moneylender	0.014	654
	(0.116)	(10,505)
Shop	0.197	34,031
	(0.398)	(161,403)

Averages are in NRP and include zero values for those without that type of debt. Standard deviations in parentheses. Averages include original and PIF beneficiaries in all treatment groups at baseline.

Table 13: Treatment effects on financial inclusion

Direct beneficiaries

	Endline 1 (2.5 years post-treatment)					Endline 2 (3.5 years post treatment)				
	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Financial inclusion index	0.262**	0.274***	0.193*	0.000	1,053	0.311***	0.249**	0.152	0.000	1,030
	(0.108)	(0.091)	(0.111)	(1.000)		(0.117)	(0.104)	(0.122)	(1.000)	
Savings group member	0.230***	0.204***	0.134**	0.655	1,053	0.129**	0.139**	0.083	0.684	1,030
	(0.050)	(0.050)	(0.056)	(0.476)		(0.061)	(0.056)	(0.054)	(0.466)	
Saved anything last month	0.137***	0.070**	0.064	0.804	1,053	0.129***	0.094**	0.081*	0.790	1,030
	(0.035)	(0.035)	(0.040)	(0.398)		(0.035)	(0.040)	(0.044)	(0.408)	
Amount saved last month	-193.815	383.265	3.934	1,089.836	1,053	583.580	129.426	-102.850	1,451.544	1,030
	(363.558)	(345.031)	(353.598)	(3,686.458)		(362.253)	(313.779)	(396.063)	(4,070.246)	
Household has savings	0.043*	0.034	0.001	0.938	1,053	0.056*	0.044	0.038	0.919	1,030
	(0.023)	(0.024)	(0.025)	(0.241)		(0.029)	(0.029)	(0.032)	(0.273)	
Amount in household savings	-10,759.669	19,549.027*	5,192.267	62,700.655	1,053	3,612.922	15,460.486	-7,446.579	94,221.434	1,030
	(8,679.194)	(11,580.008)	(10,026.716)	(104,252.275)		(12,798.827)	(16,255.406)	(12,624.248)	(158,993.354)	
Respondent has savings	0.067**	0.092***	0.057**	0.880	1,053	0.070*	0.055	0.049	0.864	1,030
	(0.027)	(0.027)	(0.027)	(0.326)		(0.040)	(0.038)	(0.044)	(0.343)	
Amount in respondent savings	-2,227.667	5,248.475	-614.915	22,174.800	1,053	7,641.389**	10,975.705***	1,310.809	20,692.188	1,030
	(3,377.627)	(3,707.524)	(3,304.039)	(37,292.969)		(3,209.948)	(3,427.419)	(3,593.195)	(32,563.261)	
Has formal loan	0.083*	-0.034	0.048	0.349	1,053	0.064	-0.010	-0.000	0.357	1,030
	(0.047)	(0.047)	(0.057)	(0.478)		(0.061)	(0.070)	(0.067)	(0.480)	
Has informal loan	-0.043	-0.037	-0.061	0.411	1,053	-0.053	-0.118**	-0.072	0.390	1,030
	(0.055)	(0.047)	(0.053)	(0.493)		(0.048)	(0.050)	(0.044)	(0.489)	
Percent of debt from formal loans	0.095**	-0.009	0.064	0.301	1,053	0.058	0.007	0.019	0.296	1,030
	(0.045)	(0.041)	(0.058)	(0.446)		(0.045)	(0.058)	(0.056)	(0.431)	

Potential PIF beneficiaries

	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Financial inclusion index	0.298***	0.169**	0.105	-0.000	774	0.043	-0.068	-0.194*	0.004	759
	(0.082)	(0.081)	(0.100)	(1.003)		(0.082)	(0.079)	(0.104)	(1.001)	
Savings group member	0.200***	0.143***	0.099	0.583	774	0.109***	0.045	-0.127**	0.697	759
	(0.055)	(0.055)	(0.066)	(0.494)		(0.041)	(0.043)	(0.058)	(0.461)	
Saved anything last month	0.144***	0.089*	0.008	0.760	774	0.080**	0.065**	-0.049	0.779	759
	(0.048)	(0.050)	(0.060)	(0.428)		(0.036)	(0.031)	(0.047)	(0.416)	
Amount saved last month	-3.144	-367.561	-310.352	1,260.536	774	-496.090	-817.451*	-785.803*	1,566.949	759
	(405.528)	(343.796)	(319.647)	(4,299.192)		(462.652)	(417.227)	(450.251)	(4,802.689)	
Household has savings	0.072***	0.064***	0.043	0.891	774	0.031	0.043	-0.024	0.923	759
	(0.025)	(0.024)	(0.031)	(0.313)		(0.024)	(0.030)	(0.034)	(0.267)	
Amount in household savings	-18,630.131	-7,118.810	-13,142.659	81,147.266	774	2,036.115	-20,258.724	-20,265.198**	101,040.769	759
	(12,698.443)	(12,651.216)	(11,373.540)	(149,610.953)		(13,349.284)	(12,660.246)	(9,671.180)	(168,913.881)	
Respondent has savings	0.098**	0.054	-0.001	0.854	774	0.008	0.001	-0.062	0.913	759
	(0.040)	(0.041)	(0.045)	(0.354)		(0.026)	(0.031)	(0.038)	(0.283)	
Amount in respondent savings	-2,477.110	3,607.673	836.580	21,411.953	774	6,212.420**	5,003.679	-3,366.687	23,191.692	759
	(2,475.115)	(3,569.547)	(2,684.611)	(34,044.948)		(2,774.297)	(3,696.756)	(2,909.478)	(40,907.600)	
Has formal loan	0.101*	0.023	0.005	0.266	774	-0.072*	-0.068**	-0.104*	0.328	759
	(0.057)	(0.050)	(0.062)	(0.443)		(0.037)	(0.033)	(0.060)	(0.471)	
Has informal loan	-0.090	-0.055	-0.099*	0.432	774	-0.129**	-0.085*	-0.113**	0.431	759
	(0.059)	(0.056)	(0.057)	(0.497)		(0.056)	(0.046)	(0.049)	(0.496)	
Percent of debt from formal loans	0.064	0.005	0.016	0.238	774	-0.056	-0.044	-0.078	0.275	759
	(0.052)	(0.045)	(0.058)	(0.415)		(0.035)	(0.033)	(0.065)	(0.420)	

FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 14: Treatment effects on psychological well-being

Direct beneficiaries

	Endline 1 (2.5 years post-treatment)					Endline 2 (3.5 years post treatment)				
	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Psychological well-being index	-0.094	-0.007	-0.151	-0.000	1,053	-0.048	-0.075	0.028	-0.000	1,030
	(0.111)	(0.115)	(0.107)	(1.000)		(0.115)	(0.097)	(0.117)	(1.000)	
Depression (CES-D score)	-0.376	-0.950**	0.221	8.927	1,053	-0.369	0.249	-0.646	9.824	1,030
	(0.483)	(0.462)	(0.431)	(4.033)		(0.394)	(0.390)	(0.412)	(4.186)	
Worries	0.269	-0.163	0.260	9.920	1,053	0.238	-0.236	0.191	10.103	1,030
	(0.302)	(0.245)	(0.275)	(2.469)		(0.241)	(0.217)	(0.275)	(2.432)	
Self-esteem	0.409*	0.120	0.347	13.876	1,053	-0.240	-0.340	-0.002	13.801	1,030
	(0.249)	(0.250)	(0.230)	(2.340)		(0.235)	(0.269)	(0.297)	(2.616)	
Life satisfaction	-0.022	-0.061	-0.015	7.196	1,053	-0.197	-0.273	0.035	7.312	1,030
	(0.223)	(0.194)	(0.206)	(1.810)		(0.180)	(0.181)	(0.194)	(1.690)	
Optimism	-0.305	-0.201	-0.358	12.258	1,053	-0.023	0.054	0.193	11.820	1,030
	(0.255)	(0.283)	(0.225)	(2.318)		(0.198)	(0.240)	(0.264)	(2.220)	
Locus of control	-0.170	-0.074	-0.252*	3.007	1,053	0.033	-0.072	0.002	2.603	1,030
	(0.105)	(0.144)	(0.132)	(1.318)		(0.124)	(0.100)	(0.110)	(1.276)	

Potential PIF beneficiaries

	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Psychological well-being index	-0.080	0.139	-0.138	0.004	774	-0.085	0.060	0.020	-0.008	759
	(0.117)	(0.141)	(0.122)	(1.001)		(0.115)	(0.091)	(0.097)	(0.997)	
Depression (CES-D score)	-0.052	-1.354***	-0.527	9.594	773	-0.214	0.111	-0.015	9.728	758
	(0.410)	(0.474)	(0.488)	(4.317)		(0.463)	(0.357)	(0.304)	(4.300)	
Worries	0.172	-0.482**	0.003	10.349	773	0.240	-0.652***	-0.266	10.328	758
	(0.280)	(0.219)	(0.223)	(2.527)		(0.264)	(0.248)	(0.292)	(2.607)	
Self-esteem	0.254	0.415*	0.308	13.542	773	0.149	0.163	0.007	13.467	758
	(0.224)	(0.229)	(0.231)	(2.504)		(0.295)	(0.310)	(0.314)	(2.938)	
Life satisfaction	-0.015	-0.294	-0.140	7.234	773	-0.038	-0.129	0.261	7.046	758
	(0.259)	(0.256)	(0.233)	(1.854)		(0.228)	(0.202)	(0.186)	(1.873)	
Optimism	-0.189	0.283	-0.091	12.010	774	-0.067	-0.218	-0.213	11.949	758
	(0.292)	(0.335)	(0.283)	(2.263)		(0.197)	(0.284)	(0.257)	(2.203)	
Locus of control	-0.081	-0.103	-0.414***	3.026	774	-0.244*	-0.029	-0.213*	2.672	759
	(0.117)	(0.151)	(0.158)	(1.238)		(0.140)	(0.143)	(0.127)	(1.266)	

FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 15: Cost-benefit analysis (10% discount rate)

Costs	*FT*	*NG*	*NVT*
Cost per direct	68,676	36,342	45,836
Cost per PIF	6,290	5,915	5,533
Cost per beneficiary	17,931	12,176	12,502
Benefits			
Direct total benefits into perpetuity	66,516	66,516	8,838
PIF Total benefits into perpetuity	58,597	58,597	26,422
Average benefits into perpetuity	59,917	59,917	23,491
Net present value (NPV)			
Direct NPV	-24,893	18,144	-52,169
PIF NPV	50,225	50,725	19,057
Average NPV	**36,050**	**43,711**	**6,851**
Benefit Cost Ratio (BCR)			
Direct BCR	0.728	1.375	0.145
PIF BCR	6.999	7.443	3.588
Average BCR	**2.510**	**3.697**	**1.412**

Appendix A Recruiting Rates

In section 4.4 we argue in favor of accepting beneficiaries' self-attested membership status in a Heifer group when calculating recruitment rates, even though Heifer provided membership rosters for each of the treated villages. However, it is still instructive to reconcile the self-reported membership from the evaluation sample with the Heifer rosters. While this will result in artificially low recruitment rates for all the reasons mentioned in section 4.4, it still sheds light on recruiting patterns across the eight treatment variants (FT vs. NG vs. NVT vs. C, and direct vs. PIF). In this appendix we describe the method we used to cross reference the evaluation sample to Heifer membership rosters, and comment on some of the observed patterns.

Heifer provided village-level membership rosters for each treated VDC included in our evaluation sample. These membership rosters included beneficiaries' first and last names in separate fields. The names were provided in the English alphabet, having been transliterated from the Devanagari script. The lists themselves are compiled by group, and thus distinguish between direct and PIF status. We carried out a combined automated/manual matching procedure to reconcile these names with respondents from endline 1 and endline 2.

The matching exercise proceeded in two stages, and each stage had two steps. In the first stage, we attempted to match the Heifer lists to each endline list using the Stata reclink command, which executes a fuzzy matching algorithm based on the Mahalanobis distance between strings. The reclink command generates a numeric match score on the 0,1 interval, where a score of one indicates a perfect match between two records on fields of interest (first name, last name, VDC name). Records were considered a potential match if the score generated by reclink exceeded 0.6. In our reclink specification, we required a perfect match on the VDC field (VDC spellings were cleaned in each set and made identical to facilitate perfect matches), and then placed greater weight on first names than on surnames, as surnames are frequently shared by many families in a village and are thus a less reliable characteristic to match on. Based on the cutoff criteria, we classified 794/1834 records as a potential match in endline 1, and 791/1794) records as a potential match in endline 2. In the second step, a native Nepali speaker visually inspected all the potential matches that did not have a perfect match score (i.e. the match score for the record was between 0.6, 0.99), and graded them as a match, likely match, possible match, or no match. Matches and likely matches were retained. In the second stage, we removed these verified beneficiaries from the original lists and repeated the first stage exactly.

This process yielded 613 individuals from treated villages in the evaluation sample affirmatively matched to a Heifer roster in endline round 1, and 599 in endline round 2. Using

only affirmatively matched observations results in considerably lower membership rates than those reported in section 4.4. For directly recruited beneficiaries, the overall "matched" membership rate is approximately 50% across all three treatments (controls excluded). In addition, the vast majority of these match to Heifer records as "direct" beneficiaries. For PIF beneficiaries the match rates are lower, as we could expect from priors and from self-reported membership. For the FT and NG groups, the matched membership rate is approximately 35% and again these individuals correctly match to Heifer records as "indirect" (PIF) beneficiaries. A little under 10% of the PIF/NVT group matches to rosters.

While these patterns mostly validate the experimental design and self-reported membership, there are two interesting an notable divergences. First, in the PIF/NVT group almost all of of the confirmed individuals matched to a roster of direct beneficiaries, which clearly should not be the case. This suggests the possibility that Heifer set up direct beneficiary groups in NVT villages, or allowed a number of PIF targets to join direct groups. Further inspection reveals that while this may have occurred, the scale of the problem is small and localized: only 18 individuals (out of 236 of this treatment), concentrated in three VDCs. Second, self-reported "direct" status among PIF beneficiaries is much lower in the verified set than it is in the self-reported set. This suggests that beneficiaries may not always a draw a distinction between the two group types, and also alleviates potential concern around noncompliance.

Appendix B Alternative Specifications

Table A1: Alternative specification of treatment effects on women's empowerment

Direct beneficiaries

	Endline 1 (2.5 years post-treatment)					Endline 2 (3.5 years post treatment)				
	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Women's empowerment index	0.140*	0.093	0.075	0.000	1,052	0.109	-0.088	-0.000	0.000	1,030
	(0.078)	(0.070)	(0.083)	(1.000)		(0.123)	(0.119)	(0.111)	(1.000)	
Empowerment over production decisions	-0.025	0.009	-0.025	0.924	1,052	0.018	-0.004	-0.007	0.890	1,030
	(0.020)	(0.016)	(0.019)	(0.266)		(0.025)	(0.034)	(0.023)	(0.314)	
Asset ownership	0.026*	0.016	-0.003	0.942	1,052	0.019	0.013	0.035	0.934	1,030
	(0.014)	(0.016)	(0.018)	(0.235)		(0.024)	(0.024)	(0.023)	(0.249)	
Access to and control over credit	0.028	-0.054	-0.027	0.622	1,052	-0.001	-0.146**	-0.057	0.592	1,030
	(0.044)	(0.044)	(0.039)	(0.486)		(0.040)	(0.065)	(0.043)	(0.492)	
Control over income	0.008	-0.000	-0.023	0.823	1,052	0.008	0.012	0.009	0.831	1,030
	(0.025)	(0.023)	(0.028)	(0.267)		(0.030)	(0.030)	(0.032)	(0.281)	
Leadership position	0.070**	0.005	0.102***	0.156	1,051	0.009	-0.020	0.015	0.169	1,030
	(0.027)	(0.028)	(0.032)	(0.364)		(0.030)	(0.031)	(0.025)	(0.376)	
Group membership	0.154***	0.149***	0.108***	0.753	1,052	0.148***	0.114**	0.061	0.750	1,030
	(0.036)	(0.032)	(0.033)	(0.432)		(0.045)	(0.052)	(0.050)	(0.434)	
Works less than 10.5 hours per day	-0.089	-0.059	-0.044	0.509	1,052	-0.057	-0.129**	-0.099	0.577	1,030
	(0.066)	(0.069)	(0.067)	(0.501)		(0.070)	(0.062)	(0.062)	(0.495)	

Potential PIF beneficiaries

	FT	NG	NVT	Control mean	N	FT	NG	NVT	Control mean	N
Women's empowerment index	-0.001	0.029	-0.217*	-0.001	771	0.054	-0.206***	-0.301***	0.001	759
	(0.095)	(0.096)	(0.117)	(1.002)		(0.082)	(0.080)	(0.088)	(1.002)	
Empowerment over production decisions	-0.037	-0.030	-0.131***	0.943	771	-0.050	-0.020	-0.094***	0.882	759
	(0.034)	(0.032)	(0.039)	(0.233)		(0.032)	(0.030)	(0.035)	(0.323)	
Asset ownership	-0.007	0.019	-0.012	0.938	771	0.033	0.024	-0.031	0.923	759
	(0.021)	(0.024)	(0.025)	(0.243)		(0.026)	(0.024)	(0.027)	(0.267)	
Access to and control over credit	-0.071	-0.052	-0.151***	0.641	771	-0.093*	-0.088*	-0.151***	0.585	759
	(0.058)	(0.059)	(0.053)	(0.481)		(0.053)	(0.047)	(0.052)	(0.494)	
Control over income	-0.031	-0.003	-0.046	0.822	771	0.038	0.024	0.000	0.803	758
	(0.032)	(0.027)	(0.036)	(0.265)		(0.031)	(0.028)	(0.035)	(0.312)	
Leadership position	0.044	0.068**	0.004	0.141	771	-0.036	-0.045	-0.036	0.169	759
	(0.028)	(0.031)	(0.026)	(0.349)		(0.033)	(0.031)	(0.030)	(0.376)	
Group membership	0.113***	0.098**	0.027	0.714	771	0.105***	0.053	-0.071	0.769	759
	(0.043)	(0.045)	(0.052)	(0.453)		(0.037)	(0.032)	(0.043)	(0.422)	
Works less than 10.5 hours per day	-0.036	-0.076	-0.044	0.490	771	0.048	-0.160***	-0.061	0.579	759
	(0.066)	(0.069)	(0.062)	(0.501)		(0.049)	(0.058)	(0.054)	(0.495)	

FT: full treatment, NG: no-goats treatment, NVT: no-values-based-training treatment. Columns 1-3: Estimates with clustered (VDC) standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.