Investing in Small-Farm Productivity: Impact Dynamics and Heterogeneity in Nicaragua

By Michael R. Carter, Patricia Toledo and Emilia Tjernström*

We report the results of a multi-year impact evaluation of a program designed to boost the income of the small-farm sector in Nicaragua. Using continuous treatment estimation techniques, we examine the evolution of program impacts over time and find that incomes in the activities targeted by the program as well as farm capital rise significantly over time. Because of the temporal pattern of impacts, shorter-term binary treatment estimators do not fully capture the impacts of the program. Additionally, panel quantile methods reveal striking heterogeneity of impacts across the sample.

Program area: Methods and Data, Sub area: Evaluation Methods
JEL: I32, O12, O13, Q12, Q16
Keywords: Agricultural productivity, Impact evaluation

Constraints such as poor infrastructure, and limited access to credit and improved technologies keep productivity low in the agricultural sector of many developing countries. The program in Nicaragua focused on potentially profitable value chains, providing participating farm households with marketing support, co-investment opportunities, as well as technical assistance and training. The value of information in rural farm settings constitutes a common thread in the literature on technology adoption in agriculture (Griliches (1957); Foster and Rosenzweig

* Carter: University of California, Davis, 2103 Social Sciences and Humanities, One Shields Avenue, Davis, CA 95616, (email: mrcarter@primal.ucdavis.edu); Toledo: Ohio University, Bentley Annex 345, Athens, Ohio 45701, (email: toledot@ohio.edu); Tjernström: University of California, Davis, 2156 Social Sciences and Humanities, One Shields Avenue, Davis, CA 95616, (email: emiliat@primal.ucdavis.edu).
Acknowledgements: We thank Anne Rothbaum, Lola Hermosillo, Jack Molyneaux, Juan Sebastian Chamorro, Carmen Salgado, Claudia Panagua, Sonia Agurto, the staff at FIDEG, and Conner Mullally. We gratefully acknowledge funding from the Millennium Challenge Corporation, as well as financial support from the US Agency for International Development Cooperative Agreement No. EDH-A-00-06-0003-00 through the BASIS Assets and Market Access CRSP.

(1995); Conley and Udry (2010)), but the effectiveness of agricultural extension interventions – programs that, among other things, aim to provide information to farmers – is mixed. Reviews of agricultural extension studies (Evenson, Gardner and Rausser (2001); Anderson and Feder (2003)) point to methodological and data quality issues – in particular the common lack of attention to issues of causality – in previous studies as major barriers to effective learning in this area.

One of the first studies to approach this topic using an experimental design Ashraf, Giné and Karlan (2009) evaluates the impact of a bundle of export-focused extension services.[1] After one year, Ashraf, Giné and Karlan (2009) find some positive impacts on incomes, but the export transactions collapse almost entirely after the evaluation ended. While this disappointing outcome probably isn't representative of extension programs more generally, it clearly demonstrates the importance of paying attention to the timing and evolution of impacts. As King and Behrman (2009) point out, programs with important learning and adoption components are unlikely to attain steady-state effectiveness soon after an intervention begins. In order to better examine the impact dynamics, we employ continuous treatment estimators in our evaluation of a rural business development (RBD) program in Nicaragua. This allows us to trace the evolution of impacts over a relatively long term, and to compare these estimates to the shorter-term local average treatment effects.

Our evaluation strategy is built around a randomized program roll-out, which involved randomly splitting eligible households into early and late treatment groups. The early households began treatment shortly after a baseline survey, while the late treatment group began treatment after a second (mid-line) survey round. The standard binary treatment effects are estimated using mid-line data, since at that time the late treatment group functions on average as a valid control group for the early treatment group. The choice of a randomized roll-out

[1] See also Feder, Slade and Lau (1987) for an earlier study of extension service intensification using a quasi-experimental research design, which uncovers positive but diminishing effects of extension services.

strategy was due to program capacity constraints – not all eligible households could be treated at once – and this also meant that treatment was staggered, and the temporal sequence in which households received the treatment was *de facto* randomized. We therefore collected a third (end-line) survey round, and exploited the panel structure of the data to estimate the continuous treatment effect, using the number of months since a household began treatment to map out the corresponding duration response functions.

Using fixed effects estimators, we show that the temporal pattern of impact indeed evolves in important ways over time. Income in the activities targeted by the program rises significantly over time, on average reaching a 28% increase after two years in the program, and these increases appear largely to be sustained over time. The program also brought about significant increases in farm capital. However, on average we observe no significant impacts on household consumption patterns, possibly because households are choosing instead to reinvest the income increases into their farm. We contrast the continuous analysis with the results of a standard binary average treatment effect estimator, and a double-sided complier estimator. The binary treatment effect estimators do detect significant impacts on farm income, but unlike the continuous estimators fail to pick up any impacts on investment.

Given the nature of the intervention (farmers had to make substantial effort, be able to learn from the training provided, and contribute part of the cost of any investment deemed necessary by extension specialists), we might suspect that it would not work equally well for all participants[2]. It seems important, therefore, to try to understand for how many and for what kind of households the program actually works – especially if we want to inform policy makers who must choose between allocating scarce resources to, say, cash transfer programs or to more complex value chain programs like the one evaluated here.

[2]For example, it might work better for those privileged by larger endowments of business acumen or other skills that are complementary to the treatment.

To explore this issue of impact heterogeneity, we employ correlated random effects panel quantile regression techniques developed by Abrevaya and Dahl (2008) to determine the extent to which estimated average impacts are indicative of the full range of impacts experienced by program participants. The analysis reveals quite striking heterogeneity in impacts. Beneficiaries in the 75^{th} conditional quantile of incomes enjoy much larger impacts than those in the lower quantiles, and those in the top conditional quantiles of investment invest many times more than those in the lower quantiles.

I. Background and Data

Agriculture has played an important role throughout Nicaragua's history, but multiple constraints have prevented agriculture from reaching its productive potential – examples include a lack of basic infrastructure, low education levels, and low access to credit and technology. The Western Region of Nicaragua, which includes the departments of Leon and Chinandega, was identified by the National Development Plan (NDP) as having particularly high potential for agricultural growth. While high-potential, the area is also quite poor: the World Bank (2008) identified that more than 50 percent of households in the Western Region live in poverty.

In July 2005, the Millennium Challenge Corporation (MCC) signed a five-year, $175-million compact with the Government of Nicaragua to develop a set of projects in the Western Region, with the objective of relaxing some of the aforementioned constraints. The compact had three components: a transportation project, a property regularization project, and the one we focus on here: a rural business development (RBD) project.[3] This latter component aimed to raise incomes for farms and rural businesses by helping farmers develop and implement a business plan built around a high-potential activity, as explained in more detail

[3] The MCC Board terminated a portion of the compact in June of 2009, reducing compact funding from $175 million to $113.5 million. While this action cut off the property regularization part of the program, the RBD Program was not affected by this partial project termination.

below.

A. Program Description

The implementing agency (the Millennium Challenge Account, or MCA) identified the productive activities most suitable for inclusion in the program: beans, cassava, livestock, sesame, and vegetables. In order to be eligible, farmers had to own a small- or medium sized farm[4], have some experience with one of these crops, be willing to develop a business plan together with extension agents, and commit to contributing 70% of the investments deemed necessary by the extension agent. A set of specific eligibility rules was developed for each productive activity, and the precise rules are shown in Appendix Table A1.[5]

Once a farmer enrolled in the program and had his business plan approved, the RBD program worked with him for 24 months. The exact benefits varied across the productive activities, but in general all farmers received technical and financial training and supplies based on their individual business plan. If investment was part of the business plan, the RBD Program could contribute up to 30% of the resources needed.

B. Evaluation Methodology

The challenge of this and all impact evaluation efforts is to identify a control group that is identical to the treatment group in every way except that they have not benefited from the intervention under evaluation. We exploit the fact that capacity constraints meant that not all eligible farmers could be brought into the project immediately. The evaluation team worked with the RBD implementation organization to identify all geographic clusters that would eventually be offered

[4] Small- or medium farm here is per the NDP definition, which included more than just farm size: small and medium farms are thought to be those that face various constraints, such as lack of access to financial services, low or variable product quality, technological weaknesses, and lack of access to markets.

[5] The impact of these eligibility criteria on the characteristics of the eligible population is described in detail in a companion paper, Carter, Patricia Toledo and Emilia Tjernström (2012), but in summary they effectively targeted direct benefits of the program toward the upper half of the rural income distribution of Leon and Chinandega.

RBD services. The evaluation team then selected a subset of these clusters for random assignment to either *early* or *late* treatment status.

Once the random assignment of early and late clusters was made, 1,600 households were sampled from the roster of all eligible producers in these clusters, split equally between early and late areas.[6] These 1,600 households completed a baseline survey in late 2007, just as the RBD program was rolling out in the early treatment clusters. The mid-line survey was applied approximately 18 months later, right before the late treatment group was initiated. As can be seen in Figure 1, the timing of the surveys meant that the late treatment group function as a temporary control group at the time of the mid-line survey, since clusters were randomly allocated to early and late treatment and those allocated to late treatment had not yet been treated. Both early and late treatment clusters were then surveyed again near the end of the program in 2011. The roll-out strategy also randomized the duration of time in the program, a feature that will prove vital in the continuous treatment estimates presented below.

64% (57%) of the eligible households in the early (late) treatment clusters chose to participate in the RBD project. While we could not foresee who would reveal themselves to be compliers at the baseline, the late-treatment households made their participation decision around the time of the second-round survey. We can therefore estimate the impacts both on eligible households and on participating or *complier* households. The evaluation will primarily focus on the complier households as we are interested in the impact of the program on the types of self-selecting individuals who adopt it, and the double-sided estimator (described in Section II.A) boosts power compared to standard local average treatment effect estimators.

[6] More detail on the design and implementation of the experimental design can be found in the companion paper to this one, Carter, Patricia Toledo and Emilia Tjernström (2012)

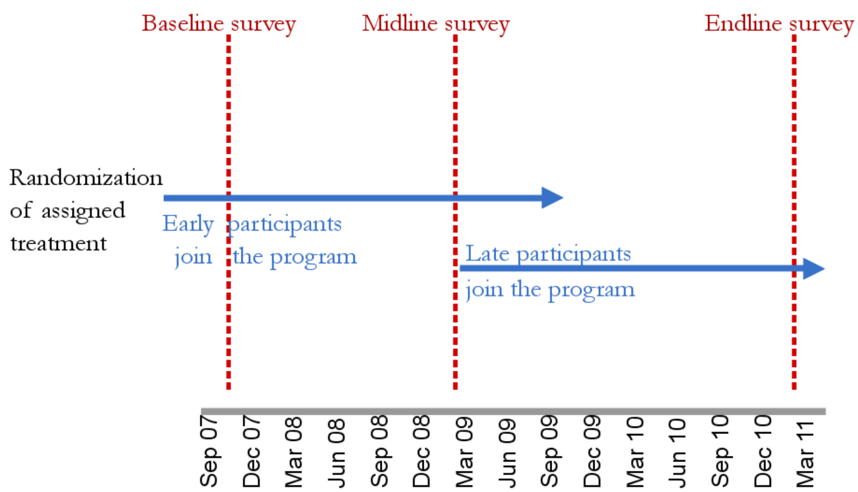


Figure 1. Timeline of Received Treatment and Timing of Surveys

C. Data

As the previous section explained, we have a three-wave panel of 1,600 households with less than 2 percent attrition by the time of the third wave. To see how well the randomization worked, Tables B1 and B2 present two sets of balance checks: one that compares eligible farmers in the early and late treatment groups, and one that compares compliers – i.e. the farm households that chose to participate and comply with the treatment – between the early and late treatment group. The randomization did not block on any of these demographic variables, but we would expect the groups to look similar on average because of the randomization.

Most of the variable averages suggest that the groups represent the same population, i.e. they are not statistically significantly different. Comparing all types of eligible farmers, the only variable where the early and late treatment groups differ statistically is on value of animals owned, and the actual difference is so small that it is unlikely to be economically significant. Some variables are significantly different once we compare early and late groups *within* activities. This indicates that even though the randomization took the productive activity of farmers into account, the number of geographical units samples was insufficient to generate on-average-equal treatment and control groups when we break the sample down by productive activity. Given that we lump all productive activities together in the evaluation, this should not be a major cause for concern (and even the statistically different variables are economically insignificant).

In a typical impact evaluation study with imperfect compliance, it is impossible to know who in the control group is a complier type. Given that our control households were also treated, we can cleanly identify compliers in both the treatment (early) and the control group (late). Therefore, we can check for balance also between early and late *compliers.* A few more variables show up as significant (value of animals, age of farmer, and whether or not the farmer has formal title to the land). We would only worry about these differences if we believed them to upset the parallel-trends assumption of the difference-in-difference estimator that we will employ (as explained in the next section). Depending on the direction of the impact of these differing variables on farmers' growth rates, the impacts of the program *could* either be over- or under-estimated. As before, we would emphasize that while the differences are statistically significant, they are unlikely to be economically significant.

II. Econometric Methodology

Our three outcome variables of interest represent both direct and indirect channels of impact. The small-farm intervention was in the first instance designed to enhance the access of small farmers to improved technologies and to markets, so we begin by examining program impacts on income in the target crops. While this measure is almost certainly an upper bound on the direct impacts of the program, we provide some evidence that the impacts are not primarily driven by household inputs being substituted into the target crop. If we observe increases in agricultural income, households would then face a key allocative choice: allocate income increases immediately to consumption, or reinvest income increases into the farm operation, postponing increased consumption until a later date. We therefore examine investment and household consumption in turn.

We evaluate the impacts of the program using two main econometric approaches. First, we adapt the local average treatment effect (LATE) estimator to our setting by using a two-sided complier estimator that allows us to gain power compared to standard approaches. Second, we employ a fixed effect continuous treatment estimator to examine the evolution of impacts over time.

To motivate our focus on continuous treatment effects, note that the workhorse impact evaluation estimators assume that program participation is a binary state – either a household receives the treatment or it does not. While this approach deals well with treatment heterogeneity across treated units (hence the derivation of local average treatment effects), it is less well-equipped to deal with the heterogeneity implied by a temporal impact curve driven by the sorts of considerations just discussed. This observation is especially important when the time span of the different impact phases is not known (*e.g.*, how long does a potential Ashenfelter dip last)?

To better frame these issues, consider the hypothetical impact relationships for the small-farm intervention illustrated in Figure 2. The solid, blue functions illustrate what we might expect to see for the early treatment group, while the dashed red functions illustrate the same for the late treatment group. The horizontal axis shows approximately the time at which the different survey rounds were undertaken (see the discussion in Section I.B). If the program had reached its full long-term impact on the early (late) treatment group by the time of the second (third) round survey, then conventional binary estimators would work well. In this case we would expect the data to trace out impact patterns similar to the step functions. On the other hand, if the impact of the program evolves more slowly over time (for example, with an initial 'Ashenfelter Dip' followed by a slow rise toward a long-run or asymptotic treatment effect), then our data

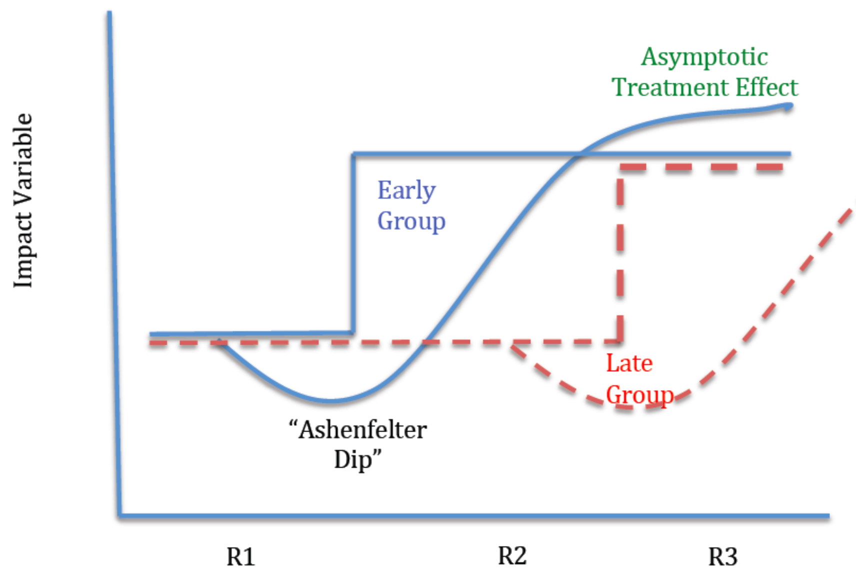


FIGURE 2. HYPOTHETICAL IMPACT PATTERNS

would be generated by a a non-linear impact or duration response function in which impact depends on the duration of time in the program. The blue and red curves illustrate what those impacts may look like for the early and late treatment groups, respectively. In the hypothetical case illustrated in Figure 2, impacts measured at mid-line using a conventional binary treatment (step function) estimator would reveal muted effects that would not accurately represent the long-run program impacts. In the remainder of this section, we will propose an econometric methodology that is general enough to encompass both cases.

A. Binary Treatment Model

Thanks to the timing of the survey rounds, our dataset allows us to identify which sampled farmers from the early treatment group were indeed enrolled before 2009, as well as which farmers from the control group (assigned to late treatment) were enrolled in or after 2009. In other words, we are able to identify participants and compliers in both the treatment and control groups and can estimate the effect of the program removing those farmers who did not participate and those who enrolled, but at a time different from the one determined by the random assignment. In focusing on this group, we restrict our attention to the subpopulation of farmers who would join a small-farm income generating program. Since the vast majority of program costs are spent on participating farmers, impacts on this subpopulation are the most relevant to policymakers.

To formalize the two-sided complier (2SC) estimator, we define three indicator variables. The variable B_i indicates assigned treatment, and equals 1 for eligible farmers who were assigned to the early treatment group, and 0 for those assigned to the late treatment group. The variable D_i indicates whether or not a farmer participates in the program when it was offered to them, so that $D_i = 1$ if the farmer participates and $D_i = 0$ if not. Finally, T_i identifies early and late compliers:

$$\begin{cases} T_i = 1 & \text{if } B_i * D_i = 1 \\ T_i = 0 & \text{if } B_i = 0 \text{ and } D_i = 1 \end{cases}$$

In other words, complier farmers in the early treatment group are those who actually enrolled in the small-farm intervention when it was offered to them, and complier farmers in the late treatment group are those who enrolled in the program when it was eventually offered to them (after the second round surveys). We can therefore use just this complier sample and the first two rounds of data to compute the effect of the program on the sub-population of compliers as the following difference-in-difference estimator:

$$\delta_{2SC} = E\left[Y_{i,2}^{E} - Y_{i,1}^{E} | T_i = 1\right] - E\left[Y_{i,2}^{L} - Y_{i,1}^{L} | T_i = 0\right], \tag{1}$$

where the superscripts indicate treatment assignment (Early & Late), and subscripts denote survey round (1,2). Since we have a third round of survey data, we can generalize the above complier estimator as follows:

$$(2) \qquad E[Y_{it}] = \alpha_0 + \lambda_2 t_2 + \lambda_3 t_3 + \delta_2(t_2 T_i) + \delta_3(t_3 T_i) + [\gamma T_i + \beta' X_i],$$

where X_i represents a vector of baseline characteristics (namely, crop indicator variables, farmer age, and farmer education), and t_2 and t_3 are binary time period indicators survey rounds 2 and 3, respectively. The parameters γ and β control for any baseline differences between early and late treatment groups. Since all the variables in square brackets are time-invariant, we can replace them with a single, household level fixed effect and rewrite the equation 2 as:

$$(3) \qquad E[Y_{it}] = \alpha_0 + \lambda_2 t_2 + \lambda_3 t_3 + \delta_2(t_2 T_i) + \delta_3(t_3 T_i) + [\alpha_i],$$

where α_i is the fixed effect.

From Equation (3), the parameter δ_2 will estimate δ_{2SC} the complier difference-in-difference estimator. If a binary step function is a good approximation of program impacts, then we would expect δ_3 (which measures the difference between the treatment and control groups in the third survey round, after the control group has been treated) to be zero, since both groups are full participants in the program by then.

An alternative and more straightforward way to write the estimating equation involves defining a fourth indicator variable, Z_{it}, which equals 1 if farm i has been treated at time t, and 0 otherwise. Using this new variable and differencing out the baseline, we can sweep away the fixed effect term and estimate the model as:

$$(4) \qquad E[Y_{it} - Y_{i1}] = \lambda_2 t_2 + \lambda_3 t_3 + \delta Z_{it}, \; t = 2, 3.$$

As before, δ estimates the complier group difference-in-difference treatment effect, and will in fact be identified entirely off of second-round variation. It's easy to verify that the parameters δ_2 from Eq. (3) and δ from Eq. (4) are numerically equivalent. In the analysis to follow, we will primarily rely on this specification to estimate binary treatment effects using the double-sided complier sample.[7] We also report standard intention-to-treat (ITT) effects and their associated (full-sample) local average treatment effects (LATE) in Appendix C.

B. *Continuous Treatment Model*

As discussed in the beginning of this section, there are a number of possible reasons why the impact of the RBD program may have evolved over time. In addition to a possible initial dip in living standards when households first joined the program and focused their resources on building up the targeted activity, there are at least three other reasons why the impact of the small-farm intervention may have changed over time. First, program beneficiaries may have experienced a learning effect with their technical and entrepreneurial efficiency improving over time. Second, the asset program may have created a crowding-in effect if the program incentivized beneficiaries to further invest in their farms. As Keswell and Carter (2011) discuss, these second round multiplier effects are what distinguish business development and asset transfer programs from cash transfer programs and other common anti-poverty policy instruments. Third, and less positively, if program impacts are short-lived (e.g., if treated farmers drop the improved practices as soon as the 24-month period of intense RBD involvement with their groups end), then impacts may dissipate over time.

One goal of this study is to estimate the impact dynamics and duration response function, and thus recover both the long-run impacts of the intervention and time path of said impacts. Both are of particular relevance from a policy perspective.

[7]Note that the validity of this estimator relies on the idea that the decision to enroll in the early and late treatment groups was structurally the same, so that we are in fact comparing like-with-like in using this estimator. This assumption is in addition to the usual exclusion assumption, i.e. that farmers who do not enroll in the program experience no effect from the treatment or the randomization. See Section I.C above for more discussion.

Indeed, it is the prospect that a skill-building program like the RBD program will facilitate and crowd-in additional asset building that makes them especially interesting as an anti-poverty program.

We begin by generalizing the double complier binary response function from Eq. (3) to the continuous treatment case:[8]

$$y_{it}(d_{it}) = \lambda_2^d t_2 + \lambda_3^d t_3 + f(d_{it}) + \alpha_i + \varepsilon_{it}, \tag{5}$$

where d_{it} is the number of months that farm i had been in the RBD program at survey time t, and $f(d_{it})$ is a flexible function that can capture the sorts of non-linear impacts illustrated in Figure ((2)) above. Empirically, we will measure duration as the number of months between when the household's farm cluster initiated activity and the date of the survey. In our data set, these durations run from 0 to 50 months.[9]

In order to gauge the shape of the function $f(d_{it})$, we employ semi-parametric analysis (see Appendix (D)) and then choose a polynomial (parametric) functional form consistent with the semi-parametric results. As in the binary treatment case, the farm-specific fixed effect term controls for all observed and unobserved time-invariant characteristics, including farming skill, soil quality, farmer education, *etc.* Importantly, this fixed effect analogue estimator controls for any systematic or spurious correlation between observables and duration of treatment.[10]

While there are several computationally equivalent ways to consistently estimate a fixed effect model like equation 5, in anticipation of later quantile re-

[8] We could alternatively follow Hirano and Imbens (2004) generalization of propensity score matching to the continuous treatment case. The Hirano and Imbens estimator only exploits observations with strictly positive amounts of treatment. In our case, this would imply dropping the baseline data for all RBD participants as well as the mid-line data for the late treatment group. For development applications that employ this estimator, see Keswell and Carter (2011) and Aguero, Carter and Woolard (2010)

[9] In a few cases, RBD activities began a few months prior to the baseline survey. For these cases, we have considered households in these clusters as treated at baseline, but their values for d_{it} can exceed the number of months between the first and third rounds of data collection.

[10] There have been some suggestions from field staff that the implementer wanted to treat the households that they considered the most promising earlier on. While this would only have affected a small fraction of the sample (field staff tried their best to limit this behavior), this methodology should deal with any differential treatment length that is correlated with observables.

gression analysis (where such models are less easily estimated), we will build on the correlated effects model of Mundlak (1978) and Chamberlain(1982,1984) and write the individual fixed effects as a linear projection onto the observables plus a disturbance:

$$\alpha_i = \psi + X_{i1}'\lambda_1 + X_{i2}'\lambda_2 + X_{i3}'\lambda_3 + \upsilon_i,$$

where X_{it} denotes a vector of observables, which includes the time dummies and the duration variables. In our case, we have little reason to believe that the way in which the time-varying observables affect the individual effects differ between survey rounds, so we use the average of the time-varying covariates and write the fixed effect as

$$\alpha_i = \psi + \bar{X}_i'\bar{\lambda} + \upsilon_i.$$

Substituting this expression into (5) gives:

$$y_{it}(d_{it}) = \lambda_2^d t_2 + \lambda_3^d t_3 + f(d_{it}) + \psi + \bar{X}_i'\bar{\lambda} + [\upsilon_i + \varepsilon_{it}] \tag{6}$$

OLS estimation of (6) allows us to recover the fixed effect estimators of the impact response function parameters of interest. As can be seen in Appendix (D), the function can be reasonably approximated by a cubic functional form, so that the actual estimating equation is:

$$y_{it}(d_{it}) = \lambda_2^d t_2 + \lambda_3^d t_3 + \zeta_1 d_{it} + \zeta_2 d_{it}^2 + \zeta_3 d_{it}^3 + +\psi + \bar{X}_i'\bar{\lambda} + [\upsilon_i + \varepsilon_{it}], \tag{7}$$

III. Results

As a prelude to the continuous analysis in this section, Figure 3 shows the histogram describing the distribution of months in the RBD program for the complier sample across all three survey rounds[11]. There is a large cluster of observations with zero treatment (comprised of the early and late treated households at baseline, plus the late treated households at the mid-line), which are not shown in the figure. Despite some bunching, the data show reasonable dispersion with households observed with as little as 1 month in the program to as much as 50 months in the program. The largest clusters of observations are around 6, 18 and 40 months of program exposure. Late treatment households comprise the first group, both early and late treatment households are found in the middle exposure group (the former at mid-line, the latter at end-line), while the latter group is comprised exclusively by early treatment households. It is this variation in length of time in the program that will be exploited in the continuous treatment estimators employed in this section.

Also, before examining the outlined regression models, we examine descriptive statistics on farm technology use and production. Table 1 displays various indicators by producer group[12]. The variable *"Manzanas planted"* is the total area that a household planted in the RBD target crop in the survey year, and can be thought of as a measure of the intensity of production in the target crop given that farmers have a (mostly) fixed amount of land at their disposal.[13] The variable *"Improved seed"* is the percentage of households that used an improved seed variety for the target crop during at least one season, and measures one aspect of farmers' utilization of improved technology. For dairy farmers, the measure is instead whether farmers applied any processing to their products before bringing them to market (*"Processing"*).

The variable *Income* represents the total value of production in the target crop,

[11]The histogram excludes 0 months for reasons of scale - including zero months of program exposure dwarfs the other bins.

[12]The producer group Hortalizas has been excluded from this analysis as the mode of production doesn't compare easily to the other producer groups in this particular framework.

[13]This variable can be thought to combine two measures: the extensive margin of cultivation and the number of times the plots are cultivated in a given time period. The distinction between the intensive and extensive margin here is not the main focus – the time period in question is one year, and most of the crops have one main planting season. If farmers have a mostly fixed amount of land at their disposal, allocating more land to the target crop then might suggest that the farmer has increased their valuation of production in the target crop .

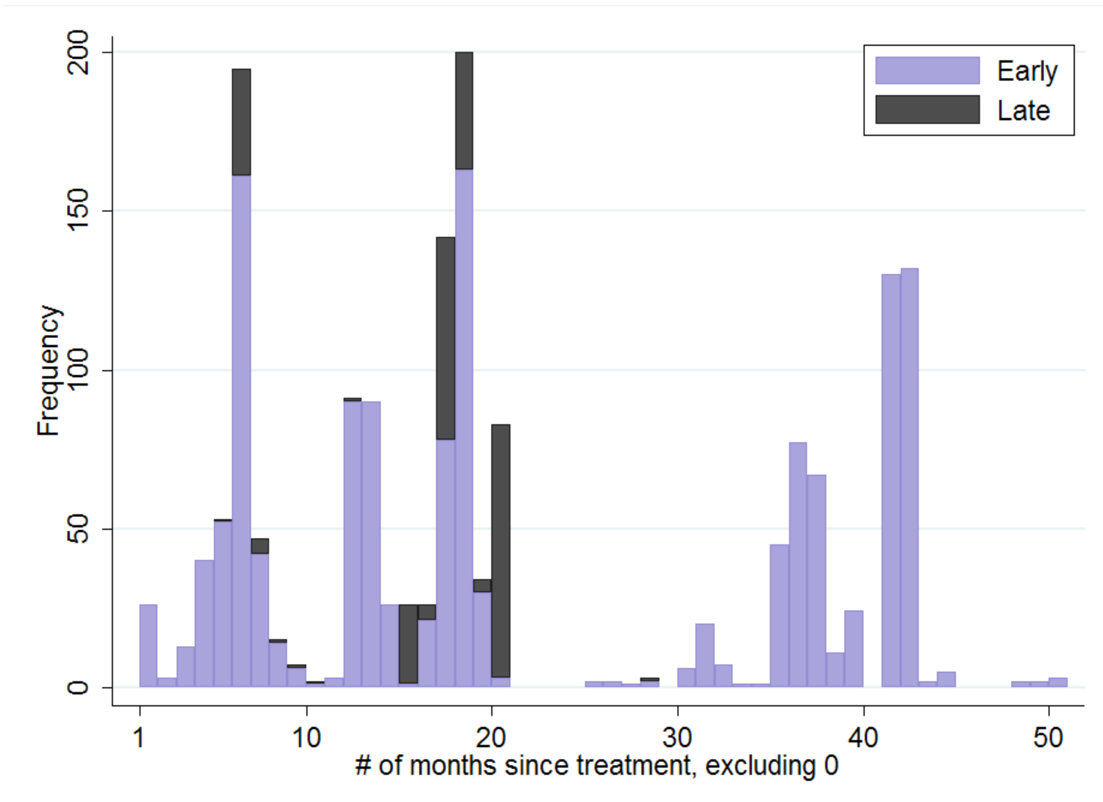


Figure 3. Distribution of the Duration of RBD Treatment (Dual Complier Sample)

Table 1—Prices, Technology and Incomes by Farmer Activity

		Baseline		Mid-line		End-line	
		Early	**Late**	**Early**	**Late**	**Early**	**Late**
Beans	Value of production	11416.13	11776.72	20652.74*	15679.03*	11461.06	10109.25
	Used improved seed (%)	.109	.087	.562***	.203***	.284	.25
	Manzanas planted (#)	3.35	3.01	4.60**	3.32**	3.53*	2.64*
	Price	433.6	422.1	823.2*	772.9*	1009.8	996.8
	N	183	69	185	69	176	65
Sesame	Value of production	28888.38	28523.96	40447.18	29561.45	48462.97	36169.35
	Used improved seed (%)	.456**	.696**	.621	.685	.439***	.807***
	Manzanas planted (#)	5.32	5.8	5.73**	3.98**	5.27	4.37
	Price	617.9***	515.3***	1276.5**	1131.7**	1409.3	1318.3
	N	86	108	86	107	66	93
Cassava	Value of production	45807	37585.3	74519.8	42177.2	32225.2	66599.97
	Used improved seed (%)	.064	.056	.171*	.0233*	.171	.077
	Manzanas planted (#)	7.78	6.89	4.84	4.56	2.93	5.06
	Price	44.7	47.6	168.8	169.3	84.6	88.2
	N	50	59	49	52	42	55
Maize	Manzanas planted (#)	2.67	2.52	1.99	1.79	2.09	2.06
	Improved seed (%)	.244	.269	.254	.276	.158	.153
	Value of production	20710.6	20242	11488.53	10893.06	8486.9	9274.5
	N	544	454	539	449	535	448
Milk	Value, livestock production	267873	292260.3	297191	277464.9	236170.9	255429
	Value, milk production	112144.3	119956.9	170900.4	167027.3	167649.2	183283.5
	Processing (%)	.013	.028	.325	.315	.598	.507
	Price	4.24	4.18	6.64	6.46	6.81	6.81
	N	208	212	203	209	201	210

The asterisks denote the statistical significance of t-tests on the equality of the early and late complier group means: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

calculated using the prices that they were able to fetch for the part of their harvest that was sold[14]. The table also includes information on maize production. While maize was not an RBD activity, it is an important staple crop that most households produce and provides a signal whether RBD crop (and income) expansion come at the expense of reduced output and income from other crops. While market prices for maize and other non-target crops were not elicited, the total value of production is calculated based on baseline market prices provided by RBD program staff.

It seems clear from Table 1 that the effects of the program on production were quite diverse across the different target crops. For bean farmers, the initial effect seems quite straightforward: farmers who enrolled in the program early planted more beans, received higher prices at the mid-line than the control group, and in addition many more of them used improved seeds – and all these differences are highly statistically significant. By the end-line, we do see a worrisome drop-off in the use of improved seeds by early treatment producers. Recalling that RBD services to farmer groups ended after 24 months, this observation suggests that certain aspects of the program may not be sustainable. A similar story can be told for sesame farmers, except for the use of improved seeds[15].

For cassava and livestock, the differences between the groups are mostly statistically insignificant, even at mid-line. It may well be that both of these programs generated substantial spillover externalities in which even untreated and late treated farmers were able to benefit from the improved processing facilities offered to early treatment livestock and cassava producers.

As for maize production, the early and late treatment groups look similar both in the baseline and at the mid-line, which suggests that the early treatment group at the mid-line did not substitute away from maize in order to concentrate on target crops. This indicates that the RBD Program did not lead to significant

[14]When no sales price is reported (i.e. when the household did not sell any part of their crop), we use the mean price by season and crop.

[15]While the differences in the value of production for early/late sesame farmers is not statistically significant, the difference between the early and late groups is larger at the mid-line than at the end-line.

crowding-out of maize production, and constitutes suggestive evidence that measured income increases from targeted activities are less likely to greatly overstate overall income. With this in mind, we now turn to examining the program effect on farm incomes.

A. Program Income

We begin the evaluation of the RBD Program by looking at its impact on income from the activities targeted by the RBD program.[16] As discussed above, observed income increases in RBD-targeted crops does not necessarily imply increased overall incomes, as productive inputs may have been reallocated from other activities to the target crops. While we have evidence that productive inputs were not reallocated away from maize production, inputs could have been substituted away from other activities (*e.g.*, off-farm labor) that we do not measure. These concerns notwithstanding, the value of the production of target crops is an important indicator.

While Table 1 generally shows that early treatment farmers had significantly higher RBD incomes at mid-line than did the late treatment farmers, we now examine this impact more carefully by employing the generalized difference-in-difference binary treatment estimator from equation 4, presented in Section II.A above.[17] The double complier difference-in-difference impact estimate of δ_{DDC} is given by the coefficient on the treatment variable, Z, in the first column of Table 2. Recall that these estimates control for household-level fixed effects. The point estimate is \$1,211, which implies an increase of about 13% over the baseline level of targeted activity income. However, the precision of the estimate is such that we cannot reject the hypothesis that the true impact is zero.

[16] RBD targeted activities are beans, sesame, or cassava for farmers in those groups, and milk for livestock farmers. Income from these activities is the total value of production in the targeted activity, valued in 2005 \$USPPP.

[17] In this and all sections of our analysis, we estimate impacts using levels (measured in \$US PPP). In results not reported here, we also estimated the same relationship in logs. The log estimates were qualitatively similar to the level estimates and did not substantially alter the statistical significance of any impacts.

As we already saw in the descriptive statistics almost 90% of the individuals reported to be the farmer/beneficiaries of the RBD program are men (see Table B2 above). The second column of Table 2 also reports a specification in which we interact the sex of the farmer with the treatment variable in the regression. As can be seen, the estimated coefficient is small in absolute value (the impact of the program is $62 higher for women than for men) and is imprecisely estimated. There is no evidence that the program differentially influenced men and women in this dimension, although this finding may simply reflect the fact that few women were actually enrolled in the RBD program.

For reasons discussed earlier, this binary treatment impact approach may fail to fully characterize the program.[18] As a first step to exploring this, we replace the binary treatment indicator in equation4 with a flexible function of months in the program ($\Delta(d)$) as in equation 5. As a first step in this analysis, we use a generalized additive model model to estimate equation 5.[19] As can be seen in Figure D, the non-parametric analysis suggests that a cubic specification would provide a reasonable parametric approximation of the shape of the duration response function.

The third column of Table 2 report the results from parametric estimate of (7), our preferred cubic specifications. Consistent with the semi-parametric results, the parametric results show that duration of time in program has a statistically significant impact on gross income in the treated activity. To draw out more fully the implications of these estimates, Figure 4 graphs the estimated cubic relationship. For comparison, the point estimate of the binary treatment effect is illustrated as a solid horizontal line in the figure. As can be seen, the estimated impact rises over the first two years of time in the program, peaking at roughly a $2,400 impact, and then flattening out after that time. As shown by the confidence intervals in the figure, there is substantial noise in these estimates, and they

[18]In addition to those conceptual reasons, around 15% of farmers in the early treatment group had actually begun receiving RBD services at the time of the baseline, contaminating the binary results. We can control for this problem using the continuous treatment estimation.

[19]The results from this analysis can be found in Appendix D

TABLE 2—PROGRAM IMPACT ON INCOME FROM TARGET ACTIVITY FOR PROGRAM PARTICIPANTS

	Binary	Binary, gender	Cubic
t_2	1778***	1778***	1687.3***
	(421.57)	(421.67)	(314.11)
t_3	363.7	364.0	241.6
	(774.80)	(774.51)	(600.35)
Z	1211.7*	1207.2*	
	(652.07)	(649.28)	
Female$^*Z$		33.1	
		(671.56)	
Months (d)			263.2**
			(108.76)
Months^2 (d^2)			-10.2
			(6.90)
Months^3 (d^3)			0.112
			(0.12)
N	2001	2001	3062
R^2	0.045	0.045	0.308
adj. R^2	0.044	0.043	0.305

Cluster-robust standard errors in parentheses

Gender indicator variable equals 1 if the participating farmer is female

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

become especially imprecise as the data thins out at higher treatment levels. This noise may reflect the fact that the program had different impacts across the different RBD-targeted activities.[20] Nonetheless, there is a significant impact over treatment durations of 1 to 36 months.

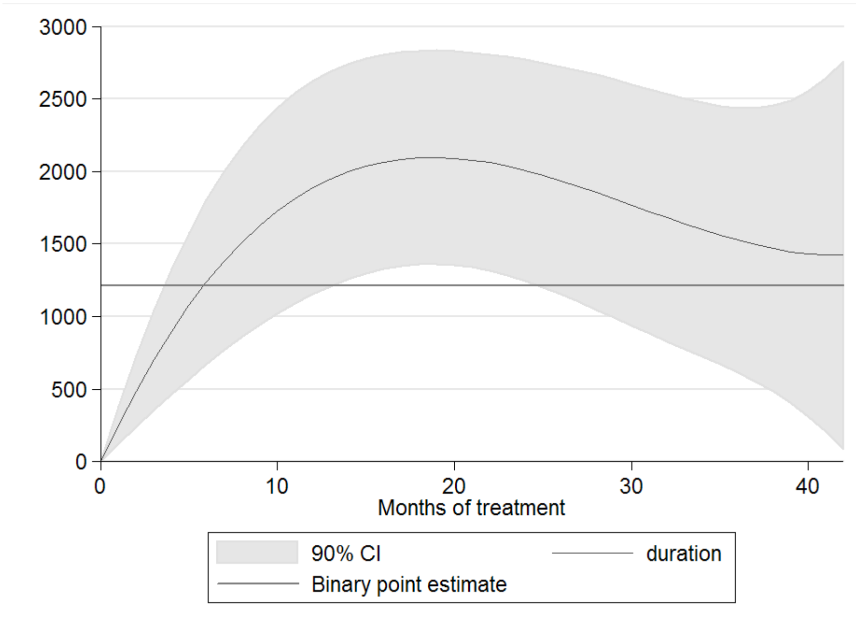


FIGURE 4. ESTIMATED IMPACT OF THE RBD ON PROGRAM INCOME

We now turn to see if these estimated gross income increases translated into increases in capital accumulation and/or consumption expenditures.

B. Investment

An important component of the business plan developed by farmers was related to the accumulation of farm assets. With the objective of increasing farmer's productivity, the program provided some equipment or supported the construction

[20]Discussions of this point with RBD staff in Nicaragua suggested that returns to some of the targeted activities declined after the initial program year (as subsidies for improved seed and fertilizers were no longer given to producers by the RBD). Unfortunately, the impact evaluation was not set up with sufficient power to distinguish impacts by specific crop activity.

of new productive installations once the business plan was approved. Therefore, weWhile the binary impact estimates (column 1, in Table 3) are not statistically significant, would expect the observed impact on incomes to go hand in hand with an increase in the stock of farm assets.

TABLE 3—PROGRAM IMPACT ON CAPITAL FOR PROGRAM PARTICIPANTS

	Binary	Binary, gender	Cubic
t_2	242.2	242.2	-531.1**
	(234.3)	(234.3)	(235.5)
t_3	4581***	4569.4***	2225.5***
	(478.1)	(478.1)	(584)
Z	503.6	660.2**	
	(311.5)	(322.9)	
Female*Z		-1106*	
		(566.3)	
Months (d)			292.8**
			(142.2)
Months^2 (d^2)			-7.9
			(10.6)
Months^3 (d^3)			0.08
			(0.18)
N	2076	2076	3147
R^2	0.137	0.138	0.206
adj. R^2	0.135	0.136	0.202

Cluster-robust standard errors in parentheses

Gender indicator variable equals 1 if the participating farmer is female

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We follow the same strategy used in the previous section to examine the effect of the program on stocks of capital. The outcome variable used here is the sum of mobile capital (tools and equipment, excluding livestock) and fixed capital (buildings, installations, and fences located on the farmer's land[21]), but the results are similar if disaggregated by type of capital.[22]

[21]Some elements of fixed capital were difficult to value as they were often constructed by the farmer rather than purchased on the market. RBD program staff assisted with the evaluation, but a few items (e.g. erosion barriers and certain types of fencing) are not included in our measure of fixed capital.

[22]The impacts of months in the program on mobile capital is stronger than on fixed capital. Results available from the authors upon request.

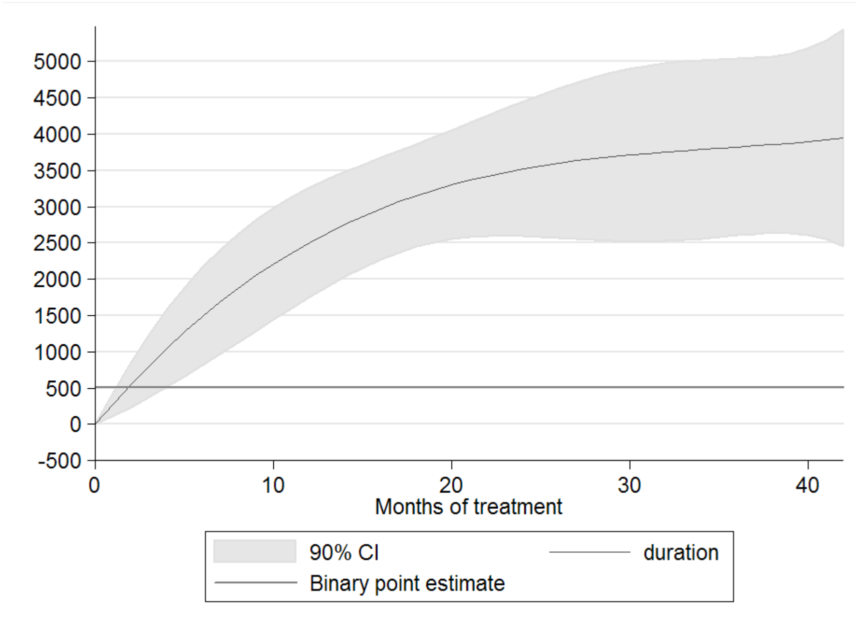


Figure 5. RBD Average Impacts on Capital

Table 3 shows the binary and continuous estimates of the program impact on capital stocks. Figure 5 graphs the average duration response of total capital. While the binary impact estimates (column 1, in Table 3) are not statistically significant, the continuous treatment model shows that the value of total capital increases significantly over the duration of the project, flattening out at more than $3,500.

Column 2 in Table 3 shows a regression where the program impact on capital accumulation is allowed to vary with the gender of the farmer. Despite the small percentage of female farmers, the estimated effect of program participation on mobile capital is significantly lower for female farmers. In addition, once the gender-specific effect if accounted for, the treatment effect becomes statistically significant. This suggests that male and female farmers respond quite differently to the program. Given that average impacts on capital start flattening out at around $3,500 dollars, the size of the coefficient is quite large as it is almost a third of the average maximum impact. In light of the absence of significant differences between the impacts on income for male and female farmers, this differential impact on investment is worth keeping in mind as we turn to the final impact variable of per capita consumption.

C. Consumption

As the final section of our analysis within the present framework, we will now look at per capita consumption. The consumption variable has been carefully constructed to take into account various aspects of household consumption and well-being. It is transformed to purchasing-power-parity adjusted US$. Further, to make it a per capita measure, it has been weighted by the number of household members (for expenditures other than food), and weighted by the number of household members that were present in the household during the appropriate recall period for food expenditures.

In order to examine the evolution of consumption estimates more closely, we report binary and continuous regression estimates in Table 4. Using the same binary regression model (4) as in the two previous sections (Column 1 in Table 4), we note that while the treatment effect is positive, it is neither statistically different from zero, nor very large. The point estimate of the binary impact implies a $187 increase in annual per capita expenditure for participant households. With around 5 people per-household on average, this point estimate would imply an increase in total household consumption expenditures of about $500 per-year.

Again, allowing impacts to differ by sex of the RBD program beneficiary does not reveal any significant gender-differentiated effects. However, the point estimate on the gender interaction term is several times larger than its counterpart in the income regression. Coupled with the size and significance of the gender-effect on capital, this suggests that perhaps women beneficiaries allocate program-induced income increases differently between consumption and capital investment than do men.

These binary estimates may again mask the underlying duration response pattern. Indeed, with consumption (as opposed to investment and program income), there are reasons to suspect an initial fall in consumption if households were to self-finance matched program investments by reducing consumption. As shown in Table B2, roughly 40% of the sample is reported to be capital-constrained in the sense of having unmet demand for loans they would like to take given the cost of capital. As can be seen in Appendix D, the semi-parametric results suggest the possible existence of an initial dip in living standards.

Table 4—Program Impact on Per Capita Consumption for Program Participants

	Binary	Binary, gender	Continuous
t_2	-476.5***	-476.5***	-380.6***
	(147.3)	(147.4)	(118.2)
t_3	-211	-208.9	87.1
	(227)	(226.5)	(251.1)
Z	186.5	157.8	
	(188)	(201.7)	
Female*Z		201.5	
		(303.7)	
Months (d)			-45.5
			(38.1)
Months^2 (d^2)			2.97
			(2.03)
Months^3 (d^3)			-0.05
			(0.03)
N	2123	2123	3198
R^2	0.006	0.006	0.196
adj. R^2	0.004	0.004	0.192

Cluster-robust standard errors in parentheses

Gender indicator variable equals 1 if the participating farmer is female

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column 3 of Table 4 report results from replacing the binary treatment indicator with a parametric cubic function of treatment duration. The individual coefficients are not statistically significant, although the key question is the statistical significance of the overall impact duration relationship.

Figure 6 displays the 90% interval estimate of the duration response implied

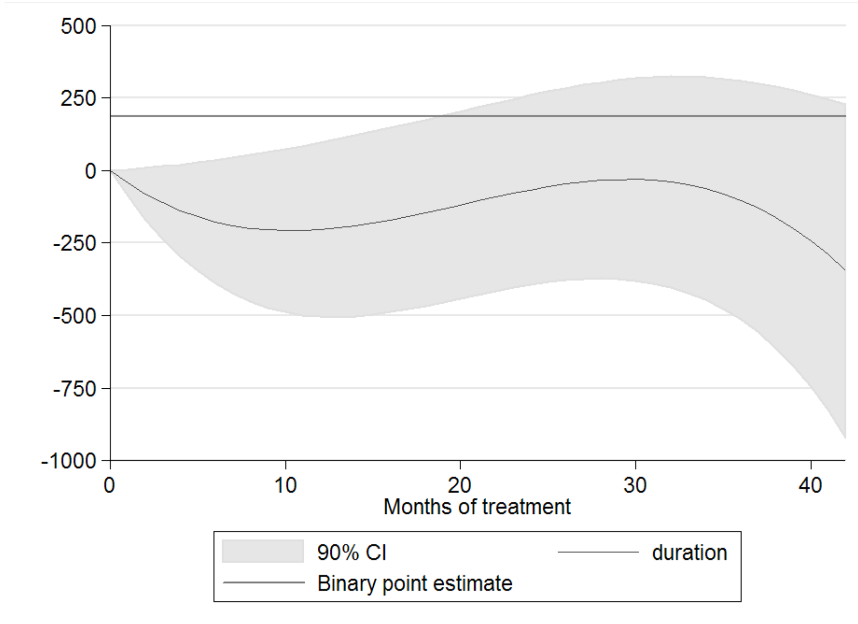


FIGURE 6. RBD IMPACTS ON PER CAPITA CONSUMPTION EXPENDITURES

by the cubic estimates in Table 4. As can be seen, the point estimates show no signs of consumption growth over the time of the program, and the interval estimator always includes zero. Given the findings of significant impacts on RBD-targeted income and on capital investment, the lack of a significant impact on living standards is somewhat surprising. There are several explanations for this, between which we cannot distinguish in this analysis. It could be that *total* income did not increase for beneficiary households (as opposed to income from only RBD-targeted activities). It is also possible that total income did increase, but that most of it was allocated to investment rather than increasing immediate living standards (especially for the predominantly male farming population). Finally, it could be that the impacts are quite heterogeneous, driven in part by differences in tenure security, capital access, etc. Section IV below returns to this issue, employing generalized quantile regression methods to investigate whether this lack of an average impact on per capita consumption levels reflects what is going on in different segments of the sample.

IV. Heterogeneous Program Effects

There are a number of reasons for believing that programs like the RBD may result in heterogeneous treatment effects. There are at least three candidate explanations as to why such heterogeneous treatment effects might occur:

- *Heterogeneous access to financial capital needed to make the most of the RBD intervention;*
- *Complementarity between unobservables (such as farming skills and business acumen) and the RBD intervention; and,*
- *Differential luck, with some succeeding and others failing for stochastic reasons.*

Earlier analysis conducted with only the mid-line data revealed substantial evidence of impact heterogeneity, with the program showing few impacts on the well-being of the poorest-performing 50% of the population (when compared against the poorest-performing segment of the untreated households), with quite high returns to the best performing segment of the treated group, when compared against top performers in the then untreated control group. Efforts reported in Toledo (2011) to unpack the reasons behind this heterogeneous performance are only partially satisfying. That analysis focused on explanation (1) above, categorizing households based on their credit-rationing status. While credit market status is of course endogenous, that analysis revealed no simple relationship between performance and contemporaneous credit rationing status. Indeed, the only factor uncovered was past credit history. RBD impacts on farms with prior credit history appeared quite large and significant. Unfortunately, the interpretation of prior credit history as a factor explaining heterogeneous program impacts is ambiguous. It seems most likely that those with past credit histories are actually those with higher levels of farming and business acumen (pointing toward explanation 2 above). It may also be that those acumen levels were themselves endogenously produced by prior random or prior program-based access to credit (and business opportunities). In this section, we examine in more depth the program effects for different segments of the population, based on the notion that the average treatment effects presented above may not tell the full story.

A. Methods and Interpretation

Conventional regression methods (such as those just employed above in Section 4) estimate average or mean relationships. They assume that the vector of covariates, x, affects only the *location* of the conditional distribution of y, not its scale nor any other aspects of y's conditional distribution. Quantile regression methods allow us to see whether the statistically average relationship is in fact a good description of the relationship in all parts of the distribution. Specifically, quantile regression allows us to recover the regression parameters that best describe the impacts on observations in different portions of the error distribution for our regression model.

Observations in the higher quantiles are those that "do better" than is predicted by the household's level of treatment and other regression variables (e.g., are in the upper tail of the conditional per capita consumption distribution). We will refer to these observations in the higher quantiles as "high performers." Conversely, observations in the lower quantiles are those are in the lower tail of the conditional distribution of the outcome variable. Quantile regression allows us to see if the marginal impact of RBD program participation at various parts of the conditional distribution of the outcome variables differs from the impacts at the mean – i.e. the average relationship estimated in Section 4. Note that if the average regression model explains the data well, the impact estimates will be the same for all quantiles. However, if there is unobserved heterogeneity in the impacts, then the impact slopes across quantiles may be different. As mentioned above, there are conceptual reasons (supported by the analysis of the mid-line data) for suspecting that the RBD program has heterogeneous impacts.

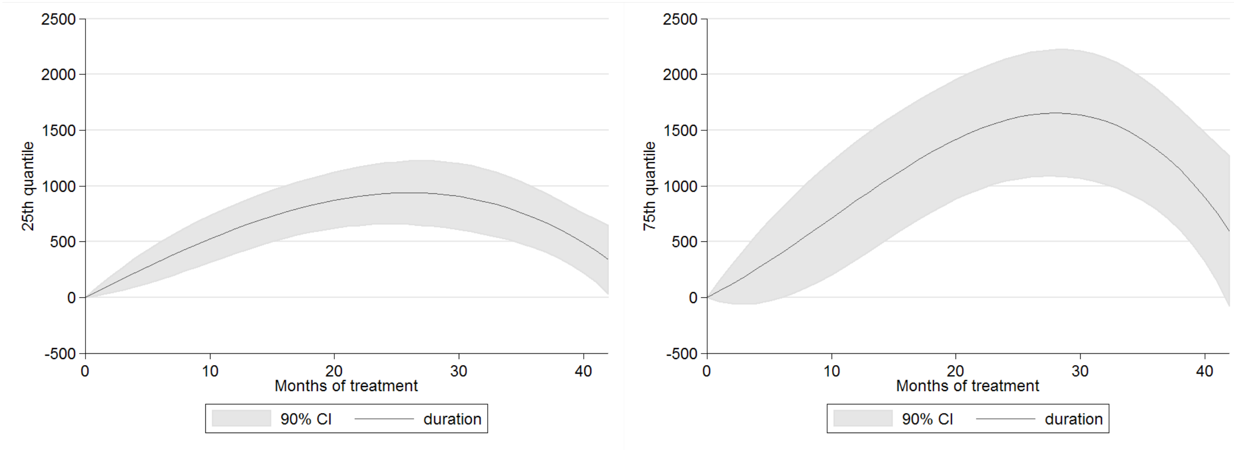
To recover conditional quantile estimates, we employ the method developed by Abrevaya and Dahl (2008) that extends a correlated random-effects framework (such as our regression model (6) above) to apply to conditional quantile models. While quantile models have been widely used in empirical studies since their development by Koenker and Bassett (1978), they are not often applied to panel data, likely because of the difficulty of differencing in the context of conditional quantiles. This problem arises because quantiles aren't linear operators, so that, simply put, the conditional quantile of a difference is not simply a difference of the conditional quantiles. Importantly, this methodology based on correlated random-effects preserves the fixed effects characteristics of the results, inoculating them against any systematic or spurious correlation between the duration of treatment and initial and time-invariant conditions.

Parameter estimates for the Abrevaya and Dahl (2008) estimator can be obtained through any quantile regression package. Standard errors are obtained through bootstrapping, drawing *households* with replacement from the sample and estimating the estimator's variance-covariance matrix from the resulting empirical variance matrix. All results in Section IV.B are based on 500 bootstrap repetitions.

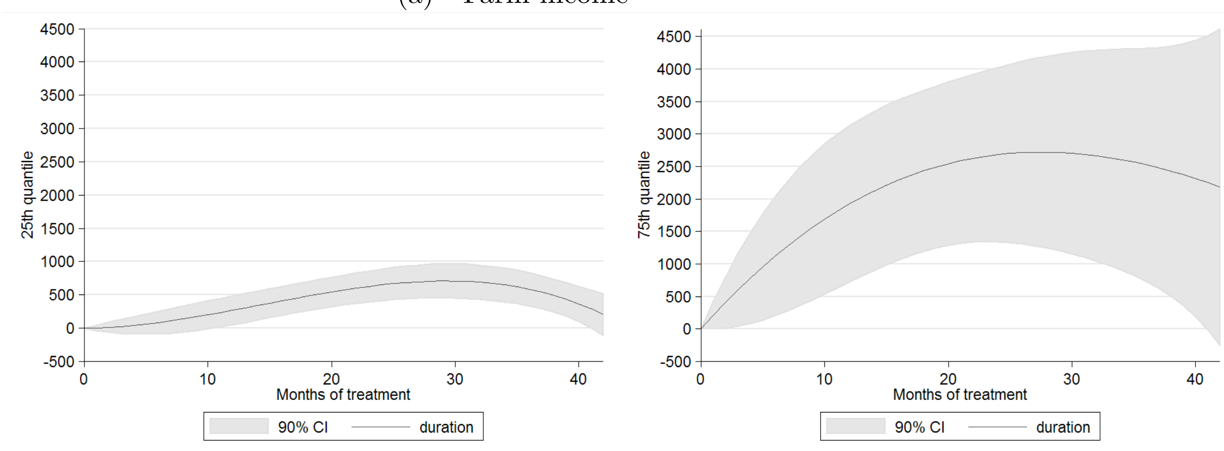
B. Quantile Results

This section explores the heterogeneity of the impact or duration response function by estimating the conditional quantile functions for our preferred (cubic) parametric continuous treatment models. In the interest of space, we present these results graphically, showing the point estimates for the 25th, 50th (median) and 75th quantiles. We represent the 90% confidence interval as a shaded area around the point estimates. The standard errors used to construct the interval estimates are based based on bootstrapped estimates (500 replications).

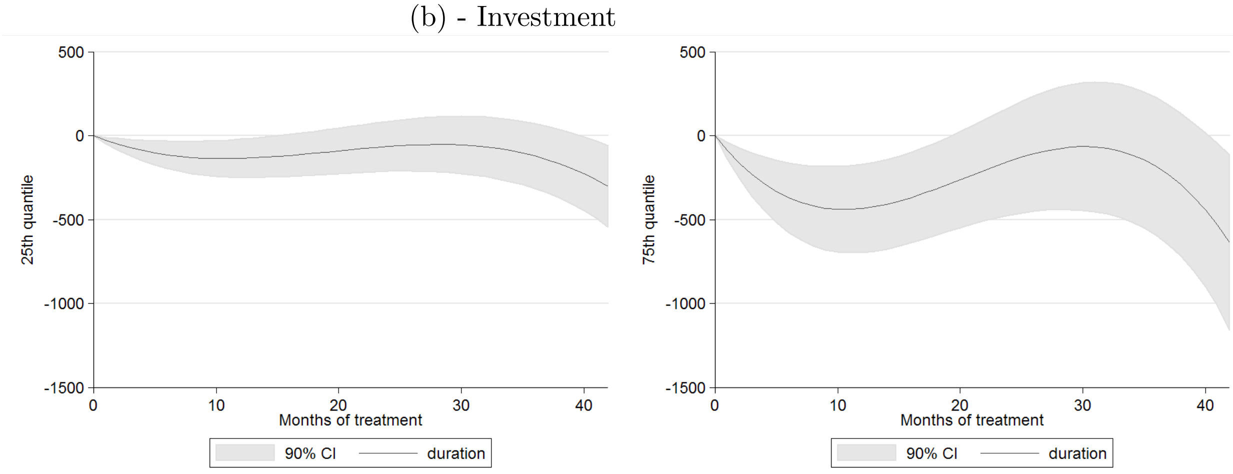
Figure 7 (a) displays the results from the quantile analysis of RBD-targeted



(a) - Farm income



(b) - Investment



(c) - Per-capita consumption

Figure 7. Generalized Quantile Impact Results

activity income. As can be seen, these estimates confirm the hypothesis that program impacts are heterogeneous across the participant population. The impacts of the program seem to be greatest at the extremes of the distribution, with the median impacts somewhat less positive than the conditional expectation function estimated in Section 4, and peaking earlier than the results at the mean of the conditional expectation function. The high performers in the upper 75th quantile enjoy a much steeper impact response function. Indeed, it peaks at about $2000, double the long-term impact level for the producer at the median of the conditional income distribution.

The effects of the RBD program on mobile capital (Fig. 7 (b)) increases as we move upwards in the conditional distribution of mobile capital – peaking at just under $300 for the lowest 25th percentile, around $1,200 at the median, and around $2,000 for the 75th percentile. Towards the end of the program, however, the amount of investment in mobile capital dips substantially for the 75th percentile, dropping down to or below baseline levels. The impacts on fixed capital can be seen in Figure , and show small, barely significant impacts for the 25th percentile and the median, but show substantial increases after 20 months since program enrollment for the 75th percentile of the conditional distribution of fixed capital.

We find significant heterogeneity in the impacts on per capita consumption, but the pattern of heterogeneity is quite different from the effects on, say, program income. The bottom panel in Figure 7 (c) shows the duration response paths for three different quantiles. The point estimates for those households at the lower end of the conditional per capita consumption distribution experience no significant increase in consumption for the duration of the program. For the median regression, consumption does appear to be increasing, but the impact is never statistically different from zero. In contrast, the high performers in the 75th quantile show very large increases in per capita consumption.

V. Conclusion

A key part of the 5-year, $175 million Nicaragua-MCC compact, the Rural Business Development Program (RBD) was designed to enhance the business knowledge of small farmers, and to improve their access to markets and technologies. RBD direct assistance lasted for 24 months, and the program also included elements of matching investment (*e.g.*, in improved milking sheds), and had average direct costs of about $US2500 per-farmer in the program. Participation in the program was subject to both administrative filters (eligibility criteria and business plan approval) and to beneficiary self-selection (eligible producers had to be willing to join and provide required matching investments).

In order to evaluate the impacts of the RBD, the evaluation and implementation teams worked together to create a randomized program roll-out strategy. No eligible households were denied access to the program, but the temporal sequence in which they received the program was randomized as eligible households were split into early and late treatment groups. Three rounds of data were collected: a baseline in 2007, a mid-line in 2009 and an end-line in 2011.

This evaluation strategy affords several advantages. First, shortly after the mid-line survey, it was possible to fully identify a two-sided complier sample as all eligible households in both early and late treatment groups had either accepted or declined the invitation to join the RBD by that date. The analysis has focused on this sub-population as it those who are inclined to join such a program that we are interested in.

Secondly, this research design randomized the duration of time that any particular producer had in the program. Using continuous treatment methods, we have been able to recover temporal time path of impact. Doing so is especially important for programs like the RBD that are intended to spur learning and co-investment, meaning that their impacts are likely to evolve over time. Given that we did not know ex ante how long it would take for these effects to take place, the continuous treatment methodology allows us to recover much more information than could have been obtained with standard binary treatment estimator approaches.

Analysis of these data indicates that RBD likely increased the incomes of most, but not all participants. Impacts built up over time, peaking and leveling off after the individual has been in the program for 2 to 2.5 years. It is important to stress that income increases in targeted activities are an *upper*-bound estimate of the change in overall household income, and hence the 39% average internal rate of return figure is also an upper-bound estimate. Note also that this figure is an average across all RBD-targeted activities that were included in the study. It is likely that returns were higher for some activities, and lower in others. The impact evaluation was not, however, designed to estimate activity-specific impact estimates nor rates of return.

While the impacts on income in targeted activities are substantial, the spillover of these income increases into improved living standards appears to be at best modest, at least over the time frame of the evaluation. Evidence that stocks of agricultural capital (especially movable capital) increased significantly with the program is consistent with this sluggish consumption response and would seem to indicate that households face binding credit constraints and an inability to borrow in order to finance investments and smooth consumption over time.

In addition to these average effects, the study employed fixed effects analogue conditional quantile regression methods to explore the degree to which the average pattern of impact faithfully reflects the experience of both high performers (those who do better than the OLS regression average) and low performers (those who do less well than the OLS regression average). Looking at the full distribution of impacts is especially important for efforts like the RBD program that target beneficiaries' income-generating and entrepreneurial capacities. In general, we find that the program is much more effective for the high performing households. Indeed, the upper quantile, high performing households exhibit a larger impact on their income in targeted activities, and their observed household living standards (as measured by per capita consumption expenditures) increase significantly 2-3 years after joining the RBD program. In contrast, the lower quantile households show no increase in living standards, even after 3-4 years in the program.

Roughly speaking, two-thirds of eligible farm households chose to participate. The one-third that did not participate had modestly lower living standards at baseline.[23] Of those that chose to participate, roughly three-quarters appeared to have benefited, while the remaining minority benefited little, if at all from the RBD. The existence of these two groups (those that did not participate, and those that did, but did not succeed) serves as a useful reminder that maybe not all small farms can upgrade and succeed. If the goal is to eliminate rural poverty, then this limitation needs to be kept in mind as other interventions may be needed to improve prospects for this sub-population and their children.

Looking forward, it may be that next generation RBD programs can reduce the size of this minority. While the analysis here was unable to identify which families failed to succeed and why,[24] it is likely that some failures were do to the natural vagaries of agriculture as a risky activity. Efforts to incorporate elements of insurance into small farm development strategies may have a key role to play in this regard, allowing a greater percentage of the small farm population to succeed over the longer term.

In addition, the RBD program did not include a direct credit market intervention. The overall MCC program in Nicaragua operated in part on the theory that improved property registration would indirectly improve smallholder access to capital by increasing their collateral and creditworthiness to the extant banking sector. Whether or not that strategy would have worked remains an open

[23] Recent work by Macours and Vakis (2008) and Laajaj (2012) on poverty and aspirations suggest that there may be some individuals who could benefit from interventions such as the RBD, but that they need smaller, confidence and aspiration building steps before they are willing to jump into a more forward-looking and entrepreneurial profile.

[24] One important message that emerged from the mid-line evaluation is that there is no evidence that farms closer to the asset minima benefited less from the program than did better endowed farmers (Carter and Toledo, 2011). While the asset floors and ceilings used to establish RBD eligibility were based on best practice intuition, it is clear from a targeting perspective that more work needs to be done to see if there is such a thing as a farm that is too small to benefit from this kind of intervention.

question, as the property registration component of the program was eliminated in early 2009 (see note 2 above). What is clear is that the pattern of increasing income, but sluggish changes in living standards (and indeed, perhaps a small initial drop in household living standards following the introduction of the program), may signal the existence of capital constraints as income increases are soaked up to self-finance future fixed and working capital investments.

References

Abrevaya, Jason, and Christian M Dahl. 2008. "The Effects of Birth Inputs on Birthweight." *Journal of Business & Economic Statistics*, 26(4): 379–397.

Aguero, J., M. R Carter, and I. Woolard. 2010. "The impact of unconditional cash transfers on nutrition: The South African Child Support Grant." Working Paper.

Anderson, J., and G. Feder. 2003. "Rural extension services." *World Bank Policy Research Working Paper No. 2976.*

Ashraf, Nava, Xavier Giné, and Dean Karlan. 2009. "Finding Missing Markets (and a Disturbing Epilogue): Evidence from an Export Crop Adoption and Marketing Intervention in Kenya." *American Journal of Agricultural Economics*, 91(4): 973–990.

Buchinsky, M. 1998. "Recent advances in quantile regression models: a practical guideline for empirical research." *Journal of Human Resources*, 88–126.

Carter, Michael, Patricia Toledo, and Emilia Tjernström. 2012. "Evaluating the Impact of Business Service on the Well-Being of Small Farmers in Nicaragua." Working Paper.

Chamberlain, G. 1982. "Multivariate regression models for panel data." *Journal of Econometrics*, 18(1): 5–46.

Chamberlain, G. 1984. "Panel Data. Handbook of Econometrics." *Griliches and M. Intriligator, eds.*

Conley, T. G., and C. R. Udry. 2010. "Learning about a new technology: Pineapple in Ghana." *The American Economic Review*, 100(1): 35–69.

Evenson, Robert E., Bruce L. Gardner, and Gordon C. Rausser. 2001. "Chapter 11 Economic impacts of agricultural research and extension." In *Agricultural Production*. Vol. Volume 1, Part A, 573–628. Elsevier.

Feder, Gershon, Roger H. Slade, and Lawrence J. Lau. 1987. "Does Agricultural Extension Pay? The Training and Visit System in Northwest India." *American Journal of Agricultural Economics*, 69(3): 677–686.

Foster, Andrew D., and Mark R. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy*, 103(6): 1176–1209. ArticleType: research-article / Full publication date: Dec., 1995 / Copyright © 1995 The University of Chicago Press.

Griliches, Z. 1957. "Hybrid corn: An exploration in the economics of technological change." *Econometrica, Journal of the Econometric Society*, 501–522.

Hirano, Keisuke, and Guido W. Imbens. 2004. "The Propensity Score with Continuous Treatments." In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives.* , ed. Andrew Gelman and Xiao-Li Meng, 73–84. John Wiley & Sons, Ltd.

Keswell, M., and M. Carter. 2011. "Poverty and Land Distribution."

King, Elizabeth M., and Jere R. Behrman. 2009. "Timing and Duration of Exposure in Evaluations of Social Programs." *The World Bank Research Observer*, 24(1): 55–82.

Koenker, R., and G. Bassett. 1978. "Regression quantiles." *Econometrica: journal of the Econometric Society*, 33–50.

Laajaj, R. 2012. "Closing the Eyes on a Gloomy Future: Psychological Causes and Economic Consequences." Working Paper.

Macours, K., and R. Vakis. 2008. "Changing households' investments and aspirations through social interactions: Evidence from a randomized transfer program in a low-income country." *World Bank Research Paper. World Bank, Washington, DC.*

Mundlak, Y. 1978. "On the pooling of time series and cross section data." *Econometrica: journal of the Econometric Society*, 69–85.

Royston, P., and G. Ambler. 1998. "Generalized additive models." *Stata Technical Bulletin*, 42: 38–43.

Toledo, Patricia. 2011. "Impact Evaluation of a Rural Business Program Using Field Experiment Data." Manuscript.

World Bank, the. 2008. "Nicaragua - Poverty Assessment (Vol I: Main report)." 39736, Washington, D.C.

Appendix A - Eligibility Criteria by Productive Activity

Table A1—Eligibility Criteria Used to Identify Farmers in Target Activities

	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 in cassava	Developed livestock activity
Water	−−	−−	On-farm water source	−−	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

Appendix B - Balance Checks

Table B1—Early versus Late Treatment Groups–All Eligible Farmers

	Livestock		Beans		Sesame		Cassava		Vegetables		All Farmers	
	Early	Late	Early	Late	Early	Late	Early	Late	Early	Late	Early	Late
Expenditures per capita	2407	2450	1194	1229	1803*	1556*	1686	1465	1462	1633	1831	1818
Mobile farm asset	69.1	69.2	9.7	10.0	59.9**	29.9**	51.1	51.1	25.6	31.0	46.6	42.6
Farm's installations	36.8	36.7	6.9	5.4	19.8	12.9	25.0	22.6	10.5	23.9	22.8	21.8
Animals	292.1**	344.7**	47.5	46.3	99.4	89.0	75.8	91.3	22.8	28.2	154.6**	178.8**
Monthly remittances	660	751	482	406	716**	278**	404	334	121	590	576	519
Land size	61.6	72.2	18.4	18.4	28.5**	22.2**	20.7	22.9	6.9	11.8	37.0	41.1
Formal tenure	55%**	63%**	21%*	15%*	48%	43%	48%	54%	43%	56%	42%	45%
In process to tenure	22%	28%	32%	34%	25%	32%	19%	20%	30%	13%	25%	25%
Credit status (%)												
With a loan	38%	42%	34%	25 %	66%	67%	41%	30%	50%	50%	42%	41%
No loan-price rationed	20%	24%	15%	23%	9%	9%	15%	21%	4%	21%	16%	20%
Quantity rationed	16%	10%	12%	12%	17%	11%	22%	14%	29%	13%	16%	12%
Risk rationed	26%	24%	39%	40%	8%	14%	22%	35%	17%	17%	26%	28%
Farmer's age	52*	54*	49	50	49	50	50	53	48	49	50	52
Farmer's education	4.9	5.0	3.4	3.3	4.6**	3.0**	3.3	2.9	4.8	5.4	4.3	4.9
Farmer male	87%	88%	88%	81%	88%	93%	82%	85%	63%	75%	86%	86%
# of household	5	4	5	5	4	4	4	6	6	6	4	4
# of household	1	1	2	2	1	0	1	0	0	1	1	1
# of observations	308	311	240	218	147	139	82	89	24	24	801	781

* Different tests used for different data types: t-tests for categorical variables, and Kolmogorov-Smirnov for continuous variables
For any test, an asterisk $*$ indicates that the p-value $[5\%, 10\%]$; $**$ indicates p-value $[1\%, 5\%]$; while $***$ means p-value $\leq 1\%$

Table B2—Early versus Late Treatment Groups–Complier Farmers Only

	Livestock		Beans		Sesame		Cassava		All Farmers		Excluding Beans	
	Early	Late	Early	Late	Early	Late	Early	Late	Early	Late	Early	Late
Expenditures per capita	2407	2308	1193	1326	1708	1511	1745	1622	1811	1866	2127	1969
Mobile farm asset	81.5*	62.3*	10.4	15.5	47.6	30.4	59.4	54.1	50.0	46.1	69.6***	51.9***
Farm's installations	43.7	33.6	7.6	8.3	21.3	13.5	25.2	19.2	26.0	22.9	35.3**	25.6**
Animals	325.6	370.4	51.2	44.8	97.1	89.4	87.6	81.8	169.8**	212.5**	232.4	244.1
Monthly remittances	734	519	468	424	608*	315*	366	377	575	436	647*	438**
Land size	69.9	75.7	19.8	22.9	29.8**	21.2**	22.7	22.5	41.1	46.1	52.7	51.8
Formal tenure	56%***	67%***	20%	18%	44%	41%	42%	52%	41%***	51%***	51**%	57**%
In process to tenure	25%**	16%**	31%**	44%**	27%	33%	17%	21%	27%	25%	24%	22%
Credit status (%)												
With a loan	42%	43%	36%	22%	67%	69%	47%	27%	45%	44%	49%	48%
No loan-price rationed	21%	23%	17%	24%	8%	6%	25%	15%	16%	19%	16%	11%
Quantity rationed	12%	10%	12%	11%	19%	10%	10%	5%	15%	11%	17%	19%
Risk rationed	25%	24%	36%	43%	6%	15%	12%	32%	24%	26%	18%	23%
Farmer's age	52*	54*	49*	52*	47	50	48**	54**	50***	53***	50***	53***
Farmer's education	5.1	4.9	3.7	3.8	4.9***	3.0***	4.3**	2.9**	4.5	4.4	4.9***	4.0***
Farmer male	87%	89%	86%	81%	91%	93 %	80%	85%	86%	88%	87%	90%
# of household	5	4	4	5	4	4	4	4	5	4	5	4
# of household	1	1	1	2	1	0	1	0	1	1	1	1
# of observations	198	208	162	72	86	108	51	61	511	449	33}5	377

See notes for Table B1

Appendix C - Binary ITT Estimates Using the Full Sample

Table C1 shows estimates of the intention-to-treat and instrumental variable estimates of the average treatment effect using the baseline and mid-line data. The intention-to-treat is estimated by the following regression:

$$y_{it} = \lambda t_2 + \gamma B_i + \delta(B_i^* t_2),$$

where t_2 is a dummy variable taking on the value 1 in round 2, and

$$B_i = \begin{cases} 1 & \text{for farmers randomized into the early treatment group} \\ 0 & \text{for farmers randomized into the late treatment group} \end{cases}.$$

This regression is estimated using the whole sample. y_{it} here is either per capita consumption, farm income or total capital. As in much of the main analysis, the parameter δ is the parameter of interest.

Table C1—ITT and IV Estimates

	Consumption		**Income**		**Total Capital**	
	ITT	IV (ToT)	ITT	IV (ToT)	ITT	IV (ToT)
λ	-516.3***	-516.2***	1479.0***	1479.1***	-65.35	-66.00
	(196.4)	(196.1)	(552.6)	(551.8)	(524.7)	(524.1)
γ	-86.55	-134.2	-467.1	-712.1	1279.5**	1984.5**
	(206.4)	(318.0)	(452.5)	(692.6)	(604.2)	(928.7)
δ	-5.004	-7.563	834.0	1269.7	495.9	754.9
	(248.8)	(381.6)	(742.2)	(1127.7)	(851.3)	(1300.4)
N	3179	3179	3040	3040	3122	3122
R^2	0.218	0.218	0.304	0.304	0.162	0.164
$\overline{R}^2$	0.216	0.215	0.302	0.302	0.159	0.161

Standard errors in parentheses

Regression coefficients for control variables (crop, farmer age and education) are suppressed.

The first and third columns of Table C1 reports standard Intention to Treat estimates (ITT) for consumption and income, respectively. For the Treatment on the Treated (ToT) estimates, we use the standard instrumental variables technique, instrument for treatment, d_i, using the assignment to early treatment, B_i. The instrument for the interaction term between treatment and the mid-line time-dummy variable ($d_i^* t_2$) is simply $B_i^* t_2$.

The ITT estimates are consistent with the average treatment effects reported in the main body of the paper, and while both the ITT and ToT estimates for consumption are negative, it's important to note that given the wide confidence intervals, the estimates are consistent with the small but positive impacts found above, as well with a zero impact.

Appendix D - Semi-parametric Estimates of the Continuous Treatment Estimator

As explained in the body of the paper, there are a number of reasons to expect that the impact of the RBD program to evolve over time in highly non-linear ways. As a prelude to the parametric continuous treatment analysis reported in Sections II.Band III, we first used flexible semi-parametric methods to explore the shape of the impact curve. We estimated equation 5 by allowing the time variables to enter parametrically, while the duration function was estimated non-parametrically. Estimation was carried out using a Generalized Additive Model (GAM). The way we specify the GAM model, baseline characteristics (education, age and gender of the farmer) and time dummies enter linearly, and a flexible function $s(d_{it})$ is fitted for the time since treatment began. The following equation shows the form of a GAM with q predictors:

$$g(\ ) = \ + \sum_{j=1}^{q} s_j(x_j)$$

where the s_j's are smooth functions of the predictors. Each predictor is assumed to be additive on the scale of $g(\cdot)$. The GAM-implementation that we use employs cubic smoothing splines to estimate the flexible function $s(\cdot)$.[25] The figures below display the results of these estimates. These were estimated using only those observation that began treatment after the baseline survey round.

As can be seen, the semi-parametric regression on consumption suggests that the impact relationship can be captured by a cubic relationship. The relationship for income suggests that a quadratic may suffice, but we choose a cubic parametrization in the analysis (the coefficient on the third degree polynomial would simply be estimated to be 0 if a quadratic were the correct fit). None of these relationships are especially supportive of the step function in Figure 2 that

[25] Stata -gam- add-on, which is based on a FORTRAN program called GAMFIT, and documented in Royston and Ambler (1998)

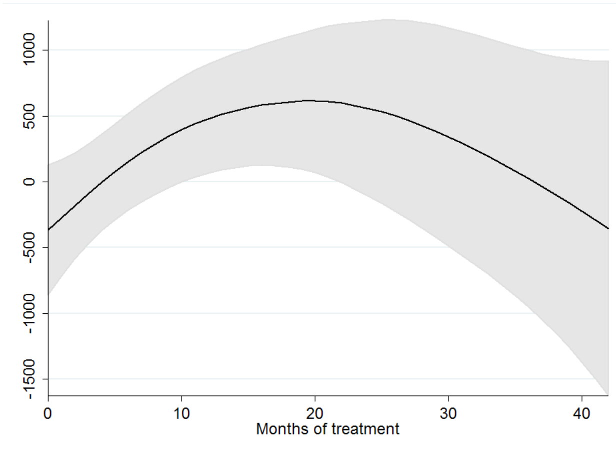


Figure D1. Semi-Parametric Results, Income

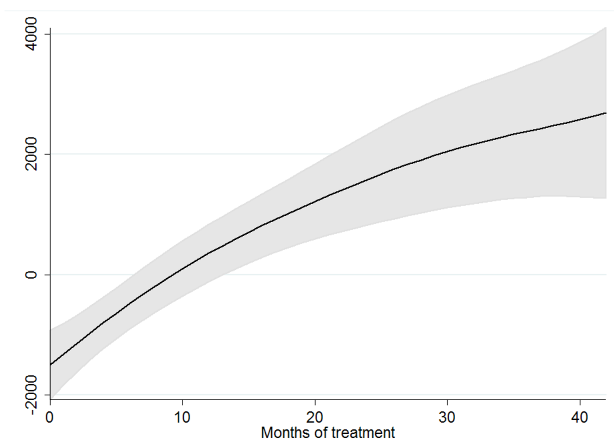


Figure D2. Semi-Parametric Results, Total Capital

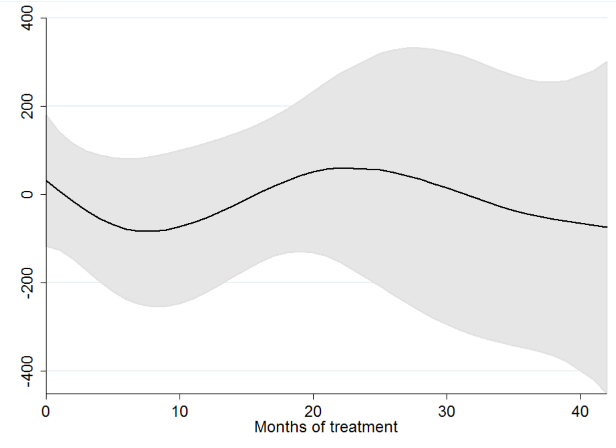


Figure D3. Semi-Parametric Results, Per Capita Consumption

would be implied by the standard binary treatment estimator.

Appendix E - Hypothesis Testing

To test whether the differences that we observe between the different quantiles are statistically significant, we employ two different testing procedures. First, we use the minimum-distance framework in Abrevaya and Dahl (extended from Buchinsky's (1998) framework to the panel data context) to test the equality of the parametric duration response variables' effects across quantiles. Since both *months*, $months^2$, and $months^3$ enter into our preferred cubic model, the relevant test is a joint test of equality. In other words, the null hypothesis is

$$H_0: \ \zeta_{1,\tau_1} = \zeta_{1,\tau_2} = \zeta_{1,\tau_3} \wedge \zeta_{2,\tau_1} = \zeta_{2,\tau_2} = \zeta_{2,\tau_3} \wedge \zeta_{3,\tau_1} = \zeta_{3,\tau_2} = \zeta_{3,\tau_3},$$

where ζ_1,ζ_2 and ζ_3 are the estimated coefficients on *months*, $months^2$, and $months^3$, respectively, and τ_1, τ_2 and ,τ_3 are the different estimated quantiles ($25^{th}, 50^{th}$ and 75^{th}).

Second, we take linear combinations of the estimated coefficients from an interquantile range regression, which estimates the *difference* in coefficients between quantiles. We then plot their bootstrapped normal-approximation confidence intervals (500 bootstrap replications). For the intervals where the confidence intervals don't overlap 0, we say that the quantiles are statistically different from each other at those months. This test is carried out for each of the differences $75^{th} - 25^{th}, \ 75^{th} - 50^{th}, \ 50^{th} - 25^{th}$.

Minimum-Distance Testing Results

We closely follow the testing framework outlined in Abrevaya and Dahl (2008), and the only changes are to allow for an additional round of data and the fact that we include averages of the time-varying regressors, instead of their value in each round. The minimum-distance test statistic has a limiting chi-square distribution, with degrees of freedom equal to the number of restrictions (in our case 6). These test statistics and their associated p-values are shown in Table E1

The estimated effect of months in the program on income and total capital

Table E1—Tests of marginal-effect equality across quantiles

Outcome variable	χ^2_6-statistic	p-value
Income	10.7	0.0971
Total capital	20.6	0.0021
Per-capita consumption	9.1	0.1678

For each outcome variable, the p-values reported are for the null hypothesis of joint equality of the marginal effects of the variables *months*, $months^2$, and $months^3$ for the quantiles .25, .50 and .75. Results are based on 300 bootstrap replications

vary significantly across the quantiles, while the effects on total capital are not statistically significant at conventional levels. There is no statistical evidence that the effect of months in the program on per capita consumption varies over quantiles.

Bootstrapped Normal-Approximation Confidence Interval

Figures E1, E2 and E3 show the point estimates and confidence intervals of the difference in coefficients between pairs of quantiles.

Consistent with the minimum-distance test, there is no evidence that the effect of months since treatment began on per capita consumption varies across the three quantiles (Figure E3 – the 90% confidence interval includes 0 everwhere. For income, the effect of time in the program does not vary between the 25^{th} and 50^{th} quantile (Figure E1 (a)), but the effect of treatment-months is significantly greater, and by a large magnitude, at the 75^{th} percentile than at the50^{th} percentile (Figures E1 & (c)). The difference in effects becomes significant about 20 months after treatment begins. Comparing the effects of months in program on capital, we can see that the most significant differences in effects show up between the 25^{th} and 50^{th} percentiles, and the 25^{th} and 75^{th} percentiles (Figure E2 (a) & (b)). There is no statistically significant difference between the effect of months on capital investment at the 50^{th} and 75^{th} quantiles (Figure E2 (c)). Overall, these are consistent with the minimum-distance test in the previous section, but they provide some additional information about what quantiles drive the differences.

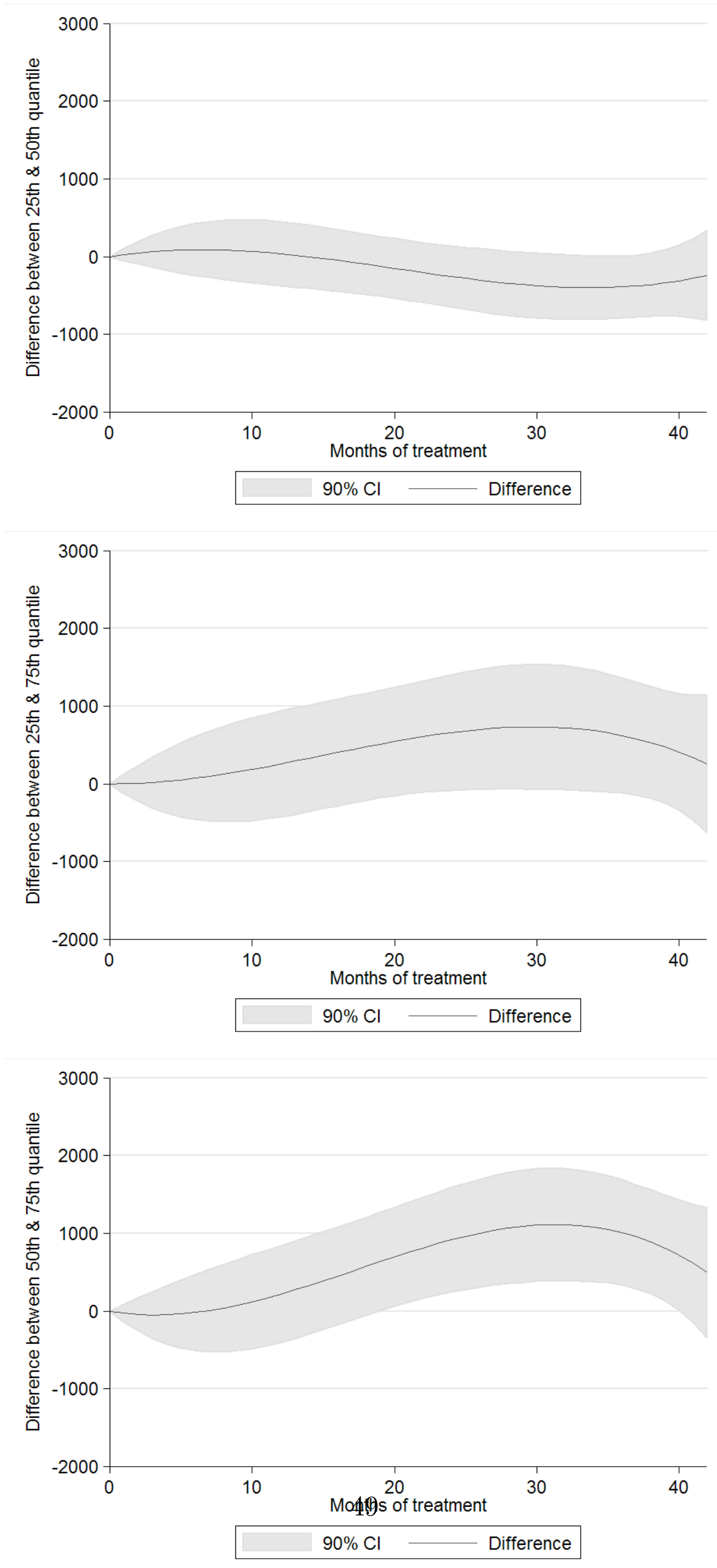


FIGURE E1. DIFFERENCES ACROSS QUANTILES, INCOME

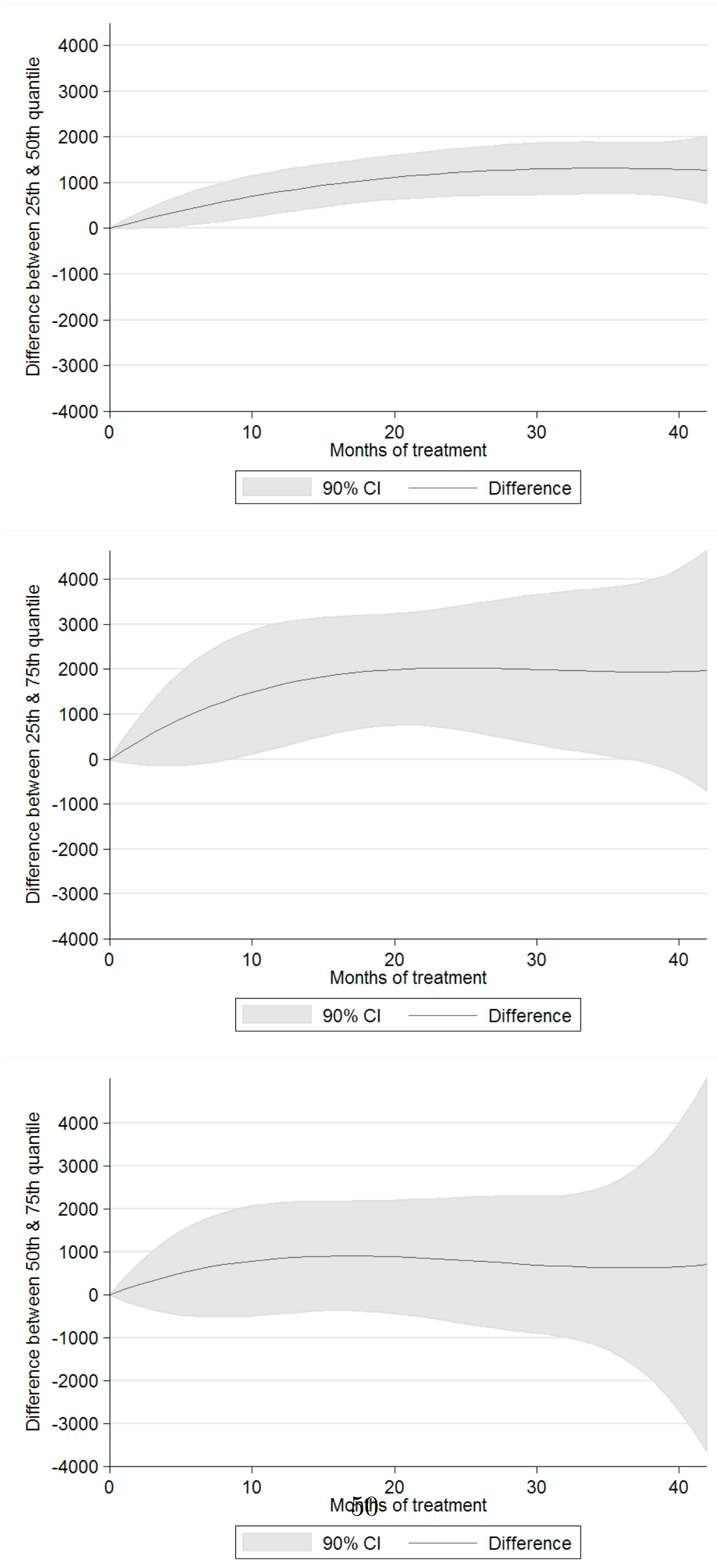


Figure E2. Differences across quantiles, capital

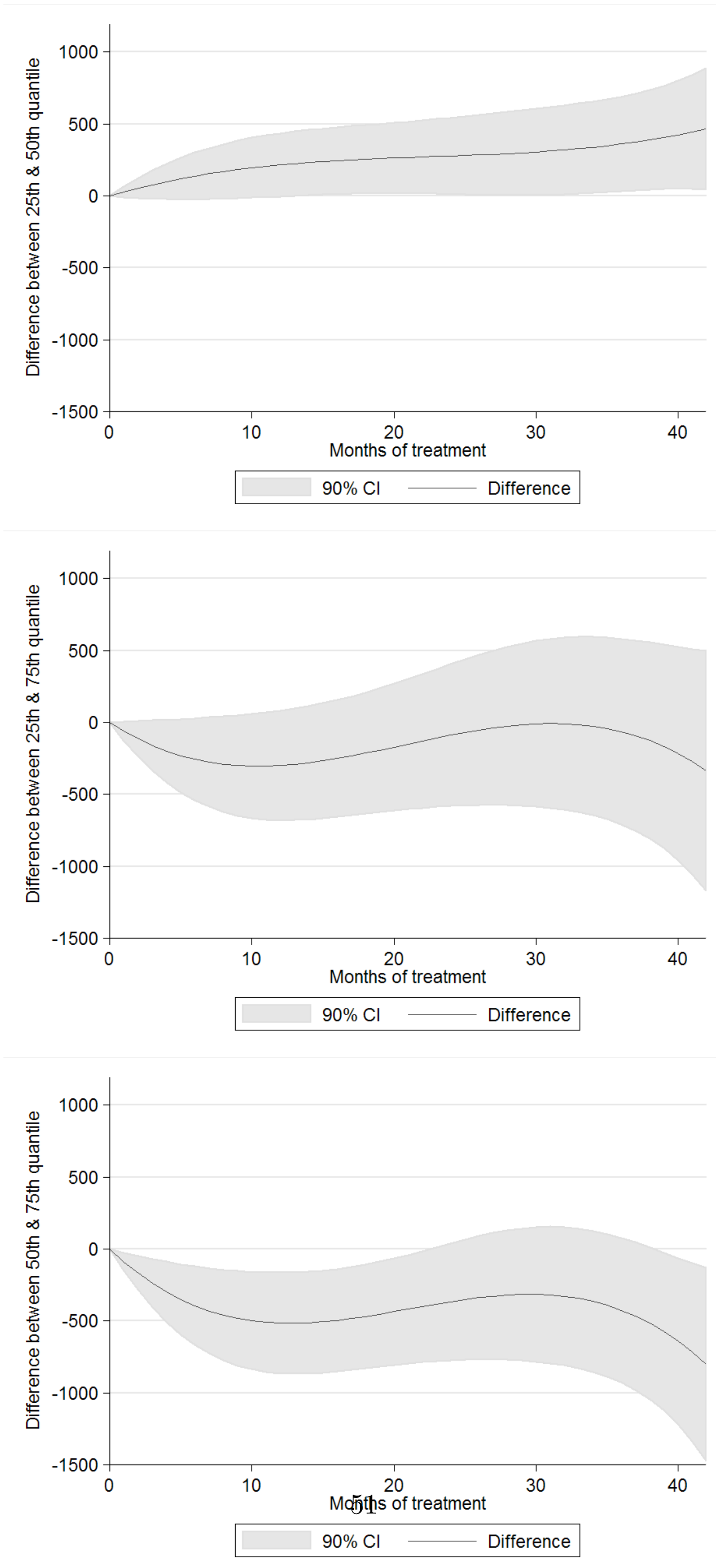


Figure E3. Differences across quantiles, consumption