Goalpost Effects: Understanding Impact Heterogeneity in Nicaragua's Rural Business Development Program

Stephen Boucher, Jonathan Malacarne, and Michael R Carter

Abstract

As evidence mounts for heterogeneous treatment effects in entrepreneurial interventions, program implementers look for ways to tilt the balance in favor of cost effectiveness. One such strategy is targeting on observable characteristics believed to be associated with positive program benefits. In this paper we show that such a strategy carries significant risks and may exclude households who stand to gain significant benefit from the intervention. The risks of sharp targeting are especially salient for multifaceted entrepreneurial programs, which benefit participants through multiple channels. Using a model that links a household's baseline characteristics to the participation decision and benefits of program participation, we highlight that for sharp targeting to be effective two conditions must hold. First, program implementers must know and observe a set of characteristics that are predictive of success. Second, it must be the case that households are not already self-selecting optimally into the program. These are strong restrictions and ones not likely satisfied in the majority of entrepreneurial interventions. We draw attention to these challenges by analyzing heterogeneous impacts in a Millennium Challenge Corporation Rural Business Development program in Nicaragua.

1 Introduction

The popularity of entrepreneurial interventions among policy makers stems from the belief that brief, intensive infusions of human and financial capital can loosen constraints to fruitful self-employment and place participants on higher paths of long run productivity. Business training programs, in particular, have become increasingly common in recent years in both developed and developing countries. At the moment, the World Bank alone has eighty-five active vocational training programs worldwide, with an additional thirteen in the pipeline (*World Bank Group* 2017). Despite being touted as having the potential to generate large returns at moderate cost, many evaluations of business training programs fail to find the anticipated impacts—especially over longer time horizons (Blattman and Ralston, 2015). This has caused concern among the organizations that fund and implement such interventions. It has also been seized on by supporters of programs that compete with entrepreneurial interventions for funding.

Evaluating the impact of entrepreneurial interventions, and comparing the returns of such programs to their alternatives, is complicated by the participation decision. In entrepreneurial interventions, program implementers neither desire nor expect universal participation from the population to which the intervention is offered. This is due to the fact that not all eligible households have the requisite abilities, nor access to the resources, required for success as entrepreneurs. Optimal program benefits are thus achieved not when uptake is universal, but rather when the program is accessible to and utilized by the set of households for whom participation will most likely translate to improvements in income. Inducing the right participants to join entrepreneurial interventions is thus an important part of getting maximum returns from a limited budget.

Program implementers have only a few tools to influence the composition of an intervention. The first tool is to fully inform potential participants of the nature of the intervention and, to the best of their knowledge, the requirements for success. Relying on informed self-selection does not require any detailed information about potential participants, only that program implementers know which household characteristics will translate into program benefits. It also requires, however, that potential participants take the implementers assessment as true. If self-selection can not be induced, sharper targeting efforts can be undertaken by setting eligibility criteria. While this tool can be used to target an intervention at a certain segment of the population, the requirements that can be used are limited to a small number of easily observable household characteristics. Going beyond this set, or even verifying such common characteristics as land holding or tenure security, can be extremely costly and quickly erode the efficiency gains of targeting. Targeting efforts are further complicated by the potential for the endogenous enhancement of household characteristics likely to influence the benefit of participation. While this is not likely to be the case with slow-to-change assets like landholding, it is much more plausible for knowledge based assets like technical efficiency and traditional constraints such as access to financial services.

In this paper, we consider the case for sharper targeting as implied by a simple model in which a household possess a suite of assets that interact with the educational, technical, and financial resources made available through entrepreneurial interventions. These assets determine the likely benefits that a household obtains from participating in an entrepreneurial intervention. To the extent that households know their stock of the assets and understand how they map to the benefits of participation, the asset stock may also affect the participation decision. The mounting body of evidence for heterogeneous treatment effects and the oft observed null average treatment effects in entrepreneurial interventions would suggest that households do not optimally select into entrepreneurial interventions. This would also suggest that there may be large programmatic efficiency gains to be made from sharper targeting.

We explore the need for and feasibility of sharper targeting in the context of a Millennium Challenge Corporation rural business development program in Nicaragua. While there is some evidence that both selection and program impacts are related to baseline stocks of household assets, the effects are non-monotonic. Rather, they might be best described as goalpost effects: households with low asset levels and high asset levels derive the greatest benefits from the program. This pattern implies both benefits and risks from pursuing a sharper targeting strategy and is consistent with an intervention that both complements existing household characteristics and builds household capacity.

2 Entrepreneurial Interventions

Evaluations of entrepreneurial interventions—and business training programs specifically—are taking place across the world, both in developing and developed countries. In a survey of impact evaluations of business training programs across developing countries, McKenzie and Woodruff (2014) report that programs generally focus on small, existing, urban businesses and last between two days and a few weeks. All the programs evaluated in the paper included classroom-based coursework, and some included additional follow-up training or individual visits. The content of the classroom training tends to focus on basic business skills and best practices, such as keeping financial records and separating household finances from business finances. More in depth interventions may include support in creating business plans and matching grants (Carter, Tjernström, and Toledo, 2016), or may use the training to augment access to financial services (Karlan and Valdivia, 2011).

While the contexts differ substantially, the conclusions coming out of the evaluations are broadly consistent— noisy estimates that fail to fully satisfy researchers or to find the expected positive impacts. In one large study in the United States, Fairlie et al (2015) consider common outcome dimensions for business training programs [1] and only find empirical support for business ownership among those households unemployed at baseline, and even then only along certain time horizons. This is consistent with the general consensus of the fourteen programs surveyed by McKenzie and Woodruff (2014). Specifically, the authors find that most evaluations show small positive effects for new business creation and survival rates, but little impact on sales or profits.

[1] Business startup, employment, business performance, household income, and work satisfaction

There are a number of reasons studies may fail to find the expected impacts. The small sample size of most evaluations and heterogeneous treatment population result in chronically low power in the majority of studies (David McKenzie and Christopher Woodruff, 2014). Carter, Tjernstrom and Toledo (2016), as well as McKenzie and Woodruff (2014) and de Mel et al. (2014) note that the time horizon at which impacts are evaluated is also important. While proponents of entrepreneurial interventions hope that the programs bump participants to higher long run income paths, the mediocre observed results are even less convincing as the time horizon increases [2].

Even the observed null impacts, however, are provocative. Estimates of program impacts tend to be noisy, and some participants do seem to derive large benefits from participation. Systematic analysis of such heterogeneity, though, has proven difficult. Due to the same small sample issues discussed by McKenzie and Woodruff (2014), most attempts to tease out heterogeneous impacts along observable dimensions have met with little success (see Fairlie et al 2015). Carter, Tjernstrom, and Toledo (2016), as well as Karlan and Valdivia (2011) and Bjorvatn and Tungodden (2010), raise the point and provide suggestive evidence that the dimensions that are most important to explain impacts may be unobservable, and that participants are almost certainly heterogeneous along these dimensions.

As noted above, one important feature of entrepreneurial interventions is they neither expect nor desire universal participation. Across the studies surveyed by McKenzie and Woodruff (2014), participation rates ranged from 39% - 92%, with an average of 65%. Clearly, potential participants are forming expectations both about themselves and about the opportunities offered by the intervention, which they then use to motivate their participation decision. Who chooses to participate is important for both programmers and participants. For programmers, the cost of supporting a single household are substantial. Karlan and Valdivia (2011), for example, report a per participant cost of about $1,300. The per participant cost of the program detailed in the later part of this paper was about $2,500 (Carter, Tjernström, and Toledo, 2016).

For participants as well, the decision to join a business training program is an investment. Depending on the program, the participating household may be committing substantial time and resources to the program. Both Karlan and Valdivia (2011) and Bjorvatn and Tungodden (2010) suggest that those who are most likely to benefit from participation are among the least likely to participate in the intervention. Specifically, Bjorvatn and Tungodden point to the result that households with lower education levels experienced larger benefits but were less likely to participate. They hypothesize that these households perceived themselves as weaker entrepreneurs and possessed a "knowledge gap" that discouraged participation (Bjorvatn and Tungodden, 2010).

The effect of this "knowledge gap", and similar factors that may both complement the intervention and be endogenously affected by it, complicate the notion of targeting an entrepreneurial intervention. While simple interventions might focus on relaxing credit constraints or providing classroom knowledge—dimensions for which program benefits

[2]Note that Tjernstrom, Toledo, Carter do find increasing impacts as time progresses, especially for investment. There are, however, no significant effects on consumption even at the longest time horizon.

may be expected to increase monotonically with the severity of the constraint—more complex interventions undertake to build technical skills, create and strengthen market linkages, and tailor support to the individual needs of a participant. In such cases, two mechanisms exist for a single household characteristic to map into program impact: a complementary mechanism and an endogenous formation mechanism. These mechanisms may interact to create a non-monotonic mapping of household characteristics into program impacts, which makes sharp targeting difficult. The problem compounds as the complexity of the intervention increases.

3 Who Participates and Who Benefits?

Imagine, for the moment, that household income is driven completely by a single asset. Further, imagine that the costs of program participation occur only once and immediately. Let a household's endowment of the asset be given by α_i. In a general form, household income (y_i) and the gains of participation for household i (I_i) can defined as follows. Let household income without the program be given by:

$$y_{iN} = f_N(\alpha_{iN})$$

and let household income with the program be given by:

$$y_{iT} = f_T(\alpha_{iN}) - g(\alpha_{iN})$$

where $g(\alpha_{iN})$ is the cost of participating in the program. The gains of participation to household i are then:

$$I_i = y_{iT} - y_{iN}$$

$$I_i = f_T(\alpha_{iN}) - f_N(\alpha_{iN}) - g(\alpha_{iN}) \tag{1}$$

We can add some simple structure to the functions $f()$ and $g()$ in order to demonstrate the role of α_i. Let:

$$f_N(\alpha_{iN}) = F_N \cdot \alpha_{iN}$$

$$f_T(\alpha_{iN}) = F_T \cdot \alpha_{iN}$$

$$g(\alpha_{iN}) = G$$

Which allows us to express Equation 1 as:

$$I_i = (F_T - F_N) \cdot \alpha_i - G$$

This simple but instructive example is illustrated graphically in Figure 1.

In this simple world, a household that understands the nature of $f()$ and $g()$ would choose to participate in the program only if α_i is greater than $\underline{\alpha} = \frac{G}{(F_T - F_N)}$. From a programmatic efficiency perspective, this is also the desirable result. If α is distributed in the population of potential participants according to some distribution $\phi(\alpha)$, self-selection of this type would generate both non-universal participation and heterogeneous impacts across household types.

A more realistic model quickly becomes complicated. We first acknowledge that both costs and benefits do not occur only once, but rather as flows into the future. Carter, Tjernstrom, and Toledo (2016), in evaluating the program we consider later, find support for this characteristic. At the time they influence the participation decision, however, both $f()$ and $g()$ are discounted present values of the expected costs and benefits of program participation. Added to these uncertain and time-dependent characteristics of the problem is the fact that there are likely numerous releveant asset dimensions. It remains true, however, that the necessity of targeting and the ability to target depend on identifying the relevant set of assets for a specific intervention and the degree to which households already self-select. If we can identify the assets that determine success but households do not self-select, there is room to improve programmatic efficiency and the welfare effects on participating households.

Figure 2 depicts three possible self-selection scenarios. In Scenario 1, represented by the solid, kinked, line, households perfectly self-select. Households with asset levels below $\underline{\alpha}$ participate with probability zero while households with asset levels above $\underline{\alpha}$ participate with probability equal to one. The dashed line of Scenario 2 captures a more nuanced perfect self-selection. Households still select into the program based on their asset level, but owing to different risk and time preferences, there is not a sharp break at $\underline{\alpha}$. Scenario 3 (horizontal, solid line) represents the effective opposite of Scenario 1. Instead of self-selecting based on α_i, households participate with probability P regardless of their asset level. In addition to these polar cases, many degrees of positive self-selection are possible. The space between Scenarios 2 and 3 can be thought of as populated by these lesser degrees of positive self-selection. Such scenarios would result from an imperfect understanding about the relationship between α_i and I_i.

Equation 2 decomposes the total benefit-cost ratio by a household's asset position relative to the "zero net benefit" point $\underline{\alpha}$–where I is total social benefit, C is total program cost, I_i is individual household benefit, and c is cost-per-participant. If only households with levels of the complementary asset above $\underline{\alpha}$ participate in the intervention, the first bin in Equation 2 is empty. As households in this bin have zero or negative expected returns to participation in the program, this is a desirable outcome for both the household and the programmers. The money saved through this sorting can be redirected to programs better able to serve these households. In fact, the optimal cutoff $\underline{\alpha}$ would not be the "zero net benefit" point, but rather the point where participation in the entrepreneurial intervention had a higher return on investment than the next best programmatic option.

$$\frac{I}{C} = \sum_{\alpha=\alpha_L}^{\underline{\alpha}} \frac{I_i}{c} + \sum_{\alpha=\underline{\alpha}}^{\alpha_H} \frac{I_i}{c} \tag{2}$$

The story presented in this section suggests that the null average treatment effect often observed in entrepreneurial interventions may be related to the existence of household assets that complement the intervention and a failure of targeting or household self-selection. Further, the complementary asset story would help explain the observed heterogeneity of impact in programs that do report positive impacts. The temptation is, then, to say that through better targeting both programmatic efficiency and household welfare outcomes could be improved. While these are, in fact, the implications of the story we present, they come with a strong word of caution. These implications rely on both the existence of a small number of complementary assets and the ability of the programmers to identify and observe the relevant assets prior to enrolling households in the intervention. Anyone who has worked in implementing or evaluating an entrepreneurial intervention will tell you this is a very, very tall order. We agree. While the temptation to target is strong, the risks are considerable even in this stylized example. Misidentifying α_i or the break-even threshold $\underline{\alpha}$ would result in the exclusion households who could gain from program participating or the inclusion of households for which participating is not likely to be profitable.

4 Multiple Impact Mechanisms

While the story told in Equation 1 and Figure 1 has some appeal and reflects characteristics we observe in evaluations of entrepreneurial interventions, it would not be consistent with assertions like those of Karlan and Valdivia or Bjorvatn and Tundgodden– that those households most likely to benefit are the least likely to participate. The mapping from household characteristics to program impacts these authors tell is one of an intervention building assets rather than complementing existing assets. In this case, it is those households with initially low levels of an asset who stand to gain the most from participation. These two stories need not be mutually exclusive. We can demonstrate this multiple mechanism scenario with slight adjustments to Equation 1. Let household income with (T) and without (N) the program be given by:

$$y_{iN} = f_N(\alpha_{iN})$$

and

$$y_{iT} = f_T(\alpha_{iT}) - g(\alpha_{iT})$$

Note that a household's level of the asset α_i is now affected by participating in the program. Program impacts for household i are once again defined as:

$$\begin{aligned} I_i &= y_{iT} - y_{iN} \\ I_i &= f_T(\alpha_{iT}) - f_N(\alpha_{iN}) - g(\alpha_{iT}) \end{aligned}$$

Which, by adding and subtracting $f_N(\alpha_{iT})$ we can now decompose into a complementary effect mechanism and an endogenous build-up mechanism as follows:

$$\begin{aligned} I_i &= f_T(\alpha_{iT}) - f_N(\alpha_{iN}) - g(\alpha_{iT}) \\ I_i &= f_T(\alpha_{iT}) - f_N(\alpha_{iT}) + f_N(\alpha_{iT}) - f_N(\alpha_{iN}) - g(\alpha_{iT}) \end{aligned} \quad (3)$$

Where the impact of the complementary mechanism is given by:

$$Comp_i = f_T(\alpha_{iT}) - f_N(\alpha_{iT})$$

and the impact of the endogenous build-up of the asset α_i obtained by households participating in the program is:

$$Endog_i = f_N(\alpha_{iT}) - f_N(\alpha_{iN})$$

There are many ways to characterize the transition from α_{i0} to α_{iT}. It can easily be thought of as a partial adjustment process in which α_{iT} is a function of α_{iN}. In other circumstances, such as considering a household's access to credit or land tenure security, it might be better to think of α_i as an indicator variable which may be affected by program participation. Regardless of how it is operationalized, adding this second effect dimension significantly complicates the relationship between program participation and household impacts. Heterogeneous impacts are still present in Equation 3 as long as the two effects do not trade off perfectly. Now, however, knowing which effect dominates is necessary in order to know whether it is households with high or low levels of α_i which can be expected to benefit most from program participation. From the point of view of a program implementer seeking to target an entrepreneurial intervention to those households most likely to benefit, this is a significant hurdle.

Earlier, we claimed that targeting was possible only if implementers could identify and observe a set of assets that predicted success in a program. The multiple mechanism extension of the model adds an additional wrinkle. Recall, however, that targeting is only necessary if households are not already self-selecting optimally into the program. We turn now to an example. Using data from a rural business development program in Nicaragua, we will consider the necessity and feasibility of sharper targeting. The necessity of targeting will be based on the criteria of sub-optimal self-selection, while feasibility will be evaluated by the presence of (ex-ante) identifiable, observable household characteristics predictive of success.

5 A Rural Business Development Program in Nicaragua

5.1 Program Description

Beginning in 2007, the Millennium Challenge Corporation (MCC) undertook a multi-faceted rural business development (RBD) program in western Nicaragua. The departments of Chinandega and Leon were selected by MCC and their partners as areas with high growth potential. This potential stems from the region's proximity to markets and centers of human capital development, ports, and transportation infrastructure. At the same time, the region has high levels of extreme poverty and large numbers of small farmers who are not currently integrated into value chains.

The rural business development project was one component of a larger compact, which included a transportation project and a property regularization project. In 2009, due to local conditions inconsistent with MCC eligibility criteria, the transportation and property regularization projects were suspended, but the rural business development project continued. Its mission and primary objectives were to contribute to the growth of income and employment in the region through the active support of small and medium-sized rural businesses and agricultural enterprises. The project also sought to eliminate restrictions to growth and promote investment and the integration of rural businesses into value chains (Millenium Challenge Corporation, 2009).

The RBD project itself was divided into various categories, each focusing on an agricultural activity. Five of these (Livestock, beans, cassava, sesame, and horticulture crops) are represented in the data from the impact evaluation. Due to the small number of horticulture projects, we focus on the other four program activities. For each agricultural activity, MCC identified a set of challenges and opportunities facing agricultural businesses. For example, market access and low yields were identified as primary challenges and points of improvement for dairy farmers in the region. Similarly, sesame farmers were determined to suffer from poor seed quality and receive disadvantageous prices for their product. The support offered by the program then focused on taking advantage of these opportunities and addressing the challenges.

To be eligible for the program, households must own a small or medium sized farm, have some experience in the program crop, and be willing to develop a business plan with the assistance of extension agents. Households also provided 70% of the investment costs set forth in the business plan, and had to meet certain asset thresholds (see Table 1). The business plan developed was to contain a description of the activity to be undertaken over the two year life of the program, as well as a cost-benefit analysis, market analysis, and expected environmental impact. It then detailed the necessary investment and the support requested—both in terms of resources and technical assistance. The business plans were then evaluated by program staff for technical, financial, and economic feasibility (See Carter, Tjernstrom, and Toledo (2016 for a full description of the program).

The multi-faceted nature of the MCC intervention creates many possible interactions between household characteristics and program impacts. Consider, for example, the stated goals of improving yields and the prices received for output. While all households can benefit by receiving better prices, the total benefit accruing to households with larger or more productive farming operations will likely be greater. On the other hand, if these households are already technically skilled they are unlikely to experience large increases in their yields by participating in the program. The reverse is likely true for households with small farm operations and/or low levels of pre-program technical skills. For these households, the dominant effect may come from increased yields. Similar stories can be told of the likely impacts of co-financed investment on households with or without access to credit and secure land-tenure status.

5.2 Data and Original Impact Estimates

The data constitute a three wave panel of 1,600 households collected over a five year period. Rollout was randomized across communities with control areas eventually being offered the treatment. In all, 62 % of households eligible for the intervention chose to participate when the program was offered to them.

The original impact analysis is presented in Carter, Tjernstrom, and Toledo (2016). The authors find positive and significant effects on income from the program activity (USD 1,345 in the specification closest to the one we estimate here). They also find significant but smaller effects on agricultural investment (USD 545), but do not find any impact on household consumption. While we use these binary treatment effects as the reference point for our analysis, the authors of the original impact analysis go on to show that both program activity income and investment effects evolve over time. Using conditional quantile regression, they also demonstrate that program benefits were heterogeneous across the population, and that households seemed to occupy stable positions in the conditional error distribution.

The current analysis is the result of a Millennium Challenge Corporation initiative encouraging researchers to undertake a "deep dive" into the data from current and past programs in order to glean lessons that can help guide future activities. As such, we begin where the authors of the original study conclude, trying to make sense of the heterogeneous pattern of impacts. Specifically, we will consider the implications of the selection and impact patterns for a program implementer trying to improve program efficiency through sharper targeting.

For comparability with the treatment effects reported in the original impact analysis, we use the same ANCOVA estimator and complier sample described in Carter, Tjernstrom and Toledo (2016). In short, the randomized roll-out of the program allows us to identify which late-treatment households chose to participate in the program when it was eventually offered to them. This group then serves as the control group for the households who were offered the program early and chose to participate. We use midline data only in our impact estimates, and the levels of the right-hand-side variables are those that would be observable to a program implementer at the time of enrollment (baseline values). Standard errors are clustered geographically. The ANCOVA model thus takes the form:

$$y_{i2} = \alpha + \beta \cdot y_{i2} + \delta T_{i2} + \epsilon_{i2} \tag{4}$$

where y_{i2} is income in the program activity at midline. y_{i1} is income in the program activity at baseline. T_{i2} takes value 1 (0) if the household was in the early (late) treatment group. And, δ gives the program impact (in this case, the impact of the treatment on the treated).

6 Searching for Positive Self-Selection

The first implication of the discussion in Sections 3 and 4 is that sharper targeting can be useful if households are not optimally self-selecting. The converse of the same idea is that if significant levels of positive self-selection are already occurring–as captured by the positive relationship between α_i and the probability of participation in Figure 2–further targeting is not necessary. In this section we look for evidence of such positive self-selection in the MCC RBD project in Nicaragua. We take the following steps to do so: First, we estimate a probit selection model, which we use to generate a household's predicted probability of participation in the program. Using the ANCOVA estimator and complier sample described above, we then estimate impacts separately for each quartile of predicted probability of participation.

Positive self-selection will be characterized by a strong selection model and increasing impacts across quartiles of predicted participation. Note that both the single and multiple impact mechanism cases described in Sections 3 and 4 will generate this pattern of impacts. Imagine, for example, that both very low and very high levels of α_i are associated with large benefits from participation and households select into the program accordingly. In the first stage probit model, both the lowest and highest quartiles of α_i will positively predict participation; households in these groups will have high predicted levels of participation and be in the upper quartile of predicted probability in the impact regressions.

To better understand the challenges facing program implementers, we start by limiting our perspective to the set of household characteristics easily observable at baseline. We then repeat the process, including household-specific levels of technical efficiency in the selection model. Technical efficiency is an example of an asset that is not easily observable to program implementers at the time of enrollment, but that likely affects the benefits of program participation. The section ends by comparing the two exercises and considering their implications on the need for sharper targeting efforts in this specific program.

6.1 Predicting Participation with Observable Household Characteristics

We begin by limiting our perspective to that of a program implementer considering a household at the time of enrollment. The asset-based regressors we include are: farm size, whether or not a household has access to formal credit, and whether or not a household has a written title to at least one plot of land. All of these characteristics could be used for targeting purposes and, in fact, do correspond to components of the program's eligibility criteria (see Table 1). In addition to being potential mechanisms for targeting, households may also self-select into the program along these asset dimensions.

We estimate a probit selection model using quintile of land holding, credit access, and land tenure security as the asset dimensions of interest. We also include controls for the program activity, the age and highest level of education completed by the household head. Tables 2 and 3 contain the results of the selection model and the associated marginal effects. Of the easily observable dimensions, only landholding significantly predicts participation. While property rights (tenure security) and credit status seem to be plausible complementary asset dimensions for the intervention, they have little effect on the probability of participation.

The selection model is quite weak, which we discuss in Section 6.3, but does provide some limited evidence of a relationship between baseline assets and participation rates. The selection model, however, is only the first part of the puzzle. To investigate positive self-selection we must relate predicted probability of participation to program impacts. To repeat our earlier conjecture, we expect that if positive self-selection is occurring, impacts should increase across quartiles of predicted probability. These impact estimates are reported in Table 4. To be comparable with our later analysis, we drop households that are missing data required to reconstruct technical efficiency measures, to which we soon turn.[3]

The model shows a pattern of impact which is compatible with some level of self-selection driven by the expectation of program impacts. Program effects are indeed increasing across quartile of predicted probability. The means of the household assets of interest are reported below the impact estimates for each quartile of predicted probability. Given that landholding was the strongest predictor of participation, it is unsurprising that total land increases strongly across quartile of predicted participation. A similar pattern is observable in the percentage of households with credit access and holding formal land title. This is true despite the fact that they were not significant predictors of participation, suggesting strong correlations to landholding.

6.2 Predicting Participation with Technical Efficiency

Earlier, in discussing the challenges of targeting, it was mentioned that many authors point out important dimensions of heterogeneity that may be unobserved by program implementers. One such dimension, which fits into the asset narrative developed here, is technical efficiency. Using the detailed data collected during the impact evaluation, we can make this asset observable via stochastic frontier estimation.

We estimate technical efficiency using STATA's *frontier* routine. The estimation process allows us to recover a farmer specific efficiency parameter ψ_i, where $0 < \psi_i < 1$. Technical efficiency is estimated separately for each activity and quartiles of technical efficiency are generated within a program activity. Some of our tables report raw technical efficiency estimates to provide a point of reference, but it is the quartile ranking that we rely on most heavily in what follows.

We re-run the Section 6.1 selection model, now including technical efficiency quartiles. The results are reported opposite the no-technical-efficiency model in Tables 2 and 3. The results are largely unchanged for the original set of characteristics. Now, however, technical efficiency emerges as a second asset dimension with noticeable effects on predicted participation. Interestingly, while the upper quartiles of technical efficiency are associated with an increased probability of participation, only the lowest quartile is associated with a lower probability of participation. Even then, the decrease is quite small. The addition of technical efficiency has not, however, significantly improved the performance of the selection model.

In order to determine if the patterns from the selection model are being driven by the expectation of program impacts, we once again estimate Equation 4 on quartiles of predicted probability. Table 5 reports the estimation results. Note that landholding is still increasing strongly across quartiles, but that the technical efficiency of households has shuffled significantly. The impact estimates are no longer what we would expect from households positively self-selecting. Accounting for technical efficiency, impacts no longer increase monotonically across quartiles of predicted participation.

[3]Using the full sample, as a program implementer would, returns almost identical results.

6.3 Implications For a Need For Targeting

While the first model of impact across quartiles of predicted probability seemed promising, overall evidence for positive self-selection is thin and sensitive to the choice of included assets and household control variables. This stems mainly from the weakness of the selection model itself. While likelihood ratio tests prefer the selection models including household assets, model selection using AIC and BIC disagree on whether or not they are preferred to a model containing only control variables. While the range of predicted probabilities is ever-so-slightly wider in the technical efficiency model, standard nested model comparison methods (AIC, BIC, and a Likelihood Ratio Test) favor the smaller model. This weak selection model and the significantly different impact results from the models with and without technical efficiency do not paint a convincing picture of positive self-selection.

Given that positive self-selection seems to be limited, there may still be room realize efficiency gains through sharper targeting. For this to be true, there must exist a small number of easily observable assets which predict the returns to participation. In the next section, we will set the selection models aside and look directly at heterogeneous impacts along a number of asset dimensions.

7 Heterogeneous "Goalpost" Effects

The second condition for sharp targeting to increase efficiency is the existence of identifiable and observable asset dimensions associated with positive program impacts. In addition to being observable, the mapping from assets to impacts must be such that program implementers can devise a targeting strategy. While a self-selection mechanism can easily handle non-linearities in the mapping from assets to program impacts, such a scenario poses more problems for active targeting. A "goalpost" pattern in which households with especially low and high levels of an asset derive significant benefits from participation is particularly difficult. Such a pattern makes the setting of simple minimum or maximum asset requirements ineffective as a targeting mechanism.

Goalpost effect patterns could arise in a number of ways. Recall the earlier discussions of multiple impact mechanisms and the multi-faceted nature of the MCC RBD program. Program impacts derived from a single asset acting through multiple mechanisms, or multiple assets each acting through a single mechanism, could easily result in a goalpost pattern of program impacts. In this section, we look for heterogeneous impacts in four different household assets. We give attention to both the magnitude and pattern of impacts across quartiles of the asset distribution.

7.1 Landholding

We begin with landholding. This is a natural starting point as it is an easily observable asset dimension and one that already plays a prominent role in the eligibility criteria of the MCC Nicaragua RBD program. For completeness, we consider both total landholding (Table 6) and land dedicated to the program activity (Table 7).

Estimating treatment impacts by quartile of total land holding reveals an interesting. The largest impacts accrue to households in the lowest and highest quartiles of total landholding, both of which exceed the approximately $1,300 cost of the intervention. The middle two quartiles, however, have impact point estimates falling below this threshold. The large estimate for the lowest quartile of landholding is especially striking given that the outcome dimension being considered—program activity income—is likely scale biased.

The impact pattern implies a significant risk to using landholding as a sharp targeting mechanism. While the selection models indicated that households with more land were more likely to participate, the Table 6 results indicate that smaller farm households may be deriving equal or greater benefit than their larger counterparts. This result is softened slightly by the results of estimating impacts across quartiles of land dedicated to the program activity. Under this specification (Table 7), the upper quartiles have larger, though noisier, impact estimates. The lowest quartile, however, still shows a positive impact estimate of USD 702. This estimate is also among the most precisely estimated and statistically significant impacts in our analysis.

7.2 Technical Efficiency

Differences in program impacts across quartiles of technical efficiency would pose a particularly vexing problem to targeting efforts. On one hand, technical efficiency seems a likely predictor of program success. On the other hand, it is not easily observable and therefore could not be used to target without significant effort and expense. If technical efficiency does predict program success, targeting along any other dimension would require knowing the relationship between that asset and the distribution of technical efficiency among the population.

Table 8 contains the results of estimating impacts across quartiles of technical efficiency. We once again see a goalpost pattern of impacts, with impacts for the first and fourth quartile of the asset estimated to be greater than the cost of the intervention. Note that the explanation for large impacts on the left-hand goal post and the right-hand goal post need not be the same. It may be that the high technical efficiency households are making use of their technical skill, while the low technical efficiency households are benefiting from another component of the program–such as a building up of the skills they lack or the loosening of a constraint on access to financing. Unlike with landholding—which we described earlier as a slow-to-change asset—technical efficiency is a plausible asset dimension for the type of endogenous formation impacts described in Section 4. In fact, it was the goalpost-shaped impacts across technical efficiency quartiles that led us to revisit the work of Karlan and Valdiva (2011) and Bjorvatn and Tungodden (2010).

Finally, note that mean technical efficiency increases only slightly across quartiles of total and program land (bottom of Tables 6 and 7. This lack of a strong relationship between the two assets means that targeting on landholding alone would not be sufficient to account for differences in technical efficiency.

7.3 Credit Status and Property Rights

The final asset dimensions we consider are credit status and property rights. Table 9 contains the results for estimating impacts separately for households according to credit access and land tenure security. Recall that neither of these dimensions was predictive of participating in the program.

This set of results is suggestive in a number of ways. Mainly, it supports the notion that the intervention might both complement existing assets (such as holding a land title) and loosen existing constraints (such as a lack of credit access). Taken together with the earlier results across quartiles of landholding and technical efficiency, the significant positive impact on households without credit access acts as a further caution to a single mechanism interpretation of the interaction between household assets and RBD program impacts.

7.4 Program Activity Differences

One factor that might be thought to drive the results presented above is differences across program crops. Indeed, participation rates and overall impacts do differ somewhat across program activities. Tables 10 - 14 contains results for estimating the impact models separately for dairy farmers and crop farmers, broken out by terciles of technical efficiency and landholding (for crop farmers only). [4] Impacts for dairy farmers seem to be concentrated in the upper tercile of technical efficiency, while the same goalpost distribution of effects persists for farmers engaged in crop activities.

8 Implications and Conclusion

While there is an inherent appeal to the idea that a brief investment in human and physical capital can result in longrun increases in household welfare, the high per-household cost and tepid results of entrepreneurial interventions have caused many to reconsider their desirability. One response to the evidence of heterogeneous returns is to seek to improve the welfare impact and efficiency of such programs through sharper targeting. This may be a valid strategy under two conditions. First, program impacts are being driven by an identifiable characteristic, preferably in a monotonic fashion. Second, households are not already self-selecting optimally. When these conditions are not satisfied, either because the characteristics driving program impacts are unobservable, impacts are non-monotonic, or households are already self-selecting optimally, an attempt to sharpen program targeting may have detrimental welfare and efficiency effects.

Both the role of household assets and the dangers of sharp targeting are born out in the example of the MCC Rural Business Development program in Nicaragua. Within the existing eligibility criteria framework, models of selection into treatment are weak and predicted probability of participation does not provide a satisfactory explanation

[4] Terciles were used instead of quartiles in the activity-specific models as the sub-sample size was decreasing the more times the data was split.

of the heterogeneous impacts observed by the original impact analysis. Such results indicate that sharper targeting may improve welfare and efficiency outcomes, provided that there exists an observable asset along which impacts are monotonically increasing. This, however, is not what we observe. On the contrary, the assets considered generate non-monotonic, goalpost-shaped, impact distributions. Such a goalpost pattern of impacts serves as a caution against efforts to sharpen targeting efforts. It also highlights the multiple pathways through which the rural business development program may have benefited participating households. For some households, the program may have complemented existing skills and assets. For other households, the program may have loosened constraints on technical skill or credit access.

9 Tables and Figures

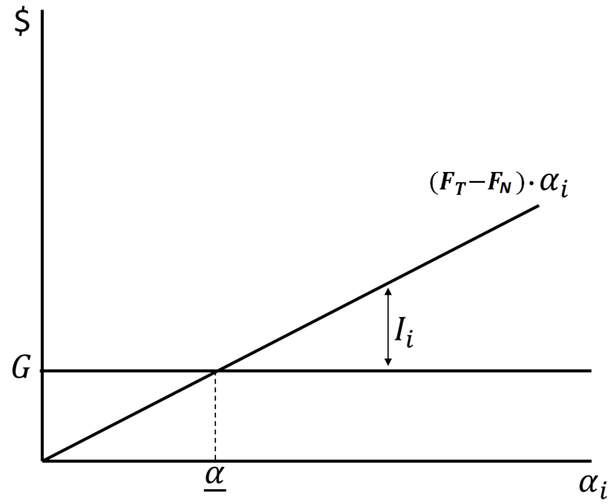

Figure 1: Simple Impacts

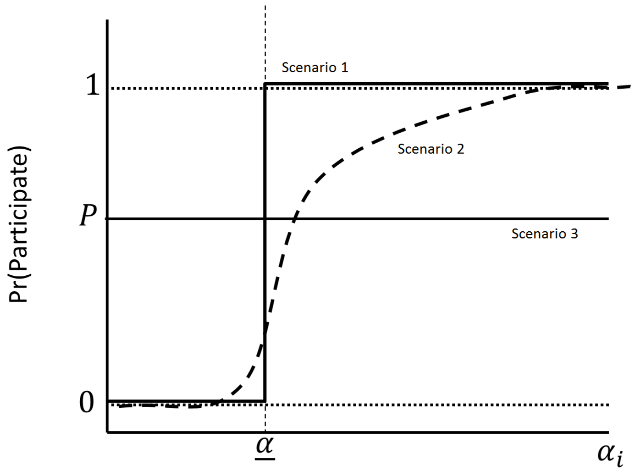

Figure 2: Self-Selection Scenarios

	Sesame	*Beans*	*Vegetables*	*Cassava*	*Livestock*
*Asset Floor**	7 hectares	3.5 hectares	1.4 hectares	3.5 hectares	10 Mature Cows
Asset Ceiling	35.2 hectares	35.2 Hectares	14.1 hectares	70.4 hectares	100 Mature Cows
Prior Experience	1.4 hectares in sesame	0.7 hectares in beans	Some Vegetable Production	1.4 hectares cassava	Developed livestock activity
Water	--	--	--	--	On-farm water source
Legal Status	Farmer has land title or is in possession of land				
Age	Farmer must be at least 20 years old				
Environment	Land located outside of national protected areas				

* Minimum farm size reduced when farm is irrigated

Table 1: MCC Nicaragua RBD Eligibility Criteria

Table 2: Probit Models

	No TE		With TE	
compliers				
Program Crop:				
Beans	-0.239*	(0.0948)	-0.233*	(0.0970)
Sesame	0.137	(0.110)	0.132	(0.112)
Casava	0.174	(0.136)	0.277	(0.154)
Age of HH Head:				
< 35	-0.105	(0.139)	-0.104	(0.139)
$35 - 44$	0.0531	(0.106)	0.0619	(0.106)
$55 - 64$	-0.0704	(0.0968)	-0.0630	(0.0971)
≥ 65	-0.123	(0.109)	-0.119	(0.109)
Edu. Level:				
None	-0.155	(0.0843)	-0.159	(0.0846)
Seconary	-0.0351	(0.121)	-0.0371	(0.122)
Past Secondary	0.00647	(0.126)	0.00771	(0.126)
Land Holding:				
Smallest 20%	-0.132	(0.110)	-0.129	(0.111)
20% – 40%	0.0761	(0.108)	0.0833	(0.109)
60% – 80%	0.164	(0.114)	0.167	(0.115)
Largest 20%	0.325**	(0.125)	0.319*	(0.125)
Title	0.0805	(0.0761)	0.0849	(0.0763)
Credit Access	0.0319	(0.0918)	0.0279	(0.0922)
TE Quintile				
1			-0.0716	(0.122)
2			0.135	(0.111)
4			0.101	(0.110)
5			0.192	(0.112)
Constant	0.291	(0.149)	0.205	(0.163)
Observations	1375		1375	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Marginal Effects of Complementary Assets on Program Participation

	No TE		With TE	
Landholding				
Smallest 20%	-0.0514	(0.0434)	-0.0500	(0.0432)
20% – 40%	0.0302	(0.0429)	0.0329	(0.0428)
60% – 80%	0.0652	(0.0454)	0.0662	(0.0452)
Largest 20%	0.129**	(0.0490)	0.126*	(0.0489)
Title	0.0316	(0.0299)	0.0332	(0.0298)
Credit Access	0.0125	(0.0361)	0.0109	(0.0360)
Technical Efficiency				
1			-0.0277	(0.0472)
2			0.0529	(0.0435)
4			0.0397	(0.0433)
5			0.0755	(0.0440)
Observations	1375		1375	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Program Impacts by Predicted Participation Probabilities, No T.E.

	Base	Q1	Q2	Q3	Q4
y_1	1.025***	0.922***	1.133***	0.779***	0.996***
	(0.0546)	(0.130)	(0.112)	(0.126)	(0.0789)
Treat	1179.3	-142.7	292.4	2165.6	3236.7
	(698.0)	(598.6)	(764.2)	(1331.2)	(1919.8)
Constant	1564.9*	1238.6*	1308.0	3154.0**	2706.8
	(629.0)	(603.9)	(760.1)	(1120.0)	(1577.1)
Variable Means					
$\hat{P}$	.636	.500	.617	.685	.742
TE	.608	.608	.603	.605	.614
TotalLand	43.7	14.1	23.8	45.9	91.4
ProgramLand	29.7	2.99	12.4	34.5	69.3
Credit	.835	.752	.827	.858	.901
Title	.639	.393	.589	.730	.845
Observations	852	214	214	211	213

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Program Impacts by Predicted Participation Probabilities: With T.E.

	Base	Q1	Q2	Q3	Q4
y_1	1.025***	1.176***	1.057***	0.878***	1.003***
	(0.0546)	(0.140)	(0.105)	(0.137)	(0.0738)
Treat	1179.3	434.1	-519.8	1709.9	3833.3
	(698.0)	(524.7)	(891.4)	(1063.7)	(1934.9)
Constant	1564.9*	725.7	2210.7***	2513.8	1996.6
	(629.0)	(489.5)	(602.9)	(1377.3)	(1598.8)
Variable Means					
$\hat{P}$	.637	.496	.618	.684	.752
TE	.608	.571	.600	.601	.658
Total Land	43.7	14.5	25.8	48.6	86.0
Program Land	29.7	3.73	13.9	38.2	63.0
Credit	.835	.732	.850	.864	.892
Title	.638	.404	.582	.746	.822
Observations	852	213	213	213	213

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Program Impacts by Total Land Quartile

	Base	Q1	Q2	Q3	Q4
y_1	1.025***	0.978***	0.836***	0.839***	0.975***
	(0.0546)	(0.215)	(0.108)	(0.109)	(0.0825)
Treat	1179.3	1688.9	988.9	456.5	1831.7
	(698.0)	(872.8)	(715.7)	(1098.2)	(1835.9)
Constant	1564.9*	195.3	1733.8**	3047.5**	4563.0*
	(629.0)	(690.0)	(635.8)	(994.9)	(1884.2)
Variable Means					
$\hat{P}$	.637	.561	.624	.646	.720
TE	.608	.581	.615	.604	.631
TotalLand	43.7	8.62	17.7	33.1	116.9
ProgramLand	29.7	3.40	7.29	16.3	93.1
Credit	.835	.770	.815	.850	.905
Title	.639	.549	.631	.602	.773
Observations	852	213	222	206	211

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Program Impacts by Program Land Quartile

	Base	Q1	Q2	Q3	Q4
y_1	1.025***	0.713***	0.646**	0.605***	0.944***
	(0.0546)	(0.159)	(0.235)	(0.103)	(0.0866)
Treat	1179.3	664.7**	1946.3*	1387.2	2257.6
	(698.0)	(207.0)	(828.7)	(1166.0)	(1691.0)
Constant	1564.9*	877.2***	1460.4	5489.5***	5540.0**
	(629.0)	(210.2)	(857.0)	(1036.3)	(1925.3)
Variable Means					
$\hat{P}$	.637	.555	.642	.653	.715
TE	.606	.581	.588	.620	.642
Total Land	43.7	17.1	20.5	27.9	110.2
Program Land	29.7	1.48	4.00	16.7	97.4
Credit	.835	.787	.801	.837	.915
Title	.638	.469	.663	.660	.793
Observations	852	249	181	209	213

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Program Impacts by Technical Efficiency Quartile

	Base	Q1	Q2	Q3	Q4
y_1	1.025***	0.979***	0.974***	0.914***	1.098***
	(0.0546)	(0.224)	(0.167)	(0.110)	(0.0772)
Treat	1179.3	1475.3	-846.0	934.4	2497.0
	(698.0)	(849.5)	(1127.7)	(1680.2)	(1259.1)
Constant	1564.9*	1075.4	4308.6***	3326.8*	-997.4
	(629.0)	(610.0)	(1176.2)	(1405.8)	(1087.0)
Variable Means					
$\hat{P}$	.637	.604	.636	.652	.659
TE	.608	.370	.577	.686	.798
Total Land	43.7	29.6	42.5	62.1	40.7
Program Land	29.7	15.6	30.4	47.0	25.9
Credit	.835	.803	.840	.845	.850
Title	.638	.662	.606	.662	.624
Observations	852	213	213	213	213

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Program Impacts by Credit Access and Property Rights

	No Credit	Credit	No Title	Title
y_1	1.030***	1.021***	0.961***	1.025***
	(0.0750)	(0.0612)	(0.0777)	(0.0663)
Treat	1822.7	1096.4	741.1	1470.4
	(911.9)	(795.8)	(722.6)	(894.1)
Constant	692.9	1723.3*	1385.6**	1840.2*
	(658.5)	(724.3)	(472.9)	(873.6)
Variable Means				
$\hat{P}$	.600	.645	.592	.663
TE	.595	.610	.611	.606
TotalLand	35.8	45.3	30.6	51.1
ProgramLand	16.6	32.3	17.3	36.7
Credit	0	1	.808	.849
Title	.582	.650	0	1
Observations	141	711	308	544

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Impacts by Program Crop

	Beans	Sesame	Casava	Livestock
y_1	0.379	0.587***	0.277	0.958***
	(0.205)	(0.0854)	(0.203)	(0.0753)
Treat	670.2**	1768.2*	3828.1	1371.2
	(226.9)	(717.8)	(2320.9)	(1091.1)
Constant	1242.3***	1416.1*	2409.2*	4452.6**
	(318.3)	(561.8)	(914.4)	(1297.0)
Variable Means				
$\hat{P}$	.513	.686	.687	.680
TE	.601	.646	.388	.645
Total Land	18.4	23.6	21.8	72.1
Program Land	1.85	5.46	5.07	62.0
Credit	.758	.829	.849	.879
Title	.396	.639	.802	.745
Observations	227	158	86	381

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Crop Program Impacts by Technical Efficiency Tercile

	Base	T1	T2	T3
y_1	0.486**	0.619*	0.647***	0.226*
	(0.146)	(0.245)	(0.171)	(0.0997)
Treat	1225.8	1766.6*	701.3	1312.6
	(616.3)	(837.2)	(685.1)	(712.0)
Constant	1499.6**	868.8	1478.5	2269.5***
	(458.8)	(523.3)	(732.0)	(426.4)
Variable Means				
$\hat{P}$	.603	.577	.612	.620
TE	.577	.370	.578	.784
Total Land	20.8	21.7	19.3	21.3
Program Land	3.65	3.86	3.57	3.52
Credit	.798	.771	.828	.796
Title	.552	.586	.554	.516
Observations	471	157	157	157

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Livestock Program Impacts by Technical Efficiency Tercile

	Base	T1	T2	T3
y_1	0.958***	1.233***	0.859***	1.033***
	(0.0753)	(0.132)	(0.139)	(0.103)
Treat	1371.2	-743.4	-757.0	4836.9*
	(1091.1)	(1385.5)	(2099.1)	(2286.1)
Constant	4452.6**	4286.0**	7593.1***	-446.0
	(1297.0)	(1298.8)	(1906.5)	(2389.2)
Variable Means				
$\hat{P}$	.680	.678	.664	.700
TE	.645	.523	.655	.755
TotalLand	72.1	63.7	76.3	76.2
ProgramLand	62.0	57.6	68.8	59.6
Credit	.879	.898	.882	.858
Title	.745	.724	.724	.787
Observations	381	127	127	127

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Crop Program Impacts by Tercile of Total Land

	Base	T1	T2	T3
y_1	0.486**	0.676	0.433***	0.405**
	(0.146)	(0.521)	(0.0899)	(0.136)
Treat	1225.8	1906.7	843.2	821.5
	(616.3)	(936.3)	(485.0)	(963.1)
Constant	1499.6**	342.0	1676.4***	2520.0***
	(458.8)	(1094.0)	(330.7)	(694.7)
Variable Means				
$\hat{P}$	.603	.543	.627	.642
TE	.577	.555	.598	.581
Total Land	20.8	8.2	16.1	38.4
Program Land	3.65	2.57	3.52	4.90
Credit	.798	.738	.827	.834
Title	.552	.543	.560	.554
Observations	471	164	150	157

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 14: Crop Program Impacts by Tercile of Total Land

	Base	T1	T2	T3
y_1	0.486**	0.572***	0.392	0.381**
	(0.146)	(0.0956)	(0.519)	(0.134)
Treat	1225.8	707.0**	1073.9	2659.6
	(616.3)	(212.8)	(1203.6)	(1402.4)
Constant	1499.6**	977.8***	2749.2	2251.6**
	(458.8)	(156.2)	(1502.5)	(743.1)
Variable Means				
$\hat{P}$	.603	.552	.618	.676
TE	.577	.580	.591	.567
Total Land	20.8	17.0	20.7	26.8
Program Land	3.65	1.48	2.91	7.47
Credit	.798	.787	.743	.843
Title	.552	.459	.662	.647
Observations	471	244	74	153

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

10 Bibliography

References

Bjorvatn, K and B Tungodden (2010). "Teaching business in Tanzania: Evaluating participation and performance". *Journal of the European Economic.*

Blattman, Christopher and Laura Ralston (2015). "Generating employment and fragile states: Evidence from labor market and entrepreneurship programs".

Carter, Michael, Emilia Tjernström, and Patricia Toledo (2016). *Heterogeneous Impact Dynamics of a Rural Business Development Program in Nicaragua.* Tech. rep. Cambridge, MA: National Bureau of Economic Research.

Fairlie, RW, D Karlan, and J Zinman (2015). "Behind the GATE experiment: Evidence on effects of and rationales for subsidized entrepreneurship training". *American Economic Journal:*

Karlan, D and M Valdivia (2011). "Teaching entrepreneurship: Impact of business training on microfinance clients and institutions". *Review of Economics and statistics.*

McKenzie, David and Christopher Woodruff (2014). "What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World?" *The World Bank Research Observer* 29.1, p. 48.

Mel, S De, D McKenzie, and C Woodruff (2014). "Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka". *Journal of Development Economics.*

Millenium Challenge Corporation (2009). *Rural Business Development Program Documents.*

World Bank Group (2017). URL: `http://www.worldbank.org/`.