Does Microfinance Reduce Rural Poverty? Evidence Based on Household Panel Data from Northern Ethiopia

Guush Berhane and Cornelis Gardebroek

Evidence on the long-term impacts of microfinance credit is scarce. We use a unique four-round panel dataset on farm households in northern Ethiopia that had access to microfinance, observed on two key poverty indicators: household consumption and housing improvements. Fixed-effects and random trend models are used to reduce potential selection biases due to time-invariant unobserved heterogeneity and individual trends therein. Results show that borrowing indeed causally increased consumption and housing improvements. A flexible specification that takes into account repeated borrowings also suggests that borrowing has cumulative long-term effects on these outcomes, implying that short-term impact estimates may underestimate credit effects.

Key words: Ethiopia, household panel data, long-term impacts, microfinance, trend model.

JEL codes: O16, G21, I32, C23.

The microfinance revolution has gained considerable momentum around the world in the last twenty-five years. The potential of microfinance as an effective tool to break through 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 in both rural and urban areas. However, important questions that need to be asked are whether and to what extent microfinance credit has contributed to reducing poverty.

Despite efforts to measure this impact, evidence on the poverty-reducing effects of long-term microfinance credit remains limited. This is due mainly to the difficulty of measuring counterfactual outcomes and the lack of follow-up data spanning sufficiently long periods to measure credit impact. Without experimental designs, evaluations based on simple comparisons between participants and nonparticipants are subject to biases from two sources (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 villages based on 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. Potential applicants themselves can choose 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.

However, even if experimental or quasi-experimental designs that randomize over potential sources of selection are implemented, impact estimates based on one period of data may fall short of capturing the complete impact of credit because it may take longer before the full effects of credit are realized (Karlan and Goldberg 2007). A recent review of the evaluation literature emphasizes that the issue of "timing and duration of exposure to programs" is just as important as, but relatively less studied than, the identification problems that often attract much of researchers' attention (King and Behrman 2009). Long period data

Guush Berhane is a postdoctoral fellow at the International Food Policy Research Institute, Ethiopia Strategy Support Program, Addis Ababa, Ethiopia. Cornelis Gardebroek is an assistant professor in the Agricultural Economics and Rural Policy Group, Wageningen University, the Netherlands. The project was supported by the Netherlands Organization for Scientific Research (NWO), grant number WB 46-450. The authors thank participants at various seminars, the editor, and two anonymous reviewers for their insightful comments.

Amer. J. Agr. Econ. 93(1): 43–55; doi: 10.1093/ajae/aaq126
Received September 2009; accepted October 2010

are, however, costly and largely unavailable. As a result, most studies so far (e.g., Coleman 1999; Pitt and Khandker 1998) have 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 studies by Khandker (2005), Copestake, Bhalotra, and Johnson (2005), and Tedeschi (2008), who used two-period data to estimate impacts.

Long-term panel data allow us to measure impact from the intensity of participation over time. 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 process is based on time-varying unobservables, such as individual motivation that is likely to change over time, standard panel data methods like fixed effects and difference-in-differences are biased (Armendáriz de Aghion and Morduch 2005, p. 210). Other, less frequently used panel data techniques such as (flexible) random trend models offer alternative approaches to reduce this problem by allowing arbitrary correlation between time-invariant unobservables as well as individual trends in time-varying unobservables and program participation (Wooldridge 2002, p. 317). These approaches are used in this study and explained in more detail in a later section.

This article uses unique four-round household panel data covering the period 1997–2006 to estimate the impact of participation in microfinance credit on annual household per capita consumption and housing improvements. The data come from sixteen villages in northern Ethiopia. We first investigate the impact of credit using a standard fixed-effects approach that accounts for time-invariant individual and village unobservables. Further, we use variants of the random trend model by Heckman and Hotz (1989) that considers both time-invariant unobservables and individual trends in time-varying unobservables. We find that program credit has significant impact on household consumption and housing improvements. Compared with the random trend approach, 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 after the third borrowing cycle. From this flexible approach, we conclude that borrowing effects last longer than one period. Additionally, 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 long-run effects may therefore underestimate the effect of MFI borrowing.

The rest of the article is organized as follows. First, we present a brief literature review of the main approaches used in impact assessment of microcredit. Next, we describe the data used in the analysis. Empirical methods and estimation results are provided in subsequent sections. We present conclusions and implications in the final section.

A Review of Microcredit Impact Studies

This section presents a brief survey of the main methodological approaches used in 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 (see, e.g., Pitt and Khandker 1998). Self-selection is a problem because, compared with nonparticipants, 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 alternatively: How would nonparticipants have performed had they participated in the program?

MFIs may also design their credit programs to fit 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, pp. 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"). A study of this nature by MkNelly and Lippold (1998) finds no effect on enterprise profits in Mali. Another similar study by Edgcomb and Garber (1998), however, finds higher enterprise profit impacts for MFI clients. Although simple to implement, this method is criticized for attributing the mean difference between the two as impact without dealing with selection problems. For a thorough discussion on various problems that may arise using this method, see Karlan (2001) or Tedeschi and Karlan (2010). 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. They find, among other things, that credit has a larger impact on consumption expenditures for female borrowers than it has for male borrowers. However, these 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 is one that compares effects with and without the program. A third approach that has received considerable attention in recent microfinance evaluation is a predesigned randomized experimental approach (Karlan and Goldberg 2007). Experimental designs that randomize over observable and unobservable attributes of participants and nonparticipants in principle would provide unbiased estimates. Such designs are, however, time-consuming and costly to undertake. Besides, they 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-effects panel data methods can provide consistent estimates by differencing out time-invariant unobserved individual and village effects (Wooldridge 2002, p. 637). Khandker (2005), Copestake, Bhalotra, and Johnson (2005), and Tedeschi (2008) relied on this assumption and used difference-in-difference and fixed-effects approaches to analyze the impact of credit.

The findings vary across these studies. Khandker (2005), building on the cross-sectional data used in Pitt and Khandker (1998), finds positive impacts on female borrowers. Copestake, Bhalotra, and Johnson (2005) use a difference-in-difference approach and find positive impacts on individual incomes but not on microenterprise sales and profits of program participants in Peru. Tedeschi (2008) finds positive impacts of loans on microenterprise profits.

While more rigorous than the cross-sectional studies, 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 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 a later section.

Brief Description of the MFI, Survey Design, and Data

Data used in this study come from rural households in the northern Ethiopian region of Tigray, where an MFI, Dedebit Credit and Saving Institution (DECSI), provides financial services for production purposes.[1] Although DECSI, under the auspices of a local nongovernmental organization, has been providing credit services in a 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 was providing loans to 210,000 borrowers with 1.4 million credit transactions amounting to 447 million Ethiopian birrs (ETB) total outstanding loans and ETB74 million total savings.[2] As of 2002, its network of 9 branches and 96 subbranches with headquarters in the capital city of the regional state covered more than 91% of the villages in the region and extended loans to about half a million borrowers (Borchgrevink, Valle, and Woldehanna 2003).

Initially, DECSI provided loans using Grameen-style joint liability–based credit mostly for farm inputs. It went on to diversify into micro- and small enterprise loans and other off-farm activities (Berhane, Gardebroek, and Moll 2009). Loans are extended once a year because production is dependent largely on monsoon rain, and depending on activity, loans mature between six and twelve months. In 2003, DECSI started individual loans packaged for specific farming activities such as beekeeping and milk production. Loan maturity of this latter loan product is one to two years. In this study, participation in borrowing is defined as having a borrowing relationship with DECSI in the years preceding the survey. No distinction is made among the different loan products.

A four-round survey with three-year intervals (1997–2006) was administered on 400 randomly selected rural borrowers and nonborrowers.[3] 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, due mainly to high age and physical incapacity, which DECSI implicitly considered as selection criteria. These respondents have been excluded from the analysis. Respondent attrition was minimal and related mostly to the Ethiopia-Eritrea border war, which started in 1998 and ended in 2000. This analysis is thus based on a balanced panel of 351 households, of which 211 borrowed and 140 did not borrow 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 and others dropping out. In general, there were 33 households that borrowed in all four periods and 40 that did not borrow at all. The other households borrowed at least once in one of these years but also had years without a loan.

An advantage of these data in studying credit impact is that the first year of observation coincides with the massive expansion of DECSI into most villages in the region. This enables us to identify impact using 1997 as baseline information for both borrowers and nonborrowers. Moreover, due to the government's as well as donors' interest in synchronizing credit services with the regular input extension programs that were running throughout the region, there is little reason to believe that DECSI's quick and massive branching out to villages has been systematic and endogenous to village outcomes. In principle, all residents were eligible at branches available in the nearest rural town. Credit was available in villages close to towns as well as in more remote villages in 1997. Note that each *wereda,* or district, has a central town, but that additional village towns are possible within a *wereda.* This means that the "nearest town" can be different from the *wereda* town. In either case, the "nearest town"

Table 1. Households' Participation and Changes in Borrowing Status over Survey Years

Survey Year	Number of Times Borrowed				
	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).

[1] Tigray is the northern regional state of Ethiopia. It covers a total area of 80,000 square kilometers and has 4.3 million inhabitants, 80.5% of whom (750,000 farm families) reside in rural areas with livelihoods depending on smallholder agriculture.

[2] In 1997, \$1 = ETB6.32; in 2006 \$1 = ETB8.94.

[3] This is a subsample of a larger baseline survey that covered 100 villages in Tigray. This baseline survey was designed by a collaborative research team from the International Food Policy Research Institute, Mekelle University, Ethiopia, and the International Livestock Research Institute and was funded by the Norwegian Research Council, Norway (Hagos 2003). In the second round survey, 400 households were selected from 16 villages (out of the 100 villages), which in turn were selected from 11 districts (*weredas*) in four main zones of the region (see fig. 1 for the locations). Twenty-five households were selected from each village, randomized at village and household levels. None of our sample villages was included in the 1994 pilot project.

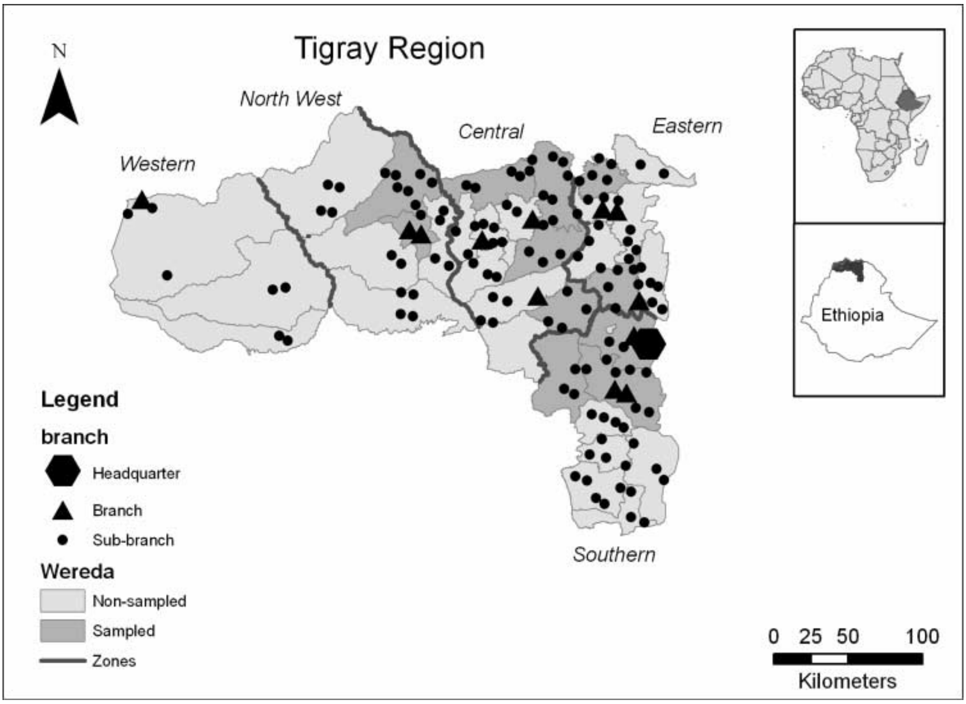

Figure 1. The distribution of DECSI's branches (triangles) and subbranches in Tigray regional state, Ethiopia, as of 2000

is unique for each village, as sample villages are distributed across the region, as indicated in figure 1. However, households may have self-selected into credit, and participation can be endogenous at the 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 baseline information to identify impact. 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 or had started to improve their houses to corrugated-sheet or iron roof anytime between the previous and the present survey year. Note that housing improvement may occur incrementally, resulting in subsequent outcomes after multiple loans. Household consumption is an aggregate of selected food and nonfood items, both from own sources and from purchases over a period of one year.[4]

Necessary adjustments are made to make measured items and units comparable across 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 heterogeneity in age among household members, per capita adult consumption is used. Summary statistics of indicators are presented in table 2.

In general, compared with nonparticipants, average participants enjoyed higher per capita consumption levels and more often improved their houses in all the years observed. Note, however, that average outcomes in table 2 are based on participation or nonparticipation status in each survey year, i.e., regardless

[4] Food items include food grains, fruits, vegetables, milk and milk products, beef, meat and meat products, cooking oil, salt, and coffee, tea, and other leisure drinks. Nonfood items include clothing and footwear, gas and fuel, schooling, health, family events, and household durables. Note that the recall period for estimating small consumption items such as fruits and vegetables was a month, and for other important items such as food grains, it was a year.

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 (constant ETB)				
Mean	442	683	651	1422
SD	523	503	371	1051
Housing improvements				
Mean	0.033	0.193	0.429	0.594
SD	0.180	0.396	0.497	0.493
Nonparticipants	140	216	225	191
Annual per capita consumption (constant ETB)				
Mean	371	675	577	1087
SD	215	543	496	715
Housing improvements				
Mean	0.027	0.042	0.102	0.115
SD	0.167	0.200	0.304	0.320

of previous status. We take such contamination effects into account in our econometric modeling and estimation. Moreover, from the table, one cannot infer whether higher consumption and housing improvement are due to borrowing or to other factors.

Empirical Method

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:

$$(1) \qquad C_{it} = \boldsymbol{X}_{it}\boldsymbol{\beta} + prog_{it}\gamma + \boldsymbol{M}_i\boldsymbol{\alpha} + u_{it}$$
$$t = 1, 2, \ldots, T; \quad i = 1, 2, \ldots, N$$

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 $\boldsymbol{X}_{it}$, a program participation variable, $prog_{it}$, and a vector $\boldsymbol{M}_i$ of time-invariant unobservable variables. In many studies, the program participation variable is defined as a dummy variable. However, given the nature of our data, we define $prog_{it}$ as the number of years the household has been in a borrowing relationship in order to account for the *degree* or *intensity* of participation, as suggested by Copestake, Bhalotra, and Johnson (2001).

Borrowing in turn depends on a set of observable ($\boldsymbol{Z}_{it}$) and unobservable variables ($\boldsymbol{W}_{it}$), i.e., $prog_{it} = \boldsymbol{Z}_{it}\boldsymbol{\psi} + \boldsymbol{W}_i\boldsymbol{\phi} + v_{it}$, where $\boldsymbol{Z}_{it}$ can be contained in $\boldsymbol{X}_{it}$. Selection bias arises if unobservables $\boldsymbol{W}_i$ and residuals v_{it} correlate with unobservables $\boldsymbol{M}_i$ and residuals u_{it} in the consumption outcome equation.[5] Households that select themselves for borrowing may do so on the basis of unobservable characteristics that also determine the consumption and housing improvement outcomes (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 drop out in 2000 or those who always borrowed were statistically different from those who never borrowed:

$$(2) \qquad C_{i,97} = \beta_1 + \boldsymbol{X}_{i,97}\boldsymbol{\beta}_2 + \beta_3 Always_{i,00} + \beta_4 Dropout_{i,00} + \beta_5 New_{i,00} + \beta_6 Branch_{i,97} + \varepsilon_i$$

where $\boldsymbol{X}$ is a vector of household characteristics. The dummy variables *Always, Dropout*, and *New* enable testing for consumption differences between borrowers and those who *Never* borrowed. The dummy variable *Branch* is 1 if a borrower knew there was a DECSI branch in the nearest town and it captures bias due to branch assignment by the MFI. If selection is indeed a problem, the impact of borrowing on consumption or housing improvement cannot be consistently estimated from equation (1) by standard pooled ordinary least squares (OLS) estimators.

Panel data models that allow program participation decisions to be correlated with unobservables affecting outcome variables reduce this problem (Heckman 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, will be elaborated on 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

[5] We follow Wooldridge (2002, p. 247) in using W and M to denote that the unobserved heterogeneity term is a random variable and not a parameter to be estimated. Therefore, we ignore ϕ and α in subsequent discussions.

interest are time invariant, since these unobservables are removed by a within or first-difference transformation (Wooldridge 2002, p. 252). However, if such individual-specific unobservables change over time—which may happen for various reasons—the estimate for γ is still biased. In our setting, there are two potential reasons for such effects. First, unobserved negative economic shocks affecting households' input endowments may pressurize households into input-bridging borrowings or repeat-borrowings to settle earlier debts. Anecdotal evidence from our sample villages indicates 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. For example, unobserved household characteristics such as entrepreneurial abilities, which may condition credit demand, may change over time depending on previous exposure to microfinance credit.

A more robust specification according to Heckman and Hotz (1989), the individual-specific linear trend model, allows both household-specific time-invariant unobservables and individual trends of time-varying unobservables to correlate linearly with program participation (Wooldridge 2002, p. 315). In other words, this specification remedies bias from time-invariant factors and linear trends in time-varying factors, but not from any remaining nonlinear factors. This model, also used by Papke (1994) to study the effect of nonrandom enterprise zone designation on unemployment and investment, is specified as:

$$(3) \qquad C_{it} = \boldsymbol{X}_{it}\boldsymbol{\beta} + prog_{it}\gamma + \boldsymbol{M}_i\boldsymbol{\alpha} + g_i t + u_{it}$$

where g_i is an individual trend parameter, which, in addition to the level effects $\boldsymbol{M}_i$, captures individual-specific growth rates over time. A consistent estimate for γ, viz., the treatment effect of an additional year of borrowing, can be obtained by eliminating the linear trend in time-varying unobservables as well as time-invariant unobservables that can potentially bias γ (Wooldridge 2002, p. 315). First, equation (3) is first-differenced to eliminate $\boldsymbol{M}_i$, which gives a standard fixed-effects model:

$$(4) \qquad C^*_{it} = \boldsymbol{X}^*_{it}\boldsymbol{\beta} + prog^*_{it}\gamma + g^*_i + u^*_{it} \qquad t = 1, 2, \ldots, T$$

where $C^*_{it} = C_{it} - C_{it-1}$, $\boldsymbol{X}^*_{it} = \boldsymbol{X}_{it} - \boldsymbol{X}_{it-1}$, $u^*_{it} = u_{it} - u_{it-1}$ and $g^*_i = g_i t - g_i(t-1)$. Second, equation (4) is consistently estimated using a standard fixed-effects approach.[6] We then second-difference equation (4) and estimate by pooled OLS. Note that γ can be estimated consistently from this specification only if $T > 3$.

Although we have only four rounds, our panel data cover ten years. An advantage of panel data covering a long period is that they enable us to estimate the impact from long-term rather than one-shot program participation. In addition to shifting the levels in each borrowing year, repeated participation may affect the rate of change of the outcome variables relative to nonparticipation. Following Papke (1994) and Friedberg (1998), we account for this by including $dumprog_{it} \cdot t$ in equation (3):

$$(5) \qquad C_{it} = \boldsymbol{X}_{it}\boldsymbol{\beta} + \gamma_1 prog_{it} + \gamma_2 dumprog_{it} \cdot t + \boldsymbol{M}_i\boldsymbol{\alpha} + g_i t + u_{it}$$

where $dumprog_{it}$ is a dummy equal to 1 if individual i participated in credit at time t. 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 for equation (3).

The specifications in equations (3) and (5) impose the restriction that borrowing in each loan cycle has the same effect. 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, p. 317) allows program indicators to reflect the frequency of participation in each year as presented in table 1. This is done by replacing $prog_{it}$ and $dumprog_{it} \cdot t$ in equation (5) with a series of program indicators for each loan cycle for which the participant has been in the program:

$$(6) \qquad C_{it} = \boldsymbol{X}_{it}\boldsymbol{\beta} + \gamma_1 prog1_{it} +, \ldots, + \gamma_k progk_{it} + g_i t + \boldsymbol{M}_i\boldsymbol{\alpha} + u_{it}$$

[6] This can be done using a within transformation or by differencing the equation (again) to eliminate g_i and apply OLS on the transformed estimation. The latter is preferred if u_{it} after the first differencing cannot be assumed white noise (Wooldridge 2002, p. 316).

Table 3. 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^2	**−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 that **branch** was available in nearest town (yes = 1)	77.345 (110.607)	0.359 (0.720)
R^2; pseudo R^2	1912	0.150
$F(14, 336)$; Wald χ^2 (14)	**7.80*****	**70.370*****
Sample size	351	351

Note: ♣OLS estimates; †logit estimates. Single asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10%, 5%, and 1%, respectively. Robust standard errors in parentheses.

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 weight to differences between households' degree of participation regardless of year of participation. More weight is also given to the timing of participation within each indicator. Estimation follows the same procedures as for equations (3) and (5). Finally, note that since one of our outcome variables, housing improvement, is a binary indicator, the estimated model is a linear probability model with panel data.[7]

[7] Binary choice models with panel data—for example, a panel probit or a panel logit model—are difficult to estimate due to the incidental parameter problem. Angrist (2001) emphasizes that rather than imposing distributional 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 at hand. Therefore, we use the simple LPM specification, which also provides estimates for the program dummy variables that can be conveniently interpreted as the effect on the mean of the dependent variable (Wooldridge 2002, p. 454–457).

Estimation Results

This section provides estimation results from the models outlined earlier, with selection bias test results being presented first. The test is carried out by estimating equation (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 at a 1% significance level for both the OLS and logit models. See table 3 for the results.

The most important results from this regression are 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. Note that our proxy for the program placement bias (i.e., "branch") is not perfect, since it does not indicate the order of opening. Remoteness of villages is associated with lower consumption. So, if DECSI targeted remote villages later, this variable could also indicate program placement bias. But again, information on order of opening was not available so that we cannot exactly infer the presence of program placement bias. Moreover, the panel data methods used to tackle self-selection bias also mitigate potential bias due to program placement by removing time-invariant variation. Second, the hypothesis that there is no significant difference between *New* borrowers and those that never borrowed (H_0: $\beta_5 = 0$) cannot be rejected in both models. However, a similar null hypothesis for *Always* is rejected at the 10% significance level in the consumption expenditure model, though not in the housing improvement model. Thus, controlling for dropouts, we do not find that those who will eventually become borrowers in 2000 had higher consumption levels than those who never borrowed. However, we find some evidence that those who always borrowed had consumption levels higher than those who never borrowed. Thus, our impact analysis must account for potential bias due to self-selection.

The basic model given in equation (1) is estimated by the standard fixed-effects approach. Since we are interested primarily in credit impact estimates, only household observables that may systematically correlate with selection are included. Moreover, since time-invariant characteristics are removed by the within transformation, only time-varying variables are included, i.e., land size and its square and household head's age and its square.

Although land in Ethiopia is state owned, farmers are given use rights. Cultivated land area therefore determines amount of input use, including credit. One implicit borrower screening criterion of DECSI is household head age. Additionally, as household heads become older, they self-select out of borrowing activities. A year dummy (equal to one for 2006, zero otherwise) is included to contrast the relatively stable and good harvest year 2006 with the earlier years, which are characterized by adverse conditions such as war and drought. The fact that 2006 was a very good year is also reflected in table 2, showing that average deflated consumption in that year was much higher.

This specification is similar to Tedeschi's (2008) fixed-effects model except that our specification considers the cumulative effect of several loan cycles versus "number of participation days" used in the former article. Results are reported in table 4. The F-test statistics indicate that for both household consumption and housing improvement models, not all parameters are jointly equal to zero at the 1% significance level.

Based on the fixed-effects estimation, credit has a significant positive effect on annual household consumption expenditure and housing improvements of borrowers compared with nonborrowers. After controlling for potential selection on unobservable fixed effects, household per capita consumption for an average borrower household has increased by ETB415 for each additional borrowing year.

Table 4. 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***** (27.584)	**0.273***** (0.015)
Women-headed household	61.058 (51.853)	−0.038 (0.028)
Additional skills other than farming	62.136 (60.823)	0.039 (0.033)
Year 2006 dummy	**264.098***** (38.227)	−0.012 (0.021)
Age of household head	10.216 (9.597)	0.004 (0.005)
Age^2	−0.059 (0.090)	-0.628×10^{-4} (0.491×10^{-4})
Cultivated land size (in *Tsimad* = 0.25 *hectare*)	−11.735 (9.378)	−0.002 (0.005)
Land size^2	0.066 (0.295)	$-0.139 \times 10^{-3}(0.162 \times 10^{-3})$
Intercept	−289.897 (246.768)	−0.168 (0.135)
Within R^2	0.215	0.257
$F(8, 1045)$	**35.77*****	**45.250*****
Household fixed effects	**Jointly significant*****	Not jointly significant
Number of observations	1404	1404

Note: Single asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10%, 5%, and 1%, respectively. Standard errors in parentheses.

Table 5. 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**** (77.065)	**160.738**** (79.016)
Random trend *borrowing participation	–	**33.858**** (16.043)
Year 2006 dummy	**323.439***** (32.594)	**324.497***** (32.517)
Age of household head	2.003 (9.428)	1.632 (9.407)
Age^2	−0.022 (0.089)	−0.017 (0.089)
Cultivated land size (in *Tsimad* = 0.25 *hectares*)	−0.496 (13.249)	−1.739 (13.229)
Land $size^2$	0.139 (0.463)	0.193 (0.462)
Intercept	−130.553 (88.088)	−113.738 (88.230)
R^2	0.164	0.169
$F(6, 695)$; F (7, 694)	**22.640*****	**20.14*****
Number of obs.	702	702

Note: Single asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10%, 5%, and 1%, respectively. Standard errors in parentheses.

To put this figure in perspective, in 2004–05 national per capita real consumption averaged ETB1,256 (at 1995–96 constant prices; Ministry of Finance and Economic Development of Ethiopia 2008), and the average loan size provided by DECSI during the study period ranged from ETB500 to ETB1,000, with a maximum loan size of ETB5,000 (Woldehanna 2005, p. 240). 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 significant in the consumption equation, indicating that for both borrowers and nonborrowers, the relatively good conditions in this year increased consumption on average by ETB264, which was already suggested by the raw data in table 2. Surprisingly, these good conditions did not lead to a significant increase in the probability of housing improvement.

The individual household heterogeneity not identified by the included variables is captured by the fixed-effects parameters. For the household consumption model, there is evidence for household heterogeneity given the significance of the fixed effects. This is, however, not the case for the housing improvement model.

As indicated in the previous section, error terms may correlate due to selection based on time-varying individual-specific unobservables. In that case the individual trend model as specified in equation (3) is more robust than the standard fixed-effects 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*. Both models for consumption and housing improvement are estimated by OLS after differencing twice to eliminate these trend components. Since the results for the housing improvement model are very similar to the fixed-effects results presented in table 4, we do not report them here. Results for household consumption are reported in the first column of table 5.

In general, removing individual-specific unobserved dynamics by including an individual trend and differencing the data twice produces more conservative results. Specifically, according to this individual-specific trend specification, per capita annual consumption increases by ETB199 per year of credit taken. This result is statistically significant, but compared with the fixed-effects result, the impact of credit is substantially lower (more than 50% lower per year of credit taken). This difference may be a consequence of the bias in the standard fixed-effects model due to time-varying individual dynamics.

A variant of the individual-specific trend model given in equation (5) not only allows individual household consumption to vary at different *trends* but also allows borrowing effects to depend on these unobserved individual-specific *trends*. Note that in this case, *trend* interacts with the participation ($dumprog_{it}$) indicator and not with "number of years in borrowing." The results of this model are reported in column 2 of table 5.

Table 6. 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^2	−0.027 (0.089)	$0.531 \times 10^{-4} (0.623 \times 10^{-4})$
Cultivated land size (in *Tsimad = 0.25 hectare*)	−0.887 (13.250)	−0.004 (0.009)
Land size^2	0.175 (0.463)	$-0.159 \times 10^{-3} (0.325 \times 10^{-3})$
Intercept	16.268 (70.153)	−0.017 (0.049)
R^2	0.170	0.044
$F(9, 692)$	**15.76*****	**3.560*****
Number of obs.	702	702

Note: Single asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10%, 5%, and 1%, respectively. Standard errors in parentheses.

The credit effect estimate is both quantitatively and qualitatively consistent with 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 ETB161 directly and by ETB34 indirectly (by changing other unobserved time-varying individual characteristics). Thus, after accounting for selection biases, credit is found to be responsible not just for changing the *levels* at which yearly per capita consumption grew for an average borrowing household in the ten years considered, but also for the *rate*. Note that other results are also consistent across the two specifications presented in table 5.

The results in table 5 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. An important follow-up question is, therefore, whether impact can be associated with the extent of repeat borrowing. This issue is analyzed using the flexible individual-specific trend model given in equation (6), which assigns indicators for the number of times each household has been involved in borrowing. The results of this model are given in table 6.

Again, the double-differencing estimation procedure reduces the risk of selection bias. Results show once again that borrowing has a significant impact on consumption, but interestingly enough, the magnitude of impact increases with the length of relationship with the MFI. Specifically, compared with nonparticipants, per capita consumption has significantly increased by ETB274 for one-year participants and by ETB319 for two-year participants.

The effect is statistically insignificant for three-year participants ($p = 0.145$) but is significant for four-year participants. For the latter, per capita consumption increased substantially (by ETB665) compared with that for nonparticipants. This pattern may be due to nonmonotonic consumption patterns of households. As their income increases, they may spend more on basic consumption items (e.g., food) at first, followed by spending on durable assets (e.g., housing improvements) and then by other consumption items included in the consumption variable (e.g., health, education, clothing).[8] This interpretation is in line with the effects of borrowing on housing improvement, which will be discussed.

Although there seems to be a cumulative effect of repeated borrowing indicated by the increasing size of the borrowing parameters, one-sided t-tests indicate that these differences are not statistically significant. Testing null hypotheses $\gamma_k = \gamma_l$ (with $l < k$) versus the alternative $\gamma_k > \gamma_l$ indicates that the null hypothesis cannot be rejected in any of these tests at a 5% critical level (p-values 0.15 for $\gamma_4 = \gamma_3$, 0.15 for $\gamma_4 = \gamma_2$, 0.12 for $\gamma_4 = \gamma_1$, 0.52 for $\gamma_3 = \gamma_2$, 0.43 for $\gamma_3 = \gamma_1$, and 0.37 for $\gamma_2 = \gamma_1$, respectively).

[8] We thank an anonymous reviewer for pointing out this idea of nonmonotonic household consumption patterns.

For the housing improvement model, the probability of improving the house significantly increased after the second round borrowing and increased to 0.244 if households borrowed for two periods, 0.555 if they borrowed for three periods, and 0.457 if they borrowed in all periods compared with nonborrowers. The relatively lower effect in the case of borrowing in all four periods compared with borrowing in three periods is not surprising, as households eventually shift attention from improving their houses to other activities. Testing null hypotheses on the equality of parameters using one-sided t-tests indicates that four years of borrowing does not lead to a statistically significant increase in the probability of house improvement over three years ($p = 0.53$) or two years ($p = 0.18$) of borrowing, but it does increase significantly compared with one year of borrowing ($p = 0.03$). Three years of borrowing leads to a statistically significant increase over two years ($p = 0.02$) or one year of borrowing ($p = 0.00$). Additionally, two years of borrowing leads to a significant increase in the probability of housing improvement compared with one year of borrowing ($p = 0.01$).

Compared with the average impact obtained from the individual-specific trend model, the results in table 6 support the hypothesis that credit has a lasting impact over time. Thus, while the impact of onetime borrowing is close to the average impact previously obtained, having borrowed multiple times leads to even higher probabilities of house improvements. Such high percentage increases attributed to credit are not surprising given the importance of credit at such low initial conditions (e.g., initial average per capita consumption is ETB442 for participants and ETB370 for nonparticipants) and the relatively long period covered by the data, during which Ethiopia experienced 8–11% GDP growth.

Conclusions

Impact evaluations are often prone to self-selection and program placement biases. This article uses panel data techniques to deal with these potential selection biases. Standard fixed-effects models mitigate selection based on time-invariant unobservables, whereas the more advanced random trend model also accounts for individual trends in time-varying unobservables. The dataset used is a unique four-round panel data set concerning households in Tigray, Ethiopia, which 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 an absence of bias due to self-selection. The analysis therefore accounts for potential selection bias.

The results indicate that microfinance credit significantly raised both annual per capita household consumption and the probability of improving housing (roofs), which is an important welfare indicator in this area. The random trend model with flexible participation indicators shows that per capita household consumption (except in the bad year 2003) and the probability of improving the house increased with the frequency of participation, although these increases were not statistically significant in the case of household consumption. Onetime borrowing has no impact on housing improvements, but it does lead to significant improvements in per capita consumption, which is plausible at such early stages of livelihood changes for households in those marginal areas. Repeat borrowing does matter in both cases, however. Although the results of the fixed-effects 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.

These findings have a number of implications. First, they show that the effect of credit on livelihoods can be multidimensional and may not be fully captured by just a single household outcome. Moreover, the effect is not monotonic over time. Second, the results also indicate that the effect of borrowing lasts longer than one or two periods. It takes time before the effect of borrowing on livelihoods is fully materialized. Therefore, impact estimates that rely on a single household indicator and focus only on one cycle of borrowing may underestimate the potentials of microfinance credit on overall livelihoods that can be achieved over time. Future research must focus on more robust specifications that incorporate temporal as well as multidimensional effects of credit on livelihoods.

Finally, an implication of these results for MFI practitioners such as DECSI is that eligible households should be encouraged not only to borrow, but also, if successful, to remain longer in a borrowing relationship in order to realize the full potentials of borrowing.

References

Angrist, J. D. 2001. Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice. *Journal of Business and Economic Statistics* 19(1): 2–16.

Armendáriz de Aghion, B., and J. Morduch. 2005. *The Economics of Microfinance*. Cambridge, MA: MIT Press.

Berhane, G., C. Gardebroek, and H. A. J. Moll. 2009. Risk-matching Behavior in Microcredit Group Formation: Evidence from Northern Ethiopia. *Agricultural Economics* 40(4): 409– 419.

Borchgrevink, A., J. H. Valle, and T. Woldehanna. 2003. *Impact Studies: the Case of DECSI in Ethiopia. SOS FAiM, Zoom Microfinance*, no. 11, September.

Central Statistical Agency of Ethiopia. 2008. *Consumer Price Index at Country and Regional Levels*. http://www.csa.gov.et/Consumer_Price_Index.htm (accessed November 10, 2010).

Coleman, B. 1999. The Impact of Group Lending in Northeast Thailand. *Journal of Development Economics* 60(1): 105–141.

Copestake, J., S. Bhalotra, and S. Johnson. 2001. Assessing the Impact of Microcredit: A Zambian Case Study. *Journal of Development Studies* 37(4): 81–100.

Edgcomb, E. L., and C. Garber. 1998. *Practitioner-Led Impact Assessment: A Test In Honduras*. AIMS project of USAID, Washington, DC.

Friedberg, L. 1998. Did Unilateral Divorce Raise Divorce Rates? Evidence from Panel Data. *American Economic Review* 83(3): 608–627.

Hagos, F. 2003. *Poverty, Institutions, Peasant Behavior and Conservation Investment in Northern Ethiopia*. PhD thesis, Agricultural University of Norway.

Heckman, J., and V. J. Hotz. 1989. Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training. *Journal of the American Statistical Association* 84(408): 862–880.

Karlan, D. 2001. Microfinance Impact Assessments: The Perils of Using New Members as a Control Group. *Journal of Microfinance* 3(2): 75–85.

Karlan, D., and N. Goldberg. 2007. *Impact Evaluation for Microfinance*. Doing Impact Evaluation no. 7, Thematic Group on Poverty Analysis, Monitoring and Impact Evaluation, World Bank.

Khandker, S. 2005. Microfinance and Poverty: Evidence Using Panel Data from Bangladesh. *World Bank Economic Review* 19(2): 263–286.

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

Ministry of Finance and Economic Development of Ethiopia. 2008. Dynamics of Growth and Poverty in Ethiopia (1995/96–2004/05). Addis Ababa. http://www.mofed.gov.et/Uploaded/Publication/DynamicsOfGrowth-and-Poverty-Final2009.pdf (accessed November 10, 2010).

MkNelly, B., and K. Lippold. 1998. *Practitioner Led Impact Assessment: A Test in Mali.* AIMS project of USAID, Washington, DC.

Papke, L. E. 1994. Tax Policy and Urban Development. *Journal of Public Economics* 54: 37–49.

Pitt, M., and S. Khandker. 1998. The Impact of Group-based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter? *Journal of Political Economy* 106(4): 958–996.

Ravallion, M. 2001. The Mystery of Vanishing Benefits: An Introduction to Impact Evaluation. *World Bank Economic Review* 15(1): 115–140.

Tedeschi, G. A. 2008. Overcoming Selection Bias in Microcredit Impact Assessments: A Case Study in Peru. *Journal of Development Studies* 44(4): 504–518.

Tedeschi, G. A., and D. Karlan. 2010. Cross Sectional Impact: Bias from Dropouts. *Perspectives on Global Development and Technology* 9(3–4): 270–291.

Woldehanna, T. 2005. The Impact of Dedebit Credit and Saving Institution on Poverty Reduction and Community Development. In *Proceedings of the Institutional Conference on Microfinance Development in Ethiopia*, ed. G. Ageba, 236–269. Addis Ababa: Association of Ethiopian Microfinance Institutions.

Wooldridge, M. J. 2002. *Econometric Analysis of Cross-sectional and Panel Data.* Cambridge, MA: MIT press.