Does microfinance reduce rural poverty?
Evidence based on household panel data from northern Ethiopia

Guush Berhane and Cornelis Gardebroek

Agricultural Economics and Rural Policy Group, Department of Social Sciences,
Wageningen University, The Netherlands
Hollandseweg 1
6706 KN, Wageningen, The Netherlands
guush.berhane@wur.nl
koos.gardebroek@wur.nl

Contributed Paper prepared for presentation at the International Association of Agricultural Economists Conference, Beijing, China, August 16-22, 2009

Abstract

This paper evaluates the long-term impact of microfinance credit from the intensity of participation in borrowing. We use a four-round panel data set on 351 farm households that had access to microfinance in northern Ethiopia. Over the years 1997-2006, with three-year intervals, households are observed on key poverty indicators: improvements in annual consumption and housing improvements. The relatively long duration in the panel enables to measure household poverty changes between consecutive periods and see the long-run effects of exposure to microfinance from the intensity of participation borrowing. The fixed-effects model is innovatively modeled to account for potential selection biases due to both time-invariant and time-varying unobserved individual household heterogeneities. Results show that microfinance borrowing indeed causally increased consumption and housing improvements. A more flexible specification that allows for the number of times the household has been in borrowing also shows that repeated borrowing is effectively increasing consumption: the longer the borrowing relationship the larger the effect partly due to lasting credit effects. Impact estimates that do not account for such dynamic effects may therefore undermine the effect of MFI borrowing.

Key words: *Microfinance, treatment effects, trend model, panel data*

1 Introduction

The microfinance revolution got considerable momentum around the world in the last two and half decades. The potentials of microfinance as an effective tool to break the vicious circle of poverty has been widely voiced. As a result, several microfinance schemes have gone operational around the world, providing financial access to millions of poor people both in rural and urban areas. Important questions are however if and to what extent microfinance credit over its long time existence has contributed in reducing poverty.

Despite efforts to measure this impact, evidence on the poverty reduction effects of long term microfinance credit remains unclear mainly due to the difficulty of measuring counterfactual outcomes and the lack of follow up data spanning over sufficiently long periods to measure the impact. Without experimental designs, evaluations based on simple comparisons between participants and non-participants are subject to biases from two sources (e.g., Pitt and Khandker, 1998; Ravallion, 2001). The first bias is due to program placement and occurs because microfinance institutions (MFIs) do not randomize over villages to place programs. They often choose on village characteristics that may not be observable to the researcher. The second bias is due to the tendency of individual borrowers to self-select into programs. From the nature of borrowing it is evident that potential applicants can choose themselves to apply for a loan. When selection into the program is based on unobservable individual attributes (e.g. entrepreneurial ability) that simultaneously affect the impact outcome, attributing observed differences to credit gives biased impact estimates.

But even if pre-designed experimental or quasi-experimental designs that randomize over potential sources of selection are implemented, estimates based on one-shot observations may fall short of capturing the complete picture because longer periods may be required before the full effects from credit are realized (Karlan and Goldberg, 2007). A recent review of the evaluation literature emphasizes the issue of ‘timing and duration of exposure to programs’ is as important but relatively less studied than the identification problems that often attract much of researchers’ attention (King and Behrman, 2009). Long period data is, however, costly and largely unavailable. As a result, most studies so far (e.g., Coleman, 1999; Pitt and Khandker, 1998) exploited program specific designs and employed innovative quasi-experimental survey methods to generate control and treatment groups from cross sectional data. A few exceptions are Khandker (2005), Copestake et al. (2005) and Tedeschi (2008) who used two-period data to estimate impacts. Long-term panel data, under certain conditions, allows to measure impact from *intensity of participation* over time by overcoming selection biases. An attractive feature of panel data is the possibility to deal with unobserved *time-invariant* individual and village heterogeneity using fixed-effects. However, when the selection processes is based on time-varying unobservables, such as individual motivation which is likely to change over time and borrowing status, standard panel data methods like fixed-effects and difference-in-difference are biased (Armendárize de Aghion and Morduch, 2005: 210). Other less frequently used panel data techniques such as *random trend*, and *flexible random trend* models offer alternative approaches to mitigate this problem by allowing an arbitrary correlation between time-invariant unobservables as well as individual trends in time-varying unobservables to program participation (Wooldridge, 2002: 317).

This paper uses unique four-round household survey data covering 1997-2006 to estimate the impact of participation in microfinance credit on annual household per capita consumption and housing improvements. The data comes from sixteen villages in northern Ethiopia. We first investigate the impact of credit using fixed-effects approaches that is standardly applied to account for time-invariant individual as well as village unobservables. Further, we use variants of the random trend model due to Heckman and Hotz (1989) that mitigates both time-invariant and individual trends in time-varying unobservables. We find that program credit has significant impact on household consumption and housing improvements of participants compared to non-participants. However, compared to the random trend approach, results from the standard fixed-effects approach that does not account for individual trends in time-varying unobservables overestimates credit impact. We also model program credit more flexibly by including the effect of loan-cycles and individual specific trends and find that credit impact on per capita consumption increases with frequency of borrowing. The effect of borrowing on the probability of housing improvement is realized after one-cycle but declines sooner after the third cycle borrowing. From the flexible approach, we conclude that borrowing effects last longer than one-period and cumulative effects are best captured the longer the time covered in the analysis. Besides, while household borrowing effects are multidimensional and cannot be captured by a single household outcome, we also conclude that effects on household outcomes are not monotonic over time. Impact estimates that do not account for such dynamic effects may therefore underestimate the effect of MFI borrowing.

The rest of the paper is organized as follows. Section 2 provides a brief review of the main approaches followed in the literature on impact assessment. Section 3 describes the nature of the data and section 4 presents the empirical method used. Section 5 provides the estimation results and section 6 concludes.

2 A review of microcredit impact studies

This section presents a brief survey of the main methodological approaches of mitigating selection bias in microfinance impact evaluations.

Measuring the impact of microcredit programs is a challenging task because establishing 'causality' between credit effects and changes in the outcome of interest is complicated by the well known problems of self-selection and program placement biases that are inherent in such programs (e.g., Pitt and Khandker, 1998). Self-selection is a problem because, compared to non-participants, participants may already have initial advantages such as better entrepreneurial ability that can translate into higher outcome variables, even without credit. Using data from a Peruvian MFI, Tedeschi (2008) finds that *"selection into credit programs is a substantial problem: those who will eventually become borrowers have significantly higher incomes than those who will not become borrowers"*. The main challenge is therefore to address the counterfactual question 'how would participants have performed in the absence of program credit or 'how would non-participants have performed had they participated in the program'.

MFIs may also design their credit programs to fit into specific villages or specific groups and screening may be based on criteria that influence outcomes of interest. Self-selection and program placement decisions in principle do not pose problems if they are based on known and measurable variables, because then they can be easily controlled for empirically. The problem is however that these decisions are often based on unobservable variables. In the absence of "comparison" and "treatment" groups, credit impact assessments that do not account for these problems are likely to be biased (Armendáriz de Aghion and Morduch, 2005:200-223; Tedeschi, 2008).

How microfinance impact studies have dealt with these problems varies. One strand of literature that is common among MFI practitioners simply compares existing clients ('treatment group') with new entrants ('control group'). Although simple to implement, this method is criticized for attributing the mean difference between the two as impact without dealing with selection problems (Tedeschi, 2008).

A second strand of literature that relies on cross sectional data deals with the selection problem employing instrumental variable and quasi-experimental techniques that exploit the nature and timing of program designs. One of the earliest and most cited studies in this line is by Pitt and Khandker (1998) who used cross-sectional data from Bangladesh and employed a quasi-experimental survey design to instrument nonrandom program placement and self-selection. However, such instrumental and experimental designs are often coincidental and difficult to replicate. Moreover, these approaches assume that the initial conditions of control and experiment villages are identical. A final problem is often that it is difficult to come up with strong and valid instrumental variables.

An ideal credit impact evaluation would have been one that compares effects with and without the program. A third approach that received considerable attention in recent microfinance evaluation is a pre-designed randomized experimental approach (Karlan and Goldberg, 2007). Experimental designs that randomize over observable and unobservable attributes of participants and non-participants would, in principle, provide unbiased estimates. Such designs are however time consuming and costly to undertake. Besides, it can be difficult to implement on ethical and political grounds (Heckman and Hotz, 1989).

A fourth strand of recent literature uses panel data to mitigate the biases present in cross-sectional studies. Assuming strict exogeneity between selection variables and time-varying unobservables that could affect the outcome of interest, fixed effect panel data methods can provide consistent estimates by differencing out time-invariant unobserved individual and village effects (Wooldridge, 2002: 637). Khandker (2005), Copestake et al. (2005) and Tedeschi (2008) relied on this assumption and used a fixed-effects approach to analyze the impact of credit. The fixed-effects estimator is however critically dependent on this strict exogeneity assumption, particularly on the assumption that the time-invariant heterogeneity is the only potential source of selection bias. Literature in empirical labor economics that studies the effect of labor-training programs on earnings under nonrandom program assignment extends the evaluation literature by allowing for individual heterogeneity to vary over time according to a linear trend (Heckman and Hotz, 1989). This approach is used in this study and is explained in more detail in section 4.

3 Brief description of the MFI, survey design and the data

Data used in this study comes from rural households in northern Ethiopia where a microfinance program, Dedebit Credit and Saving Institution (DECSI), provides financial services only for production purposes. Although DECSI, under the auspices of a local NGO, started providing credit services in few trial villages since 1994, it officially launched credit and saving programs in 1997 and expanded quickly into almost all villages in Tigray. By 2000, it extended loans to 1.4 million rural households with total outstanding loans of 447 million ETB and savings of 74 million ETB[1]. As of 2002, DECSI covered more than 91% of the villages in the region and extended to about half a million borrowers (Borchgrevink, et al., 2003). Initially DECSI provided Grameen style joint liability based credit mostly used for farm inputs, which eventually diversified into micro and small enterprise loans and other off-farm activities. Loans are extended once a year because production is largely monsoon rain dependent, and depending on activity, mature between 6-12 months. In 2003, DECSI started individual loans packaged with some specific farming activities such as bee-keeping and milk production activities. Loan maturity in this latter loan product ranges between 1-2 years. In this study, participation in borrowing is defined as being in a borrowing relationship with DECSI in the year preceding the survey and no attempt is made to make a distinction between the different loan products provided.

[1] ETB stands for Ethiopian currency, ‘Birr’; (annual average) USD conversion rate of 6.32 ETB in 1997 and 8.94 ETB in 2006.

Table 1 Households' participation and changes in borrowing status over survey years

	Number of times borrowed up to the survey year				
Survey year	*Never*	*Once*	*Twice*	*Thrice*	*Always*
1997	140	211	-	-	-
2000	87	182	82	-	-
2003	61	143	112	35	-
2006	40	102	130	46	33

Source: Survey data (1997- 2006)

A four-round survey with three-year intervals (1997-2006) was administered on randomly selected 400 borrower and non-borrower rural households. The dataset covers household- and village-level information ranging from household characteristics, consumption, assets, credit and savings to village infrastructure, markets, and credit contracts. Asked about access to credit in 1997, only a few respondents indicated that they were ineligible to borrow mainly due to old age and physical unfitness, which DECSI implicitly considered as selection criteria[2]. These are excluded from our analysis. Respondent attrition was minimal, mostly related to the Ethio-Eritrea border war, which started in 1998 and ended in 2000. This paper is thus based on a balanced panel of 351 households, out of which 211 borrowed and 140 did not in the 1997 survey. Table 1 gives a summary of the evolution of borrowing status over time. Borrowing status changed in subsequent years with some households joining, while others dropped out. In general, there were 33 households that borrowed in all four periods and 40 that never did. The other households borrowed at least once in one of these years but also had years without a loan.

An advantage in this data set to study impact is that the first survey coincided with the massive expansion of DECSI into most villages in the region, which gives the opportunity to identify impact using the 1997 as baseline information for both borrowers and non-borrowers. Moreover, due to the government's as well as donors' inherent interest to synchronize credit

[2] We also test if there was no significant difference between participant and non-participant groups in the base year in terms of our outcome variables.

services with the regular input extension programs that was running through out the region, there is little reason to believe DECSI's quick and massive branching out to villages has been systematic and endogenous to village outcomes. All residents were, in principle, eligible to branches available in the nearest rural town. E.g., credit was available for all in the most nearest-to-town villages as well as remote villages in 1997. However, households may have self-selected into credit and participation can be endogenous at individual level, which we explicitly tackle in the empirical analysis.

Although credit is given for productive purposes (e.g. fertilizer, oxen), eventually this will lead to higher per capita consumption. Our survey interval of three years is considered as an advantage in this respect, since this higher consumption is expected to materialize in years after having experienced higher output due to increased input use made possible by borrowing. The time lag needed to translate borrowing into outcomes also strengthens the usefulness of the first-round survey as a baseline information to identify impact.

Table 2 Summary statistics of household per capita annual consumption and housing improvements

Survey years	*1997*	*2000*	*2003*	*2006*
Participants	211	135	126	160
Annual per capita consumption				
Mean	442	683	651	1422
Std. Dev.	523	503	371	1051
Housing improvements				
Mean	0.033	0.193	0.429	0.594
Std. Dev.	0.180	0.396	0.497	0.493
Non-participants	140	216	225	191
Annual per capita consumption				
Mean	371	675	577	1087
Std. Dev.	215	543	496	715
Housing improvements				
Mean	0.027	0.042	0.102	0.115
Std. Dev.	0.167	0.200	0.304	0.320

We measure credit impact on two welfare indicators in Tigray, i.e. annual household consumption and housing improvements. Household consumption is a continuous variable and housing improvement is a binary indicator. Households were asked if they had improved their roof to corrugated- sheet of iron anytime between the last and the present survey year. Household consumption is aggregated from food and nonfood consumption of selected items, both from own sources or purchased over a period of one year. Necessary adjustments are made to make measured items and units comparable over the survey years. A consumer price index for the region is used to adjust for price changes over time (Central Statistical Agency of Ethiopia, 2008). To minimize measurement error from age structure heterogeneities among households, per capita adult consumption is used.. Summary statistics of indicators are presented in table 2.

In general, compared to non-participants, an average participant enjoyed higher per capita consumption levels and more often improved her house in all years observed. Note however that average outcomes in table 2 are based on participation or non-participation status in each survey year, i.e., regardless of previous status. We take such contamination effects into account in our econometric modeling and estimation. Moreover, the table doesn't indicate whether higher consumption and housing improvement can be ascribed to borrowing or whether they have increased due to other factors

4 Empirical Methodology

In this section the origins of selection bias in estimating impact from long term panel data and the panel data techniques to control for it are discussed. Consider the following generic specification for program evaluation:

$$C_{it} = X_{it}\beta + prog_{it}\gamma + M_i\alpha + u_{it} \qquad t = 1,2...,T; \quad i = 1,2,...N \tag{1}$$

where the outcome variable consumption, C_{it} for household i at time t, is determined by a vector of observable household-, village-, and MFI-level characteristics X_{it}, a binary program participation variable, $prog_{it}$ (=1, if participated in borrowing at t, zero otherwise), and a vector M_i of time-invariant unobservable variables[3]. Borrowing in turn depends on a set of observable (Z_{it}) and unobservable variables (W_{it}), i.e. $prog_{it} = Z_{it}\psi + W_i\phi + v_{it}$, where Z_{it} can be contained in X_{it}. Selection bias arises when unobservables W_i and residuals v_{it} determining borrowing, correlate with unobservables M_i and residuals u_{it} affecting consumption. Or households that select themselves for borrowing may do so on the basis of unobservable characteristics that may also determine the outcomes consumption and housing improvement (Heckman and Hotz, 1989). This is a testable hypothesis from the first year survey and we follow Tedeschi (2008) to test whether or not the 1997 consumption and housing improvement outcomes for those who eventually become borrowers or those who always borrowed were statistically different from those who never borrowed:

$$C_i = \beta_1 + X_i\beta_2 + \beta_3 Always_i + \beta_4 Dropout_i + \beta_5 New_i + \beta_6 Branch_i + \varepsilon_i \qquad (2)$$

where X is a vector of household characteristics, the dummy variables *Always, Dropout,* and *New* provide the test against those *Never* borrowed, and the dummy *Branch* is one if borrower knew there was a DECSI branch in the nearest town and instruments for bias due to branch assignment by the MFI[4]. If selection is indeed a problem the impact of borrowing on consumption or housing improvement cannot be consistently estimated from (1) by standard pooled OLS estimators. Panel data models with specifications that allow program participation decision to be correlated with unobservables affecting outcome variables provide unbiased impact estimates (Heckman

[3] We follow Wooldridge (2002:247) to use *W* and *M* to denote the unobserved heterogeneity term is a random variable and not a parameter to be estimated and thus ignore ϕ and α in subsequent discussions.
[4] We assume the further away a branch was located from a village in 1997, the less known it would be for villagers.

and Hotz, 1989; Papke, 1994). Three such specifications, i.e. the standard fixed-effects model, the random trend model, and a flexible random trend model are elaborated below and used in our analysis.

The standard fixed-effects estimator provides a consistent estimate of the borrowing parameter, γ, under the assumption that all unobservables that influence the outcome of interest are time-invariant, since these unobservables are removed by a within or first-difference transformation (Wooldridge, 2002: 252). If such individual-specific unobservables change however over time, which may happen for various reasons, the estimate for γ is still biased. In our setting, there are two such potential reasons. First, unobserved negative economic shocks affecting households' input endowments, may pressurize households for input-bridging borrowings or repeat-borrowings to settle earlier debts. Anecdotal evidence from our sample villages indicate that households indeed resort to microfinance borrowings after experiencing a negative shock. Moreover, some repeat-borrowings may follow failure on an earlier one. Second, as argued earlier, credit may have lasting effects on unobservables on which selection is based. E.g. unobserved household characteristics such as entrepreneurial abilities, which may condition credit demand, may change over time depending on previous exposure to microfinance credit. Under these conditions, a more robust specification is required to remedy bias in the parameter estimates of interest.

A more robust specification due to Heckman and Hotz (1989)- the individual-specific trend model- allows both household specific time-invariant unobservables and individual trends of time-varying unobservables to correlate with program participation (Wooldridge, 2002: 315). This model, also used by Papke (1994) to study the effect of nonrandom enterprise zone designation on unemployment and investment, is specified as:

$$C_{it} = X_{it}\beta + prog_{it}\gamma + M_i\alpha + g_i t + u_{it} \tag{3}$$

where g_i is an individual trend parameter, which in addition to the level effect M_i, captures individual-specific growth rates over time. A consistent estimate for γ, viz. the treatment effect of borrowing, can be obtained by wiping out the time-varying unobservables and the trend in time-invariant unobservables that can potentially bias γ (Wooldridge, 2002: 315). First, eq. (3) is first-differenced to eliminate M_i, which gives a standard fixed-effects model:

$$\tilde{C}_{it} = \tilde{X}_{it}\beta + pro\tilde{g}_{it}\gamma + \tilde{g}_i + \tilde{u}_{it} \qquad t=1,2,\ldots,T \tag{4}$$

where $\tilde{C}_{it} = C_{it} - C_{it-1}$, $\tilde{X}_{it} = X_{it} - X_{it-1}$, $\tilde{u}_{it} = u_{it} - u_{it-1}$ and $\tilde{g}_i = g_i t - g_i(t-1)$. Second, eq. (4) is consistently estimated using a standard fixed-effects approach, i.e. using a within transformation or by differencing the equation (again) to eliminate g_i and then estimate by OLS. The latter is preferred if u_{it} after the first differencing cannot be assumed white noise but at the cost of losing one period information in each transformation (Wooldridge, 2002: 316). Note that γ can be estimated consistently from this specification only if $T > 3$. In short panels like ours, it may be reasonable to assume u_{it} to be serially uncorrelated after first-differencing. However, using a second differencing transformation has an extra advantage of not assuming homoskedasticity of the first-difference of u_{it} (Wooldridge, 2002:316). We therefore second-difference eq. (4) and estimate by pooled OLS.

Although we only have four rounds of panel data, still our data covers a period of ten years. An advantage of panel data covering a longer period is that it enables to estimate the impact from long-term rather than one-shot program participation. Repeated participation may, in addition to shifting the levels in each borrowing year, affect the rate of change of the outcome variables relative to nonparticipation. Following Papke (1994) and Friedberg (1998), we account for this by including $prog_{it} \cdot t$ in eq. (4):

$$C_{it} = X_{it}\beta + \gamma_1 prog_{it} + \gamma_2 prog_{it} \cdot t + M_i\alpha + g_i t + u_{it} \quad (5)$$

This specification provides impact estimates robust to random periodical changes by allowing the individual-specific trend to vary on participation over time. Estimation follows the same procedures as in eq. (4).

The specifications in (3) and (5) however impose the restriction that each successive loan-cycle's borrowings have uniform effects as their preceding borrowing. Initial borrowings may however entail lasting effects on incentives as well as on consumption levels, which alter the scale of the effects of borrowings later. A more flexible specification suggested by Wooldridge (2002: 317) allows program indicators to reflect the frequency of participation in each possible participation year as presented in table 1. This is done by replacing $prog_{it}$ and $prog_{it}.t$ in eq. (5) with a series of program indicators for each loan-cycle the participant has been in the program:

$$C_{it} = X_{it}\beta + \gamma_1 prog1_{it} +,...,+ \gamma_k progk_{it} + g_i t + M_i\alpha + u_{it} \quad (6)$$

where $progj_{it}$ =1 if household i has been in the program for exactly j years in year t and zero otherwise; k is the maximum number of (observed) years a household can be in the program. Program indicators attach more weights to differences between households' degree of participation regardless of year of participation. More weights are also given to the timing of participation within each indicator[5]. As before, eq. (6) is first-differenced and then transformed again by a within or another difference procedure.

Finally, note that since one of our outcome variables, i.e. housing improvement, is a binary indicator, the model is basically a limited dependent with binary regressor of the type discussed in Angrist (2001). Binary choice models with panel data are problematic to estimate due to the incidental parameter problem. Angrist (2001) emphasizes rather than imposing distributional

[5] E.g., for household i and j that borrowed twice each, but i borrowed in the first two years and j borrowed in the last two years, the model attaches the same weights for both i and j (i.e., prog1 =1 and prog2 = 1, for i and j, i ≠ j). However, in the within observations, prog2 gives more weight to i (i will have more ones in prog2) than to j.

assumptions which may complicate estimation and yield inconsistent estimates, a simpler estimator such as the linear probability model (LPM) is attractive and consistent for answering the question of interest, mainly estimating the effect of binary regressor in models with limited dependent variables. Thus, we stick to the simple LPM specification, which also provides an estimate conveniently interpreted as effect on the mean of the dependent (Wooldridge, 2002: 454-457).

5 Estimation results

In this section estimation results from the models outlined in section 4 are provided. Selection bias test results are first presented. The test is carried out by estimating eq. (2) using OLS for the 1997 consumption expenditure outcome and using a logit model for the 1997 binary housing improvement outcome. The null hypothesis that all parameters of interest are simultaneously equal to zero is rejected by the *F-statistic* test at 1 per cent significance level for both the OLS and logit models, indicating that both fit the data set well. Results are given in table A.1 (appendix).

The most important test results are given by the parameter estimates for *Branch, New* and *Always.* First, in both models, the insignificance of the proxy for DECSI branch in 1997 suggests that there is no bias due to program placement. Second, the hypothesis that *New* is different from zero is also rejected at acceptable significance level in both models, but the same hypothesis cannot be rejected for *Always* at 10 per cent significance level in the consumption expenditure model. Thus, controlling for dropouts, we cannot confirm those who will eventually become borrowers in 2000 had higher consumption levels than those who never did. However, we find evidence that those who always borrowed had consumption levels higher than those who never borrowed. Thus, our analysis here after must account for potential bias due to self selection but not due to program placement.

The basic model given in eq. (1) is estimated by the standard fixed-effects estimator where instead of a binary participation variable, the number of years the household has been in a borrowing relationship is used to account for the *degree* of participation as suggested by Copestake et al., (2001). Since we are primarily interested in credit impact estimates, only household observables that may systematically correlate with selection even after controlling for effects of time-invariant unobservables are included. One implicit borrower screening criteria of DECSI is household head age. Besides, as household heads become older, they self-select out of borrowing activities. Since most household variables collected are time-invariant we included only time-varying variables that may be systematically correlated to participation, mainly, land size and its square, gender of household head, household head's age and its square as other explanatory variables related to selection into the program. Although land is state owned in Ethiopia, farmers are given user rights. 'Ownership' of land and size cultivated therefore determines amount of input use, including credit. A year dummy (equal to 1 for 2006, zero otherwise) is included to contrast the relatively stable and good harvest year 2006 to the earlier years that are characterized by adverse conditions such as war and drought. That 2006 was a very good year is also reflected in table 2 that shows that average deflated consumption in that year was much higher. Note that household head's gender and skills are time variant. This specification is similar to Tedeschi's (2008) fixed-effect model except that our specification considers the cumulative effect of several loan-cycles as compared to 'number of participation days' used in the former paper. Results are reported in table 3. The *F*-statistics (at 1 per cent significance level) indicate that for both household consumption and housing improvement models all parameters are not all jointly equals to zero.

Based on the fixed-effects estimation, credit has a significant positive effect on annual household consumption expenditure and housing improvements of borrowers compared to non-borrowers. After controlling for potential selection on unobservable fixed-effects, household per capita consumption for an average borrower household has increased by ETB 415 for each additional borrowing year. Moreover, the probability of improving the house increases on average by 0.27 per year of credit taken. Note that the parameter for the 2006 dummy is also statistically highly significant in the consumption equation, indicating differential impacts on participants and non-participants due to aggregate macroeconomic variability, which also includes specific events which may have occurred due to aggregate effects (e.g., death of livestock due to drought and death of key labor in the household due to war). Compared to non-participants, participants have seen ETB 264 more consumption in the good year 2006.

Table 3 Household fixed-effects estimates of the impact of credit

Dependent variables	*Per capita annual household consumption*	*Housing improvements*
Number of (observed) years borrowed	**414.665*****	0.273***
	(27.584)	(0.015)
Women headed household	61.058	-0.038
	(51.853)	(0.028)
Additional skills other than farming	62.136	0.039
	(60.823)	(0.033)
Year 2006 dummy	264.098***	-0.012
	(38.227)	(0.021)
Age of household head	10.216	0.004
	(9.597)	(0.005)
Age-squared	-0.059	-0.628×10^{-4}
	(0.090)	(0.491×10^{-4})
Cultivated land size (in *Tsimad = 0.25hectare)*	-11.735	-0.002
	(9.378)	(0.005)
Land size-squared	0.066	-0.139×10^{-3}
	(0.295)	(0.162×10^{-3})
Intercept	-289.897	-0.168
	(246.768)	(0.135)
Within R-squared	0.215	0.257
F (8, 1045)	35.77***	45.250***
Household fixed-effects	Jointly significant***	Jointly insignificant

Number of observations	1404	1404

*, ** ,*** significant at 10%, 5% and 1%, respectively; Standard errors in parentheses

The individual household heterogeneity not picked up by the variables included is captured in the fixed-effects parameter. For the household consumption model there is evidence for household heterogeneity given the significance of the fixed-effects. This is not however the case for the housing improvement model, suggesting for a pooled estimation. Estimating it by pooled regression also provides qualitatively the same results as the fixed-effects results. Note that the FE within procedure also has the benefit of removing potential selection bias due to time-invariant unobservables. As indicated in section 4, error terms may correlate due to selection based on time- varying individual-specific unobservables. In that case the individual trend model as specified in eq. (3) is more robust than the standard FE model since it allows selection to be based not only on individual averages of unobservables (i.e., fixed-effects) but also on individual-specific unobservable *trends*. This model is estimated by OLS after differencing twice to eliminate the trend component. This is done for both consumption and housing improvement outcomes. Since results for the housing improvement model are very similar to the fixed-effects results presented in table 3 they are not reported here. Results for household consumption are reported in the first column of table 4.

In general, removing individual-specific unobserved dynamics by including an individual trend and differencing the data twice provides more conservative results. Specifically, according to this individual-specific trend specification, per capita annual consumption increases by ETB 199 per year of credit taken. This result is statistically significant and credit impact is substantially reduced (by a more than 50 percent a year) compared to the fixed-effects result. This difference is the bias in the standard fixed-effects result due to time-varying individual dynamics.

Table 4 Household specific trend model results of credit impact on per capita annual consumption

Variables	*Individual trend model*	*Individual trend model, and trend based on participation*
Number of (observed) years in borrowing	**199.317****	**160.738****
	(77.065)	(79.016)
Random trend *borrowing participation	-	**33.858****
		(16.043)
Year 2006 dummy	323.439***	324.497***
	(32.594)	(32.517)
Age of household head	2.003	1.632
	(9.428)	(9.407)
Age-squared	-0.022	-0.017
	(0.089)	(0.089)
Cultivated land size (in *Tsimad = 0.25 hectares)*	-0.496	-1.739
	(13.249)	(13.229)
Land size-squared	0.139	0.193
	(0.463)	(0.462)
Intercept	-130.553	-113.738
	(88.088)	(88.230)
R-squared	0.164	0.169
F (6, 695); F (7, 694)	22.640***	20.14***
Number of obs.	702	702

*, ** ,*** significant at 10%, 5% and 1%, respectively; standard errors in parentheses

Consistent results are obtained when the same specification is estimated by fixed-effects after first-difference. After the first-difference, the fixed-effects error component is however jointly insignificant favoring estimation by pooled OLS. Second-differencing eliminates the trend and provide results more robust to second-order serial correlation and heteroscedasticity.

A variant of the individual-specific trend model given in eq. (5) allows individual household consumption not just to vary at different *trends* but also allows borrowing effects to depend on these unobserved individual-specific *trends.* Note that in this case, *trend* is interacted with participation ($prog_{it}$) indicator and not with 'number of years in borrowing'. Results are reported in column 2 of table 4. The credit effect estimate is both quantitatively as well as qualitatively consistent to the results in column 1 in table 4, but again more conservative than the standard fixed-effects estimate. After controlling for both time-invariant and time-varying selection bias, each borrowing cycle increases per capita consumption by ETB 161 directly and by ETB 34 indirectly (by changing other unobserved time-varying individual characteristics). Thus, after accounting for selection biases, credit has been responsible not just to change the *levels* but also the *rate* at which yearly per capita consumption grew for an average borrowing household in the ten years considered. Note that other results are also consistent across the two specifications presented in table 4. A consistently significant negative intercept in both specifications captures a general consumption decline trends not captured by our aggregate shock variable.

The results in table 4 provide interesting insights into how effective microfinance can be for households trying to extricate themselves from poverty in those villages, other factors remaining the same, by keeping their relationship with the MFI.

Table 5 Result of flexible random trend model with participation indicators

Dependent variables	*Household per capita annual consumption*		*Housing improvements*	
One year borrowing	**273.936****	(107.526)	-0.004	(0.075)
Two years borrowing	**319.132****	(137.706)	0.244**	(0.097)
Three years borrowing	310.697	(213.204)	0.555***	(0.149)
Four years borrowing	**665.024****	(337.707)	0.457*	(0.237)
Year 2006 dummy	326.079***	(31.954)	-0.019	(0.022)
Age of household head	2.578	(9.432)	-0.007	(0.007)
Age-squared	-0.027	(0.089)	0.531×10^{-4}	(0.623×10^{-4})
Cultivated land size (in *Tsimad = 0.25hectare*)	-0.887	(13.250)	-0.004	(0.009)
Land size-squared	0.175	(0.463)	-0.159×10^{-3}	(0.3245×10^{-3})
Intercept	16.268	(70.153)	-0.017	(0.049)
R-squared	0.170		0.044	
F(9, 692)	15.76***		3.560***	
Number of obs.	702		702	

*, ** ,*** significant at 10%, 5% and 1%, respectively; standard errors in parentheses

Important follow-up questions from a policy point of view are whether impact can be associated to the extent of repeat-borrowing. This is analyzed using the flexible individual-specific trend model given in eq. (6), which assigns indicators for the number of times each household has been involved in borrowing. Results are given in table 5.

Again, the double differencing estimation procedure reduces the potential of selection bias to a minimum. Results show once more that borrowing has a significant impact on (future) consumption, but interestingly enough the magnitude of impact substantially increases with the increase in the length of relationship with the MFI. Specifically, compared to non-participants and other participants, per capita consumption has significantly increased by ETB 274 for one year participants and by ETB 319 for two year participants (which for them adds to the first year effect of ETB 274). The effect is (slightly) statistically insignificant for three year participants (p-value 0.145). Since for most three year participation the third cycle coincided with the occurrence of one of worst droughts the country has seen, the effect seems to have been neutralized by the overall drastic consumption shortfall for participants and non-participants alike. However, having participated in the previous three years has had an inertia effect for four year participation such that per capita consumption has increased substantially (by ETB 665) only for households that participated in all four cycles. It means that while the cumulative effect of the pre-shock participation might have slightly helped to overcome consumption shortfalls for participants compared to non-participants, those that participated for four years, including after the shock, have benefited substantially in terms of per capita consumption increases. For the housing improvement model, the probability of improving the house has significantly increased after the second round borrowing and raises up to 0.244 if households borrowed for two periods, 0.555 if borrowed for three periods and 0.457 if borrowed in all periods. The relatively lower effect in the case of borrowing in 'all periods' would not be surprising as households eventually shift attention from improving their houses to other activities.

Compared to the average impact on the participant obtained from the individual-specific trend model, this finding supports the lasting impact of credit over time by uncovering the specific impacts on each cohort of participants. Thus, while the impact of one time borrowing is close to the average impact previously obtained, it also uncovers having borrowed three and four times leads to even higher increases in consumption and probability of house improvements. Such high percentage increases attributed to credit is not surprising given the importance of credit at such marginally low initial conditions (e.g. initial average per capita consumption is ETB 442 for participants and ETB 370 for non-participants) and the relatively long period covered in which 8-11 per cent GDP growth was registered in the country.

6 Conclusions

Impact evaluations are often prone to self-selection and program placement biases. This paper uses panel data techniques to deal with these potential selection biases. Standard fixed effect models mitigate selection based on time-invariant unobservables, whereas the more advanced random trend model also account for individual trends in time-varying unobservables. The dataset used is a unique four-round panel data set among households in Tigray, Ethiopia that covers a period of ten years, so that lasting effects of credit can be established.

The analysis started with tests of program placement and self-selection biases. While there was no indication of bias due to systematic program placement, the data did not confirm absence of bias due to self-selection. The analysis therefore accounts for any potential selection bias. Results indicate that microfinance credit significantly raised annual per capita household consumption. It also significantly raised the probability of improving housing (roofs), which is an important welfare indicator in this area. The random trend model with flexible participation indicators, which considers frequency of participation, shows that per capita household consumption (except in the bad year 2003) and probability of improving the house substantially increased with the frequency of participation. One time borrowing has no impact on housing improvements but significant improvements in per capita consumption, which is plausible at such early stages of livelihood changes for households in those marginal areas. Repeat-borrowing did matter in both cases however, but with a slight decline of the probability of housing improvements for household that borrowed frequently.

These findings have both substantive as well as methodological significance. First, they reflect the effect of credit on livelihoods is multi dimensional and cannot be fully captured by just a single household outcome. Moreover, the effect is not monotonically the same over time on all livelihood indicators used to measure impact. Second, it is also imperative that the effect of borrowing lasts longer than one or two periods. It therefore takes time before the effect of borrowing on livelihoods is fully materialized. Methodologically, impact estimates that rely on a single household indicator and only one-cycle of borrowing may undermine the potentials of microfinance credit on overall livelihoods that could be achieved over time. Future research must focus on more robust specifications that incorporate temporal as well as multidimensional effects of credit on livelihoods.

The implication for MFI practitioners such as DECSI is that eligible households should not only be encouraged to borrow, but also, if successful, to stay longer in a borrowing relationship in order to realize the full potentials of borrowing. As such, early *graduation* from microfinance (in our case, as early as before ten years) might be pre-mature in terms of achieving the required goal of eradicating poverty and careful weighing is necessary before graduation takes place. The flexible specification results also suggest that those that were able to continue borrowing even after a major shock in 2003 have seen even higher consumption levels after that shock. This implies that rescheduling repayment so as to provide, rather than deny, access to future borrowing after a shock may help poor borrowers to bridge their consumption and regain economic normalcy after a shock. Finally, although the results of the fixed-effect and trend models deviate somewhat, due to different assumptions, specifications and estimation techniques, they all strongly suggest that microfinance in this part of Africa has been useful in terms of measured outcomes.

Appendix

Table A.1 Test results for selection bias using base year data

Variables	*Per capita Consumption expenditure*♣		*Housing improvements*†	
Intercept	312.295	(499.573)	-9.894*	(4.568)
Household characteristics				
Age of household head	55.171***	(19.222)	0.227	(0.161)
Age-squared	-0.551***	(0.179)	-0.002	(0.002)
Women headed (yes=1)	-707.499***	(113.386)	1.934*	(0.995)
Special skills other than farming (yes=1)	388.856	(281.885)	1.325	(1.102)
Household head's education (literate=1)	411.233	(268.719)	-0.020	(0.994)
Number of oxen owned	53.220	(56.508)	0.521	(0.460)
Per capita land size owned	431.512*	(212.639)	-3.851	(2.330)
Shock occurred (yes =1)	-206.042*	(100.323)	0.378	(0.755)
Village characteristics				
Micro dam available (yes=1)	229.822*	(125.988)	0.163	(0.652)
Village is remote (yes=1)	-237.003*	(103.117)	-0.270	(0.837)
Borrowing status				
Always	249.392*	(142.423)	1.505	(1.107)
Dropout (in 2000)	191.481	(126.477)	0.024	(1.089)
New (in 2000)	-91.490	(124.259)	0.859	(1.132)
Knew **branch** was available in nearest town (yes=1)	77.345	(110.607)	0.359	(0.720)
R-squared; Pseudo R-squared	1912		0.150	
$F(14, 336)$; Wald χ^2 (14)	7.80***		70.370***	
Sample size	351		351	

♣ OLS estimates; †Logit estimates; *, ** ,*** significant at 10%, 5% and 1%, respectively; *Robust*-std. errors in parentheses