CREDIT Research Paper

No. 10/08

Has the Introduction of Microfinance Crowded-out Informal Loans in Malawi?

by

Richard Disney Eleonora Fichera Trudy Owens

Abstract

This paper uses household data to test whether microfinance institutions created by the Malawian government in the mid-1990s under its Poverty Alleviation Programme crowded out access to informal loans. As in several recent studies, the paper adopts policy evaluation techniques to identify a causal relationship between access to government-sponsored credit programmes and informal loans. After taking treatment heterogeneity into account with a multiple treatment model, the paper finds strong evidence of crowding out of formal group lending on informal sources. In particular, participation in the most widespread microfinance programme has a negative and significant effect on borrowing from informal sources, reducing on average the amount that members borrow from informal lenders by more than 70 percent of the average loan value.

JEL Classification: O16 O17

Keywords: informal lending microfinance evaluation methods

Centre for Research in Economic Development and International Trade,

University of Nottingham

CREDIT Research Paper

No. 10/08

Has the Introduction of Microfinance Crowded-out Informal Loans in Malawi?

by

Richard Disney   Eleonora Fichera   Trudy Owens

Outline

1. Introduction
2. Description of data and credit programmes
3. Descriptive statistics
4. Evaluation strategy
5. Conclusions

The Authors

* School of Economics, University of Nottingham, and Research Fellow, Institute for Fiscal Studies.

† School of Community-based Medicine, University of Manchester, and School of Economics, University of Nottingham.

‡ School of Economics, University of Nottingham.

Acknowledgements

Research Papers at www.nottingham.ac.uk/economics/credit/

1. Introduction

In developing economies such as those of sub-Saharan Africa and South Asia, complete informal insurance against income shocks is usually infeasible. This is for a variety of reasons, including the scarcity of household collateral, poor legal enforcement and the co-variance of household risks in village environments (see, for example, Deaton, 1992a, 1992b; Rosenzweig and Wolpin, 2004; Udry, 1994; Dubois, Jullien and Magnac, 2008). At the same time, the level of local information that is required for efficient market coverage by large credit institutions in such markets is limited. Commercial banks have funds to lend, but often lack adequate information on creditworthiness and enforcement mechanisms to recover the loans. One public policy response to these market failures is to create microfinance institutions that can acquire information on borrowers in innovative ways. By mimicking and exploiting some of the features of informal lending, banks can design credit contracts that harness local information and give borrowers incentives to use their own information on their peers to the advantage of the bank (Armendariz and Morduch, 2005; Ray, 1997).

Within this setting, the 2001 World Development Report notes that “informal and formal strategies are not independent: public policies and the availability of formal mechanisms heavily influence how extensively informal arrangements are used and which kinds are used” (World Bank, *ibid*, p.140). A natural question, therefore, is: do governments displace the informal loan market by introducing formal microfinance institutions? To the extent that such ‘crowding out’ occurs the effectiveness of policy interventions of this type is reduced. Whether this is desirable depends on whether such informal credit channels are seen as desirable in themselves.

In the mid 1990s, the Malawi Government published its Policy Framework for the Poverty Alleviation Programme (PAP). Among the strategies proposed in the PAP were the provision of credit facilities and the promotion of micro and small enterprises. As a consequence, new credit facilities were created by the government of Malawi of which some received funds from external agencies such as the World Bank. Loans were delivered to small groups for farming activities, such as the acquisition of agricultural inputs (i.e. fertilizers, seeds and farm equipment), and for small-scale trading activities. To examine the hypothesis that these new institutions crowded out use of informal credit sources, this paper uses the 1995 Malawi Rural Financial Markets and Household Food Security Survey (FMHFS), described more fully in the next section of the paper. The survey contains information about households' borrowing behaviour from both informal lenders and group-lending institutions.

We find significant evidence of crowding out of lending from informal sources for at least some of the programmes investigated. The results show that participation in one specific microfinance programme, that provided by the Malawi Rural Finance Company (MRFC), had a significantly negative effect on borrowing from informal sources. The average absolute size of the effect ranged from 20 to 32 Malawian Kwacha (around 1.3 to 2 U.S. dollars at 1995 exchange rates) depending on the specification. In relative terms, it reduced the amount that members borrowed from informal lenders by more than 70 percent.

The relatively large literature on crowding out in the last fifteen years has found no consensus on the effect of government sponsored programmes on pre-existent private schemes. Most of this literature tests the crowding out hypothesis with regression techniques where the dependent variables are private transfers or remittances and the independent variables include, among other controls, some form of public transfers. Typically either probit or tobit models are used, although non-parametric specifications have also been implemented (Jensen, 2003). The problem with these studies is the endogeneity bias that arises from the non-random selection of participants into the public programme. Several recent studies have attempted to resolve this issue by using alternative methodologies, such as instrumental variables, randomised treatments, and evidence from pre- and post-programme participation data (e.g. Albarran and Attanasio, 2002; Attanasio and Rios-Rull, 2000; Cox et al., 2004; Jensen, 2003; Kaboski and Townsend, 2006; McKernan et al., 2005).

Following these later studies we adopt policy evaluation techniques in order to identify a causal relationship between access to formal credit programmes and reduction in the use of informal loans in Malawi. An innovation of this paper is that it departs from the standard single treatment approach, showing that this method can lead to "aggregation bias" whenever treatment heterogeneity is not taken into account. We develop a model with multiple treatments where households are classified as members of one, or more than one, group-lending programme. In doing so, we follow the evaluation literature on training programmes (for example, Brodaty et al., 2001; Frölich et al., 2004). This approach allows a comparison between the effectiveness of different mixes of credit programmes, as well as between different groups of households. For example, crowding out could differ according to the economic status of the household, with relatively constrained (unconstrained) households more (less) likely to reduce borrowing from informal lenders (as in Cox et al., 1998; Cox and Jimenez, 1992; Navajas et al. 2003) if more constrained households switch to group-lending institutions in order to reduce borrowing costs.[1]

The evaluation of the impact of group-lending institutions on access to informal loans requires the use of an untreated group that is similar to the group of treated households who do participate in group-lending. We choose past members of group-lending institutions as the untreated group: we justify this decision later in the paper. Propensity score matching is then implemented to match participants in group-lending institutions with households that have similar observed characteristics (the so-called "control group"), but are not current members of any group-lending institution.

The paper evaluates both the effect of being a borrower and a member of one or more group-lending programme. This allows us to test the crowding out hypothesis in the presence of expected transfers, as opposed to nearly all the literature which has focused on crowding out in the context of realised transfers. We know households' demand for informal loans is affected by their membership of a microfinance programme and not just by actual borrowing (Cox and Fafchamps, 2008). The data set also provides information on self-reported demand for loans and on credit limits. Such data are not commonly available in studies of crowding-out and allow us to attempt to disentangle demand and supply factors.

[1] Few empirical studies have tested the crowding out hypothesis in the context of group-lending institutions (although see McKernan et al. 2005). Morduch (2000) has recognized the importance of analysing the role of group-lending institutions in markets where there are a variety of other lenders, but most of the economic literature on group-lending institutions has been concerned with the impact of these institutions on clients (as in Morduch, 1998; Pitt and Khandker, 1998; Wydick, 1999) and with the ability of joint-liability schemes to overcome information problems affecting formal lenders (Besley and Coate, 1995; Ghatak, 1999; Stiglitz, 1990).

In addition, we develop a rigorous sensitivity analysis by adopting a variety of matching algorithms, and test for hidden biases arising from unobservable factors that affect simultaneously the assignment into one of the programmes and the outcome variable.

The structure of the paper is as follows. In the next section we describe the Malawi Rural Financial Markets and Household Food Security survey. Section three presents some descriptive statistics of the sampled communities and of households' borrowing behaviour. The evaluation strategy is explained in section four. Section five concludes.

2. Description of data and credit programmes

2.1 The Malawi Rural FMHFS survey

The Malawi Rural Financial Markets and Household Food Security survey (FMHFS)[2] was conducted by the International Food Policy Research Institute (IFPRI) in cooperation with the Rural Development Department of Bunda College of Agriculture. It was part of a study on the determinants of access to and participation in existing formal and informal credit and saving programmes, and their effects on agricultural productivity, income generation and food security. The Malawi FMHFS was collected in three rounds during 1995: the first round took place between February and April, the second in July-August and the last in November-December. The survey includes detailed information on land tenure and agricultural production, assets, food and non-food consumption, credit and savings, and wage and self-employment income (see Diagne, 1999).

The sample includes 404 households in 44 villages in five districts of Malawi. The five districts were Dedza, Dowa, Mangochi, Nkhotakota and Rumphi (see map overleaf). The data were collected using a stratified sampling procedure to ensure that half of the sample was a member of a programme, and then a random selection within each stratum was interviewed[3].

---

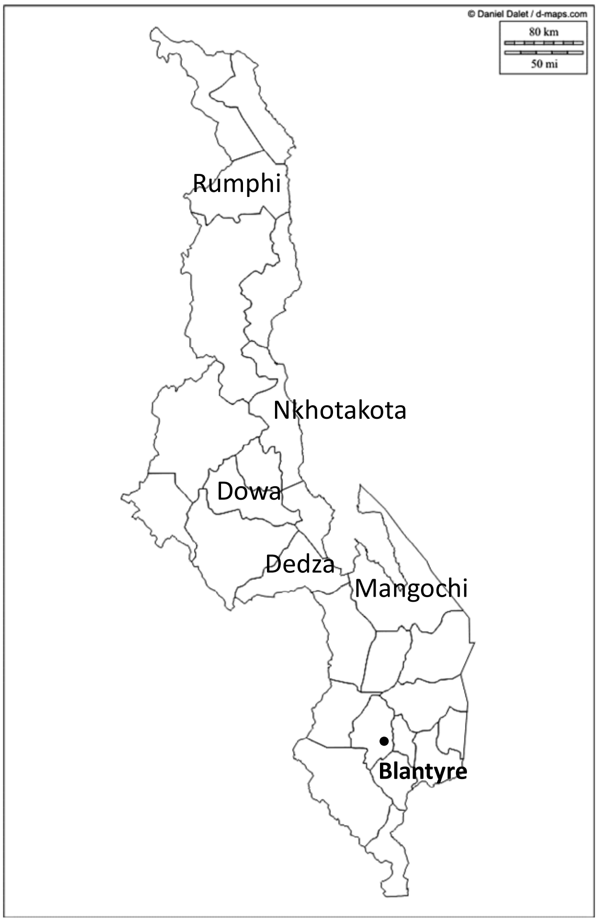
[2] Funding for this survey came from the Rockefeller Foundation, and GTZ/ Malawi via the Ministry of Women, of Children Affairs, of Community Services and Social Welfare (MOWCACDSW), and through UNICE- F/Malawi and USAID/Malawi.

[3] The sample is choice-based because households were selected according to their participation in the credit programmes.

2.2 The credit programmes

This paper focuses on four microfinance programmes created as a result of the 1995 PAP. Half of the stratum-selected sample participated in four credit programmes: the Malawi Rural Finance Company (MRFC); the Malawi Muzdi Fund

Map: Sites of districts in Malawi



(MMF); the Promotion of Micro-Enterprises for Rural Women (PMERW); and the Malawi Union of Savings and Credit Cooperatives (MUSCCO). The other half of the sample had either, previously participated in one of the credit programmes, or had never participated in a formal credit programme.

The credit and savings module of the questionnaire was administered to every member of the household who was over 17 years of age. Information was collected on the credit characteristics of all individuals who reported a positive demand for loans, whether from a formal or informal source, and for both those who had loan applications rejected and granted.

The next section describes each of the programmes. The Malawi Rural Finance Company (MRFC) is funded by the World Bank. It provides in-kind seasonal agricultural loans for fertilizers, seeds and pesticides for hybrid maize and tobacco. It also offers short-term (two year) and medium-term (five year) loans for farm equipment. Targeted individuals are formed into jointly-liable groups of 5-10 smallholder farmers. The MRFC also offers two savings deposit services to its borrowers: ordinary and contract savings accounts. With a contract savings account, clients can choose the amount and timing of deposits. By honouring loan commitments, participants either get a bonus interest or are offered a credit limit without having to secure it with collateral.

The Malawi Mudzi Fund (MMF) most closely resembles the standard micro-credit facility; indeed it was designed to replicate the Grameen Bank in Bangladesh. It is funded by the World Bank and by the International Fund for Agricultural Development (IFAD). The MMF targets poor households with less than one hectare of land and provides loans for non-farming income generating activities. To qualify for loans first-time Mudzi borrowers are not required to provide collateral, but are required to undergo a six-month training period and form groups - which are based on the Grameen model of groups of five and centres of 20-25.

The Malawi Union of Savings and Credit Cooperatives (MUSCCO) is the apex organization for Savings and Credit Co-operatives (SACCOs) which offer a range of services including credit, savings and insurance. It was created in 1980 and is financially supported by the United States Agency for International Development. The MUSCCO is the principal Malawian financial institution actively promoting savings mobilization. It has not experienced the default rates that have characterized other lending operations, primarily because it is member-based and funds loaned represent members' own savings.

The Promotion of Microenterprises for Rural Women (PMERW) is a credit programme financially supported by the German Agency for Technical Cooperation (GTZ). It was started in 1986 by the Ministry of Women and Children's Affairs and Community Services ('PMERW1' in Table 2). The most recent version of the credit programme targets groups of women of 5-10 who are skilled in business activities ('PMERW2'). The structure is similar to the saving and credit clubs except that members can borrow up to MK 1,000[4] and receive loans directly from the Central Bank of Malawi. Credit members are selected among those who have excellent credit and business management skills.

3. Descriptive statistics

This section provides some descriptive statistics of the sampled communities and of households' borrowing behaviour. The statistics have been weighted to correct for choice based sampling.

3.1 Community level

The community survey was undertaken in 1995 and includes information concerning the demographic characteristics, infrastructure and agricultural production of the villages. Among the five sampled districts, Mangochi stands out as being the most populated, with greater access to irrigation and a larger share of cultivated land. Table 1 displays the availability of tarred or gravel roads in each district which proxies for access to basic services. Large differences between districts are evident. In Mangochi, all roads to the government office, credit office, post office and commercial bank are tarred or gravel, as well as 80 percent of roads to primary schools and health centres. Distances in Mangochi are also relatively small: on average it is less than 10 Km to a government office, credit office, post office, primary school and health centre, although the nearest commercial bank is more distant at 102 Km. At the other extreme, in Dowa there are no tarred roads to the government office, credit office, post office, commercial bank and health centre.

Table 2 shows the existence of formal credit groups and informal moneylenders in each district. Not all formal credit programmes are available nationally. MRFC groups are present in all districts and on average have existed for more than two years. MUSCCO groups exist only in Dowa and Nkhotakota. These credit groups are newer than MRFC groups. MMF groups are also less widespread: they only exist in

[4] MK = the Malawian Kwacha. In 1995 the exchange rate of the Malawian Kwacha to the US dollar was around MK17=US$1; by late 2009 the Kwacha had depreciated to the extent that the ratio was MK140=US$1.

Mangochi and Rumphi and are also relatively new. Although Dedza has more villages than other districts, it only hosts MRFC groups. In Dowa, Mangochi and Nkhotakota there are a few moneylenders (two, three and five, respectively). In Dedza there are no

Table 1: Infrastructure by district: tarred or gravel roads

District	*Sample size*	*% road access to gov. office*	*Average distance to gov. office (Km)*	*% road access to credit office*	*Average distance to credit office (Km)*	*% to road access post office*	*Average distance to post office (Km)*	*% road access to primary school*	*Average distance to prim-ary school (Km)*	*% road access to commerci al bank*	*Average distance to com. bank (Km)*	*% road access to health centre*	*Average distance to health centre (Km)*
Dowa	56	0	20	0	6	0	20	33	4	0	20	0	20
Mangochi	102	100	10	100	0.3	100	8	80	2	100	102	80	8
Nkhotakota	70	71	4	71	4	71	1	86	1	57	69	71	24
Rumphi	75	0	6	22	4	11	4	22	1	0	29	0	6
Dedza	101	70	26	80	9	20	23	30	3	55	44	55	34

Source: Own calculations based on FMHFS, community data.

Table 2: Credit sources by district

District	***MRFC***		***MMF***		***MUSCCO***		***PMERW1***		***PMERW2***		*N. of money-lenders in the district*	*N. of money-lenders out of district lending to HHs*
	N. of groups	*Average n. of years of existence*	*N. of groups*	*Average n. of years of existence*	*N. of groups*	*Average n. of years of existence*	*N. of groups*	*Average n. of years of existence*	*N. of groups*	*Average n. of years of existence*		
Dowa	1	10	0	0	2	5	0	na	0	na	2	1
Mangochi	4	4.2	4	1	0	0	3	na	1	na	3	0
Nkhotakota	3	2.4	0	0	1	1.7	3	na	4	na	5	0
Rumphi	4	1.6	1	Na	0	0	3	na	3	na	0	0
Dedza	5	1	0	0	0	0	0	na	0	na	0	1

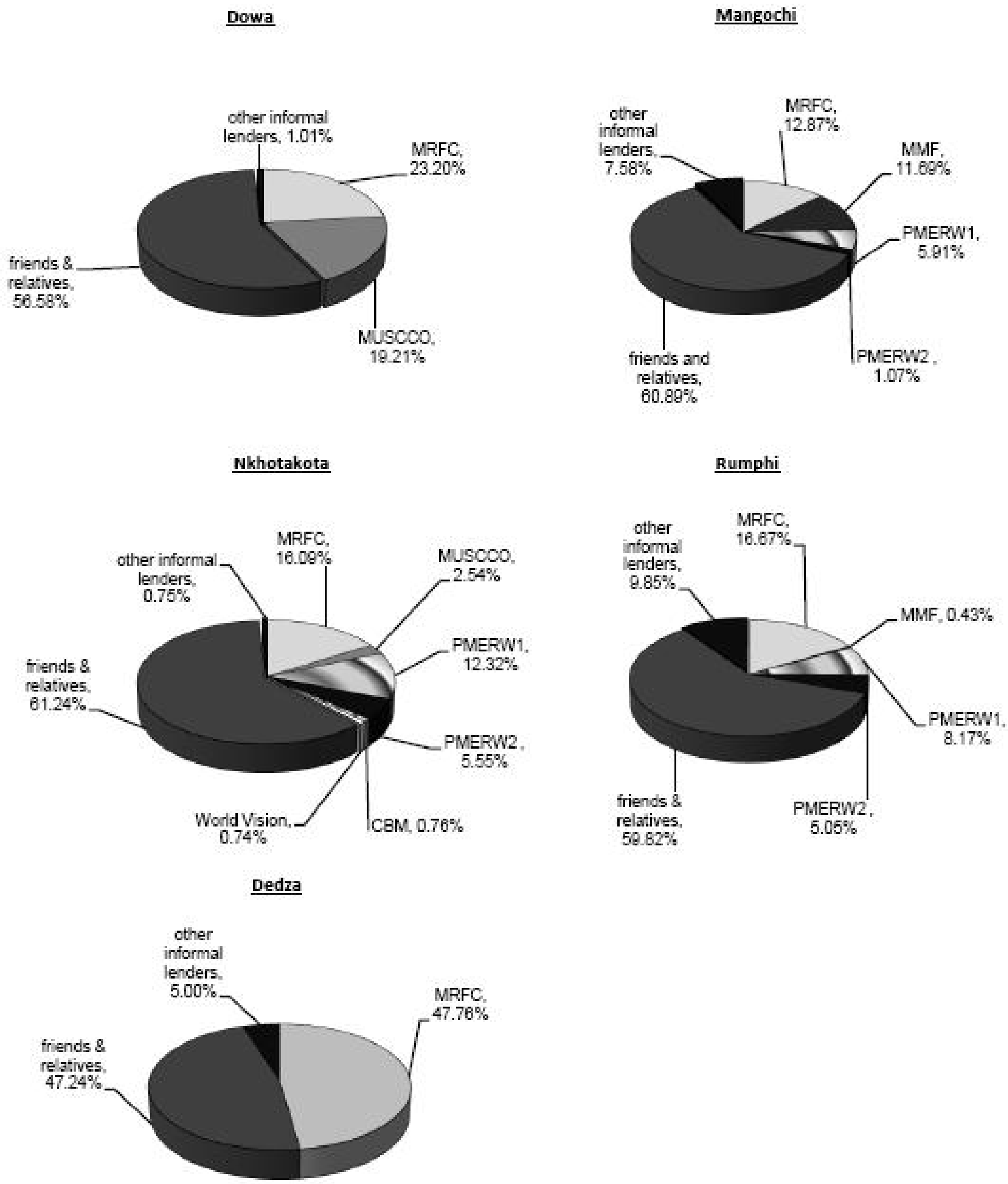
Source: Own calculations based on FMHFS, community data. na = information not available.

local moneylenders, but an outside moneylender does operate in the district.

3.2 Household level

Figure 1 displays the distribution of loans by source and district reported by sample

Figure 1. Distribution of loans sources by district



Source: Own calculation based on FMHFS.

respondents. Not all loan sources are utilised in every district, partly reflecting availability (as described above). Friends and relatives, and other informal lenders are used in all districts; indeed friends and relatives remain the dominant source of finance overall. Among the formal programmes, the MRFC programme is the most diffused across districts – in Dedza district it is the dominant source of finance.

Table 3. Characteristics of households by loan status

	Participants in formal loans	Rejected by formal lenders	Non-applicants	Past participants
Household characteristics				
Female headed HHs (%)	22.0 (541)	20.5 (51)	38.1 (311)	10.6 (159)
Average HH size	5.8 (541)	5.3 (49)	4.5 (311)	5.7 (159)
Average number of children 0-15	3.0 (541)	2.6 (49)	2.2 (311)	3.2 (159)
HH's head main occupation: agriculture (%)	48.8 (540)	71.9 (51)	69.9 (307)	86.7 (159)
Households' assets and shocks:				
HHs affected by negative income/health shocks (%)	59.4 (535)	42.9 (48)	53.9 (309)	51.6 (154)
Average land size (ha)	2.2 (517)	2.4 (47)	1.4 (288)	2.1 (153)
Share of land owned by spouse	16.2 (517)	3.2 (47)	16.2 (288)	24.5 (153)
Average value of house (MK)	1055 (517)	700 (47)	463 (288)	846 (153)
Share of assets held as land (%)	56.4 (517)	62.9 (47)	56.8 (288)	63.7 (153)
Average food expenditure (MK)	13.7 (517)	10.4 (47)	12.3 (288)	9.5 (153)

Source: Own calculation based on FMHFS. Numbers in parentheses report response numbers by category.

The dataset allows us to discriminate between several different groups of households: those who borrowed from at least one of the credit programmes; those rejected by formal lenders; non-applicants (i.e. households who never participated in a formal credit programme); and past members (i.e. households who once were members of one of the formal credit programmes). Table 3 displays the characteristics of these four groups: by sample construction just over half the households are borrowers. Non-applicants represent over a quarter of the sample and past participants just under one fifth. There are only 51 households reporting that they were rejected for a formal loan.

Among the four groups, non-applicants have the highest percentage of female headed households (38 percent). Only 11 percent of past participant households were headed by a female. Households who participate in credit programmes have, on average, the highest number of members (around six). Households with current loans from formal organisations tend to be better-off (as measured by asset values) and are more likely to have a household head not involved in agriculture. Not surprisingly, current loan holders are more likely to report that they have recently experienced an income or health shock. However, roughly half of households, including non-participants (who may have obtained loans from other informal sources) also report they have experienced a shock. Rejected applicants have the highest land holdings (2.4 hectares)[5].

4. Evaluation strategy

We examine the impact of credit programme participation on the use of loans from the informal sector. We follow the standard treatment model in the evaluation literature (Roy, 1951; Rubin, 1974) which involves the estimation of the effect of participation in a programme on a hypothetical outcome.

We evaluate the effect of credit programmes on the amount borrowed from informal lenders. We define two mutually exclusive groups: $T=m, l$, and denote the outcome variables as Q^m, Q^l.

The first group, $T=m$, denotes membership of any of the above defined credit programmes. The main advantage of this approach is that it allows us to keep the largest number of possible observations. However, by pooling different credit programmes we neglect the heterogeneity of the credit institutions which could bias our results. This (potential) bias will be discussed in more detail in the next section.

The second group, $T=l$, denotes the case of no treatment. Our untreated group is composed of past members of formal credit programmes. We argue that these individuals might plausibly be assumed to have similar (time-invariant) unobservable

[5] This is a result of the ceiling on land set by the programmes' eligibility criteria.

characteristics (e.g. entrepreneurship ability) as the group of current participants. We do not use households who never participated in any credit programme as the untreated group due to concern that unobservable factors such as entrepreneurial ability may affect selection into the programmes – all credit programmes deliver loans for farming and related activities. Moreover, because we have no information on *past* access to the informal sector, we have no information on whether the group of households who never participated in credit programmes ever demanded credit.

Following Lechner's (1999a) approach we compare participation in a particular treatment with the outcome for the no treatment option. The average treatment effects for a comparison between treatment *m* and *l* are given by:

$$\gamma_0^{m,l} = E(Q^m - Q^l \mid X) = EQ^m - EQ^l \tag{1a}$$
$$\vartheta_0^{m,l} = E(Q^m - Q^l \mid T = m, X) = E(Q^m \mid T = m, X) - E(Q^l \mid T = m, X) \tag{1b}$$

where equation (1a) denotes the average treatment effect (ATE) of *m* relative to treatment *l* for the population; and equation (1b) is the average treatment effect on households treated by programme *m* (ATT).

The evaluation problem lies in the fact that the difference in equation (1b) cannot be observed for the same household. We can only observe the effect on informal borrowing for a household that has been treated by one of the programmes at each point in time. This identification problem could be solved using the Conditional Independence Assumption (CIA), according to which, given a set of covariates *X*, potential treatment outcomes are independent of participation status (Rubin, 1974). The CIA can be applied to the case where a large set of covariates is available using balancing scores (Rosenbaum and Rubin, 1983).

The main drawback of this procedure is that it only holds when observable characteristics are controlled for. Heckman et al. (1997) show, however, that even after conditioning on observables, outcomes of participants and non-participants may still be significantly different. For example, selection into the programmes may be conditioned on unobserved characteristics, and differences in outcomes may arise when participants and non-participants live in different districts or regions. We argue that our untreated group, being composed of past members matched to treated groups in each district reduces selection on unobservables. However, in order to check the robustness of our untreated group, we discuss (possible) selection on unobservables at the end of the paper.

To account for these issues our approach involves three analytical stages. First, we estimate a standard model where the single 'treatment' is defined as participation in at least one credit programme. We construct the propensity scores of participation and then perform the Mahalanobis metric matching algorithm. We apply the constructed propensity scores to estimate the average treatment effects. The outcome variable of interest is the *total* amount households borrow from informal sources. Hence, for those in option m, the mean effect of option m rather than option l is estimated as the mean difference in the amount borrowed from informal lenders between households in option m and the matched households in option l.

Applying this approach, we find no evidence of a treatment response (in terms of a reduction in borrowing from informal sources) but significant evidence of treatment heterogeneity. We consider whether this arises from heterogeneity of the treatment or heterogeneity of the treated, and find strong evidence of the latter.

This suggests the need to unpack our treatment variable. To explore this further in the second stage we apply propensity scores to multiple treatments defined as participation in one programme as opposed to participation in one or more credit programmes, all relative to past membership. When we adopt this multiple treatment perspective, we find evidence of a significant crowding out for the most important credit programme - the MRFC.

The final stage of the analysis tests whether the results depend on the methodological assumptions of our evaluation procedure. Our sensitivity analysis adopts several alternative specifications: a) changing the regressors of the model and the matching algorithm; b) changing the definition of treatment and outcome; and c) changing the model used to estimate the propensity scores. The results are robust to each specification.

4.1 First stage: propensity scores with single treatment

We first estimate the propensity scores of participation in formal credit programmes. Consider the single treatment $T=m$ to denote membership in any credit programme. $T=1$ denotes the case of no treatment. Our untreated group is composed of past members of credit programmes. The propensity scores are the predicted values $\hat{T}^{k}_{ij(i)} = \hat{P}(T = k \mid T = m, l)$ where i indicates the ith household, $j(i)$ indicates the village where household i lives and $k=m,l$. In this context, m is participation in any one or more of the credit programmes and l is past membership.[6]

We model participation in any credit programme with a logit model. From a theoretical point of view, only those variables that affect both the participation decision and the outcome should be included. We might suspect that, in anticipation of participation, poor households might decrease their effort to increase income (by, for example, reducing their search for employment or their effort to increase production). Ashenfelter (1978) discovered a similar result when evaluating the treatment effects on earnings (the so-called Ashenfelter's Dip; for examples, see Fitzenberger and Prey, 2000; Heckman et al., 1999; Heckman and Smith, 1999)[7].

In order to avoid reverse causality between the covariate X and the participation decision, variables should be either fixed over time (i.e. gender) or should be measured before participation. Because the data set does not contain information about the starting date of membership, the following relatively time-invariant variables were included: household and community characteristics, and semi-fixed factors that affect eligibility such as land size. According to the eligibility criteria set by the credit programmes, credit is delivered to small farm holders and poor households. We can therefore include land size as a covariate in the estimation of the propensity score because it does not change much over time. However, we cannot use agricultural income because it displays variability across seasons, and both affects and is affected by participation.

A further issue arising in this evaluation is that samples are choice-based. In general, choice-based sampling leads to an over-sampling of participants relative to the eligible households in the population. Sampling weights are required to estimate

---

[6] More formally, the propensity scores are given by: $\hat{P}^{m|ml}(x) = \dfrac{\hat{P}^{m}(x)}{\hat{P}^{m}(x) + \hat{P}^{l}(x)}$

[7] This effect can be ignored if the introduction of a new programme is unanticipated. Ideally, this hypothesis could be tested by looking at households' income before and after the creation of microfinance institutions. However, because the data entails only one year (1995) and some of the programmes were created before 1995, we cannot test this hypothesis.

consistently the probability of participation in the credit programmes (Smith and Todd, 2005). However, Heckman and Todd (2009) show that matching methods can be applied even with the propensity scores without weights. This is because the ranking of the observations is simply shifted by a scalar and the same observations are matched. We check the robustness of our results by dropping the weights in the sensitivity analysis and find our main results do not change. Similarly, Frölich et al. (2004) found that dropping the sampling weights did not change the results of their evaluation of a Swedish rehabilitation policy.

We model the choice of participation with a logit model where the treatment denotes participation in any of the credit programmes. The untreated group is past membership. As mentioned above, the choice of the single treatment allows us to keep more observations, but neglects programme heterogeneity. In order to (partially) overcome the latter problem, we include covariates that affect eligibility for all programmes.

Define the logit model as follows:

$$T_{ij(i)}^{*k} = x'_{ij(i)}\beta_0 + C_{j(i)}\beta_1 + u_i \qquad (2)$$

where the subscript $i=1,2,...\ N$ indicates the ith household; and $j(i)$ indicates the village where household i lives. Also, $T_{ij(i)}^{*k}$ is the unobserved propensity to participate where $T_{ij(i)}^{k} = 1.(T_{ij(i)}^{*k} > 0)$; $k=m,l$ indicates the treated and untreated group (i.e. m=any credit programme; and l=past members). The model includes a vector $x_{ij(i)}$ of households' characteristics, education and the occupation of the household head and a vector $C_{j(i)}$ of community characteristics that vary only across villages but not across households. In addition, we include district and time dummies (the latter to pool across different seasons). Results are presented in Table 4.

The propensity scores are the predicted values, $\hat{P}^{m|ml}(x)$, estimated from equation (2) where m is participation in any credit programme and l is past membership. The choice-based corrected probabilities, $\omega\hat{P}^{m|ml}(x)$, are obtained using Manski-Lerman weights (1977).

Table 4. Logit model of participation: single treatment model

Pr(Participation in ...)	Any credit programme vs. past members
Households' characteristics:	
hh size	1.06 *(0.13)*
age head	1.02 *(0.02)*
female head‡	0.43 *(0.25)*
no. children 6-10	1.17 *(0.37)*
no. of days sick (hh head)[1]	1.03 *(0.04)*
Education & occupation of hh head:	
MSCE certificate‡	0.002 *(0.00)****
professional training‡	0.62 *(0.29)*
occupation in agriculture‡	1.67 *(0.78)*
Households' assets:	
land size (ha)	1.50 *(0.18)*
share of land owned by spouse (%)	1.00 *(0.00)*
no. of gifts	1.77 *(0.65)*
Community characteristics:	
total number of households	1.00 *(0.00)***
electricity‡	4.82 *(1.98)****
distance to government office (Km)	1.04 *(0.03)*
distance to credit office (Km)	1.04 *(0.02)***
Dowa‡	1.49 *(1.08)*
Nkhotakota‡	0.73 *(0.67)*
Rumphi‡	4.18 *(3.20)**
Round 2‡	0.51 *(0.28)*
Round 3‡	0.30 *(0.17)***
No. of obs.	1167
Pseudo R^2	0.18

Source: own calculation from FMHFS. Notes: odds ratios displayed and robust std. errors in parentheses). Weighted regressions. ‡dummy variables. [1]month before interview. *** p<0.01; **p<0.05; *p<0.1.

Table 4 displays the odds ratio, e^{β}, for each of the two logit models[8]. We briefly comment on the sign of the coefficients because the regression is only used to predict the propensity scores.

[8] The coefficients should be interpreted as follows: for a unit change in the regressor, the odds are expected to change by a factor e^{β}, holding all other variables constant. The correspondent coefficient, β, can be found by taking the logarithm of the odds ratio. The sign of the coefficient is positive when $e^{\beta} > 1$ and negative otherwise. For binary variables, going from 0 to 1, one can interpret the odds ratios directly without any transformation.

We find the probability of participating in any credit programme decreases if the household head has a MSCE[9] certificate[10]. Members of any credit programme are more likely to live in villages that are larger and have electricity. The fact that the probability of participating in any credit programme increases with the distance to a credit office could be due to the aggregation of programmes. Indeed, we cannot exactly identify which credit programme is supported by which office.

4.1.1 Matching algorithm

Having obtained the scores from the first stage, we match by selecting a control group from the pool of untreated households in which the distribution of observed variables is as similar as possible to the distribution in the treated group. Mahalanobis metric matching with propensity scores is used. This algorithm is implemented by randomly ordering households and then calculating the distance between the first treated households and all controls, where the Mahalanobis distance between a treated household *i* and a control household *j* is defined by:

$$d(i, j) = (P_i^m - P_j^l)'V^{-1}(P_i^m - P_j^l)$$

where P_i^m and P_j^l are the propensity scores in options *m* and *l* for treated household *i* and control household *j*. *V* is the sample covariance matrix from the full set of households. According to this matching algorithm, the control household *j* with the minimum Mahalanobis distance is used as a match for treated household *i* and both households are removed from the pool. The process goes on until matches are found for all treated households.

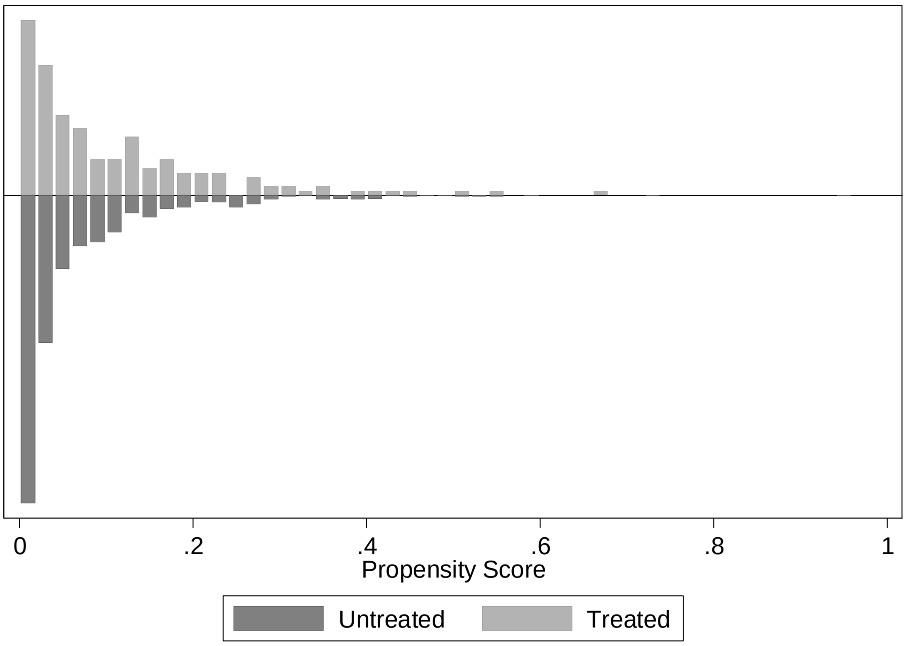
There are several ways to assess the quality of matching. Figure 2 shows the distribution of the predicted propensity scores between treated and untreated groups for participation in any credit programme. Relatively good matching is achieved in the lower tail of the distribution - up to 0.2 - where the distribution of scores between treated and untreated groups is most similar.

[9] MSCE=Malawi School Certificate of Education corresponds to high school certificate at age 16-17.

[10] As described earlier, we did not include income in order to avoid reversed causality. The inclusion of education and professional training which are highly correlated with income partially controls for this omitted variable.

In addition to the analysis of the overlapping regions, we check that matching has reduced the bias between treatments and controls for the set of covariates[11] used to estimate the propensity scores. Table 5 displays the reduction of bias and the two-sample *t*-test for the selected characteristics that had a bias higher than 10 percent prior to matching. In the policy evaluation literature (e.g. Austin and Mamdami, 2006; D'Agostino, 1998; Manca and Austin, 2008), a standardised bias higher than ten percent (and sometimes 20 percent) is taken to denote an imbalance in a covariate between treatments and controls. If matching has worked, the covariates should be balanced and no significant differences should be found after matching.

Figure 2. Bar charts of propensity scores



Source: Own calculation based on FMHFS.

[11] The matching algorithm is conditional on the districts in which control and treated households live.

Table 5. Selected characteristics by treatment groups after matching (when bias prior to matching >10%)

	Any credit programme vs. Past members		**Group comparison**	
	Treated	*Controls*	*t-stat*	*% reduction |bias|*
household characteristics:				
household size	5.99	5.86	0.51	63.70
age head	49.45	48.84	0.43	87.30
female head‡	0.19	0.16	0.69	71.70
number of children 6-10	0.92	0.87	0.24	75.30
education of household head				
occupation in agriculture‡	0.86	0.86	-0.00	100.0
household assets:				
land size (ha)	2.39	1.97	3.18***	6.40
share of land owned by spouse (%)	28.87	26.38	0.56	81.60
characteristics of the community:				
total number of households	245.60	264.81	-0.55	85.70
electricity‡	0.20	0.20	-0.00	100.00
distance to the government (Km)	20.13	18.96	0.74	82.40
distance to the credit office (Km)	9.61	9.42	1.50	96.20

Source: own calculation from FMHFS. ‡dummy variables. [1]month prior to interview. ***p<0.01, **p<0.05, *p<0.1.

The first column of table 5 show the means of treated and control groups for the single treatment model with membership to any credit programme. The last column displays the group comparison for the single treatment model based on the *t*-test and on the absolute percentage reduction in bias obtained by comparing the standardised bias (SB) of treated and control groups before and after matching. The higher the reduction of bias, the better the balance that has been achieved for that covariate. In most cases this reduction in bias is very large – indeed as high as 100% – showing that the matching procedure is successful in balancing the observed characteristics of treatment and control groups. Only land size is still significant after matching.

4.1.3 Estimation of the average treatment effects

We then calculate the average treatment effects for the amounts that households borrow from informal lenders. It is important to re-emphasise that our outcome variable (the impact of access to formal credit institutions on informal lending) refers to a variety of informal lenders: friends, relatives and other informal sources such as moneylenders and traders. However, as illustrated in sub-section 3.1, the majority of households borrow from friends and relatives rather than other informal sources.

Table 6. Average treatment effects with Mahalanobis matching (single treatment)

<table>
<tr><td rowspan="2">Outcome:</td><td colspan="2">(a) Average Treatment Effect on Treated (ATT)
$\vartheta_0^{m,l} = E[Q^m \mid T = m, P^{m|ml}(X)] - E[Q^l \mid T = m, P^{m|ml}(X)]$</td></tr>
<tr><td colspan="2">m=Any credit programme; l=Past members</td></tr>
<tr><td></td><td>Difference</td><td>t-stat.</td></tr>
<tr><td>Credit from informal lenders‡</td><td>–0.17
(–1.52%)</td><td>–0.03</td></tr>
<tr><td rowspan="2">Outcome:</td><td colspan="2">(b) Average Treatment Effect (ATE)
$\gamma_0^{m,l} = E(Q^m - Q^l) = EQ^m - EQ^l$</td></tr>
<tr><td colspan="2">m=Any credit programme; l=Past members
difference</td></tr>
<tr><td>Credit from informal lenders‡</td><td colspan="2">–9.91</td></tr>
</table>

Source: own calculation from FMHFS. ‡Value in MK (Malawian Kwachas); for conversion to US $ see footnote 4 above.

Table 6 reports the results for the average effects: the upper panel displays the estimated average treatment effects on treated (ATT) and the lower panel shows the average effects for the population. The estimated average treatment effect on treated is reported in absolute and relative terms. As Larsson (2000) points out, the absolute size of the effects allows for a comparison between treatment and control households, whereas the relative effects reported in percentage points indicate the size of the effect.[12]

As can be seen from Table 6, in the ‘single treatment’ model, we find no significant evidence of crowding out of informal credit from the credit programmes. This result, however, could be affected by a “treatment aggregation bias”. In other words, because we have pooled all credit programmes in a single treatment we have neglected treatment heterogeneity. We address this (potential) bias in the next subsection.

4.2 Second stage: propensity scores with multiple treatments

As discussed, the insignificant result in Table 6 might arise as a result of bias if there is treatment heterogeneity. There are sound reasons for thinking that there is indeed heterogeneity in this case. This can be illustrated by disaggregating the single treated group (that is, those who participated in any credit programme) into two groups: (i) those who participated only in the MRFC; and (ii) those who participated in more than one credit programme (here defined as the MRFC and at least one other credit programmes).

Table 7 shows a selection of characteristics of borrowers from each of the two treatments and the corresponding Wald test of the difference in means adjusted with survey weights. Whilst there are not many significant differences in demographic characteristics, the two treated groups differ substantially in their assets and expenditure patterns – in particular a highly significant difference in the value of equipment, food and non-food expenditure exists between the two groups. This is evidence that treatment heterogeneity occurs and that the single treatment model

[12] The relative effect has been calculated as percentage change of the average informal borrowing between treated and control households.

suffers from an aggregation bias. Further disaggregation might yield other distinctions but the lack of data points limit the scope for more complex multiple treatments.

Table 7. Characteristics of groups of borrowers

	MRFC only	**MRFC+ 2nd or more programmes**	***Wald test (p-value)***
Household demographics:			
Female headed HHs (%)	23.1 (153)	21.4 (388)	0.88
Average HH size	**5.2 (153)**	**6.2 (388)**	**0.10***
Average number of children 0-15	2.7 (153)	3.2 (388)	0.33
Household head's main occupation: agriculture	53.2 (153)	46.1 (387)	0.65
Household affected by negative shocks (Yes/No)	58.9 (152)	59.7 (383)	0.66
Average land size (ha)	**1.9** (150)	**2.4** (367)	**0.10***
Share of land owned by spouse	23.3 (150)	11.7 (367)	0.15
Average value of equipment (MK)	**193** (153)	**450** (375)	**0.00*****
Average value of non-food expenditure‡ (MK)	**3.1 (153)**	**7.7 (387)**	**0.00*****
Average food expenditure (MK)	**10.7** (150)	**15.5** (367)	**0.00*****

Source: Own calculation based on FMHFS. Note: household types are defined according to participation in the credit programmes. Weighted results. Number of observations in parentheses. Expenditure deflated by the square root of household' size. ‡Non food expenditure on candles, cigar, lighter, fuel, batteries, soap.

We therefore apply the same propensity score matching approach explained in section 4.1 in a multiple treatment setting, where the treatment option $T=m$ now denotes two treatments with $m=1, 2$. $T=1$ is membership only in the MRFC; and $T=2$ is membership in the MRFC and in other formal credit programmes (i.e. participation in at least two credit programmes). $T=l$ denotes the case of no treatment as before.

The multiple treatment model is constructed as before whereby propensity scores are estimated and then used to perform matching with the Mahalanobis distance.[13] The average treatment effects are then calculated for each pairwise combination of the treatments with the control group.

The main difference in the case of multiple treatments is how we handle the estimation of the conditional participation probabilities. These can be either estimated for each combination of treatments as binary choices or they can be modelled with a multinomial model including all relevant choices.[14] Both these models have advantages and disadvantages and although we decided to model the decision process with a series of logit models, we also checked the robustness of our analysis in a multinomial context and found that our results remain unchanged.

For the core analysis we rule out the use of the multinomial logit model because of the violation of the Independence of Irrelevant Alternatives (IIA). According to the IIA, the inclusion or exclusion of some programmes does not alter the relative probability of choosing one programme rather than another. However, as outlined by Larsson (2000), the IIA may not hold in a multiple programmes context because the relative probabilities of one choice, as opposed to another, may change whenever programmes are at least partial substitutes for each other. If we were to use a multiple choice model, the multinomial probit model (MNP) would seem to be the best option because it would not require the IIA assumption. However, the MNP could not converge on our data so we follow Lechner (2002) in using a series of logit models[15].

Bryson et al. (2002) highlight shortcomings in using a series of binary choice models in this context. As the number of treatments increases, the number of binary combinations of treatments also increases rapidly. However, in terms of outcomes, Lechner (2002) find little difference in the performance of the MNP model and the series of binomial models. In particular, the matching quality (measured by

---

[13] The results of the matching algorithm are available from the authors upon request.

[14] Both approaches have advantages and disadvantages and although we decided to model the decision process with a series of logit models, we also checked the robustness of our results in a multinomial logit model (MNL). The results of the MNL are available from the authors upon request.

[15] The non-concavity of the likelihood function causes problems in estimating the likelihood function of the MNP. The results of the series of logit models are available from the authors upon request.

standardised bias) achieved with the MNP is similar to a series of binary choices. The latter approach is in fact more flexible because it allows modelling of each of the binary choices with a different set of covariates. In addition, binary models are more robust to errors since a misspecification in the model of any pair of treatments will not compromise the other binary treatments (Dorsett, 2001).

Table 8. Average treatment effects with Mahalanobis matching (multiple treatment)

Outcome:	**(a) Average Treatment Effect on Treated (ATT)** $\vartheta_0^{m,l} = E[Q^m \mid T=m, P^{m\mid ml}(X)] - E[Q^l \mid T=m, P^{m\mid ml}(X)]$			
	(I) *m*=MRFC; *l*=Past members		**(II) *m*=2nd programme; *l*=Past members**	
	Difference	*t-stat.*	*Difference*	*t-stat.*
Credit from informal lenders‡	**–25.44**	–2.06**	1.29	0.04
	(–74.0%)		*(1.95%)*	
Outcome:	**(b) Average Treatment Effect (ATE)** $\gamma_0^{m,l} = E(Q^m - Q^l) = EQ^m - EQ^l$			
	(I) *m*=MRFC; *l*=Past members		**(II) *m*=2nd programme; *l*=Past members**	
	Difference		*Difference*	
Credit from informal lenders‡	**–23.15**		4.14	

Source: own calculation from FMHFS. **p<0.05; significant results in **bold**.
‡Value in MK (Malawian Kwachas); for conversion to US $ see footnote 4 above.

After the scores are obtained and matching is performed, the average treatment effect for each pair-wise combination of treatments is calculated. Table 8 reports the results for the average effects: the upper panel displays the estimated pair-wise average treatment effects on treated (ATT) and the lower panel shows the pair-wise average effects for the population.

First, we describe the effect of participating only in the MRFC programme on the amount households borrow from informal sources. Compared to past members, those who participate only in the MRFC borrow significantly less from informal lenders. We find that membership of MRFC and no other microfinance programme reduces borrowing from informal sources by 25.5 MK (approximately 2 U.S. dollars at that time). In relative terms, it reduces the amount members borrow from informal lenders by approximately 74 percent. The average treatment effect on the population confirms this result, but with a slightly smaller impact of approximately 23.2 MK (approximately 1.5 U.S. dollars). In contrast, there is no evidence of crowding out when households participate in more than one credit programme.

The size of the crowding out of informal loans we find is larger than in some of the previously-mentioned empirical studies, such as McKernan et al.(2005), who find for any unit of income transfer, there is a decrease in informal transfers of on average 0.25 for women and 0.31 for men.

We next explore the difference between the findings of the single treatment model in Table 6 and the multiple treatment model in Table 8 that could arise from heterogeneity of the treatments or heterogeneity of the treated. Considering the former, is it apparent that all groups, both in the single and multiple treatment model, share a common factor – namely, participation in the MRFC programme. As illustrated in Table 2 and Figure 1, households in some of the sampled districts do not participate in the other programmes – indeed some of these programmes are not available in some districts. Nevertheless it is hard to think of any plausible reason why the combinations of credit programme usage should change the loan characteristics in such a way as to change treatment effects.

Based on evidence reported in Table 7, a more plausible 'story' for the discrepancy in results between the two treatment groups lies in the heterogeneity of the treated. Such an explanation is also in line with the findings of Navajas et al. (2003) and Cox and Jimenez (2005). Table 7 shows that the group of participants that are involved in more than one credit programme are significantly different from participants of the MRFC programme only. The former group are relatively better off on average, in terms of assets, food and non-food expenditure. One interpretation of the results is that participation in more than one credit programme is an indicator of being a relatively less credit or collateral-constrained household. As in Navajas et al. (2003), less capitalized borrowers tend to switch from an informal credit contract to a loan contract when a microfinance institution is made available (i.e. the MRFC) whereas relatively wealthier and unconstrained households may not substitute one source for the other, but simply increase the overall demand for credit once the supply of formal loans increases – this behaviour simply reflecting a greater preference for borrowing by such households. The insignificant effect for the group of households participating in more than one microfinance programme may also be affected by the fact that we pool different types of programmes in both the treated group and the control group. Unfortunately, there are not enough observations to disentangle the effect of each microfinance programme in isolation.

4.3 Third stage: sensitivity analysis

The last stage of the evaluation involves checking the robustness of the results. Three sensitivity analyses are performed: a) changing the model specification and matching algorithm; b) changing the treatment and outcome definition; and c) changing the model used to estimate the propensity scores. Each of these robustness checks is analysed in turn.

a) Changes in the matching algorithm

Table 9 provides sensitivity checks for the two groups of treatments: MRFC participants in the upper panel and multi-programme participants in the lower panel. The last two rows in each panel show the absolute and relative values of the average treatment effects together with the *t*-statistic in parentheses, under alternative assumptions.

In model A, we drop the sampling weights. Heckman and Todd (2009) show that, with nearest neighbour algorithms, it does not matter whether matching has been performed on the odds ratio with or without weights since the ranking of the observations is identical and the same neighbour will be selected. Here, as in Frölich et al. (2004), we find that the most significant results remain largely unchanged – the average treatment effect on treated (ATT) households in the MRFC programme is still negative and significant and both the relative and absolute effects remain almost unchanged.

The last two columns of table 9 report the average treatment effects on treated households obtained after performing two alternative matching algorithms. Model B, nearest neighbour matching, involves finding for each treated household, the control household with the closest propensity score. This procedure is implemented with replacement, that is, while each treated household has only one match, the control household may be matched to more than one treated household. Dehejia and Wahba (2002) found that nearest neighbour with replacement produces better matching.

In order to improve the quality of the match, we have also selected control households within a preset amount (or caliper) of the treated household's estimated propensity score. In other words, the nearest neighbour matching with replacement and caliper imposes an *a priori* common support region. More formally, keeping the same notation as before, for a pre-specified $\delta > 0$, treated household i is matched to untreated household j such that:

$$\delta = \left| P_i^m - P_j^l \right| = \min_{k \in C} \left\{ \left| P_i^m - P_j^l \right| \right\}$$

where P^k, with $k=(m,l)$ are the propensity scores for the two options and C is the set of neighbours of treatment households in the untreated group.

Smith and Todd (2005) point out that a drawback of this algorithm is that it is difficult to determine *a priori* the size of caliper. We set our caliper $\delta = 0.02$ as a result of a maximization in the bias reduction and a minimization of loss of observations.[16] Model B in table 9 confirms the results obtained by Mahalanobis matching for both treatment groups, that is, a significant evidence of crowding out on informal borrowing for participants in the MRFC and no significant effect for the MRFC and other programmes.

To further check the robustness of our results we perform a non-parametric estimator, Kernel matching, as Model C in Table 9. Kernel matching is like a weighted regression where the counterfactual outcome is constructed as a weighted average of all households in the control group. Relative to Model B, nearest neighbour with replacement, an advantage of this approach is that the variance is smaller as a result of the use of more information.

[16] We lose four observations in models (II) and (III) of panel (a) and 45 observations in models (II) and (III) of panel (b).

Table 9. Average effects from Mahalanobis matching: sensitivity tests

Outcome: *Credit from informal lenders*‡	**(a) *m*=MRFC; *l*=Past members**		
	Alternative model specification	**Different matching algorithm**	
	Model (A): no weights	*Model (B): Nearest Neighbour*[1]	*Model (C): Kernel matching*
ATT	**–25.29** *(–2.24)***	**–21.46** *(–1.98)***	**–20.14** *(–2.38)***
% points	**–73.9%**	**–71.6%**	**–70.3%**
Outcome: *Credit from informal lenders*‡	**(b) *m*=MRFC+2nd programme; *l*=Past members**		
	Alternative model specification	**Different matching algorithm**	
	Model (A): No weights	*Model (B): Nearest Neighbour*[1]	*Model (C): Kernel matching*
ATT	–4.19 *(–0.13)*	31.23 *(0.83)*	31.00 *(0.84)*
% points	*–5.85%*	*71.3%*	*70.3%*

Source: own calculation from FMHFS. t-stats in parentheses; significant results in **bold**. **p<0.05.
[1]Nearest Neighbour has been performed with caliper and replacement.
‡Value in MK (Malawian Kwachas); for conversion to US $ see footnote 4 above).

The application of the Kernel algorithm involves the choice of the Kernel function and of the bandwidth. DiNardo and Tobias (2001) show that the choice of Kernel does not greatly affect the results. We have used a standard Epanechnikov Kernel. As shown by Silverman (1986) and Pagan and Ullah (1999), the choice of bandwidth involves a trade-off between bias and variability. A large bandwidth decreases the variance by providing a better fit with a smoother density function. On the other hand, as the bandwidth increases the bias increases as well. We set the bandwidth to be equal to the caliper size in model (B). Once again, as Model C demonstrates in Table 9, the average treatment effect on treated households in the MRFC programme is negative and significant. The absolute effect is slightly smaller with a value of 1.3 U.S. dollars (approximately 20 MK). However, the ATT in panel (b) with the use of multiple credit institutions is again positive but insignificant.

b) Changes in treatment and outcome definitions

In this section, we change the definition of treatment. Previously, we estimated the effect of being a member of a microfinance programme on the amount households borrow from informal lenders. Now, we examine what happens if we apply a stricter definition of treatments, that is, if we define a treated household to be both a member and borrower of a microfinance programme or programmes.

In order to answer this question, we repeat the three stages of the evaluation procedure with the new definition of treatments.[17] Table 10 reports the average effects for the two groups of the newly defined treatments. Although the magnitude of the ATT for the MRFC treatment is similar to those described in Tables 8 and 9 – the crowding out effect is approximately 1.6 U.S. dollars, the effect has decreased.

In addition, we change the outcome variable. We now know that participation in the MRFC programme reduces the amount borrowed from informal lenders. But does this happen because households demand less or because informal lenders give them less credit (or both)? This ambiguity arises from the fact that demand and supply issues cannot be disentangled by simply looking at the amount borrowed from informal lenders. The second row of table 10 therefore looks at whether crowding out applies also to the self-reported demand for credit from informal lenders. The logit models used are the same as the ones for the multiple treatment model[18]. We find a large and significant reduction in the demand for informal finance for households who participate in the MRFC (−75.2 percent). Again, however, there is no evidence of crowding out for households who participate in more than one credit programme (panel (II)).

The third row of table 10 disentangles the supply from the demand of informal loans by looking at the self-reported credit limit. The credit limit variable is the maximum amount that the borrower thinks the lender is willing (or able) to lend, and can be thought to be the limit on the "supply" of informal loans. This approach allows us to test whether transfers from informal lenders are crowded out by the introduction of microfinance programmes. Although the coefficient on this version of the 'crowding out' hypothesis is similar to the other specifications of Model I, it is not quite significant. Again, the sensitivity analysis does not find a

[17] The results of the logit models can be obtained from the authors upon request.

[18] These are also available from the authors upon request.

significant coefficient for Model II. However, caution is needed when treating credit limits directly as the supply function of credit to the individual or household since they are self-reported values.

Table 10. Sensitivity analysis of ATT to changes in treatment and outcome definitions

Outcome:	**(a) Average Treatment Effect on Treated (ATT)** $\vartheta_0^{m,l} = E[Q^m \mid T = m, P^{m\mid ml}(X)] - E[Q^l \mid T = m, P^{m\mid ml}(X$			
	(I)m=MRFC; l=Past members		**(II)m=2nd programme; l=Past members**	
	Difference	*t-stat.*	*Difference*	*t-stat.*
Credit from informal lenders[‡]	**–24.34**	*–1.75**	39.07	*0.68*
	(–70.6%)		*(54.9%)*	
Demand from informal lenders[‡]	**–29.75**	*–2.30***	4.45	*0.14*
	(–75.2%)		*(6.2%)*	
Credit limit from informal lenders[‡]	–67.62	*–1.55*	28.02	*0.31*
	(–69.1%)		*(19.5%)*	

Outcome:	**(b) Average Treatment Effect (ATE)** $\gamma_0^{m,l} = E(Q^m - Q^l) = EQ^m - EQ^l$	
	(I)m=MRFC; l=Past members *Difference*	**(II)m=2nd programme; l=Past members** *Difference*
Credit from informal lenders[‡]	–13.54	60.87
Demand from informal lenders[‡]	–27.92	6.35
Credit limit from informal lenders[‡]	–76.33	12.05

*Source: own calculation from FMHFS. t-stats in parentheses. **p<0.05, *p<0.1. [‡]Value in MK (Malawian Kwachas); for conversion to US $ see footnote 4 above).*

5. Conclusion

The role of microfinance institutions in markets where there are other informal lenders is relevant for examining appropriate policy interventions. It is established in the literature on developing countries that informal lending offers partial, but incomplete, insurance against income shocks, but that large scale ‘modern’ banking institutions rarely reach small-scale rural borrowers. A proposition is that governments wanting to reach small borrowers could create new lending institutions that mimic the features of informal lending arrangements by, for example, adopting joint liability schemes that enable borrowers to select safe fellow group members so to avoid the risk of default. The empirical question for policy-makers is whether these credit institutions impact on households' access to informal sources, or just serve a different segment of households leaving competition in the credit market unchanged?

This paper addresses the question of whether microfinance institutions crowd out informal credit using data on several credit programmes introduced in Malawi in the mid-1990s. By using evaluation techniques, the paper joins a relatively small group of recent empirical papers which demonstrates some evidence of 'crowding out' of local informal lending.

After showing that the standard single treatment model can provide biased results when treatment heterogeneity is not taken into account, the paper adopts a "multiple treatment" framework. Borrowers are differentiated into those who borrow only from the most geographically extensive microfinance institution, the Malawi Rural Finance Company (MRFC), and those who borrow from multiple sources. We find strong evidence of significant crowding out of informal credit among those who borrow only from the MRFC, but little evidence of crowding out for those who borrow from several institutions. We suggest that the latter result could arise from heterogeneity of the treatment or heterogeneity of the treated, but argue that the latter is most likely the case, since multiple borrowers tend to be better off in terms of expenditure levels and access to collateral (such as land and housing), notwithstanding ceilings on access to credit from these institutions determined by farm size.

The paper demonstrates that most informal lending among survey respondents in Malawi is obtained from friends and families rather than from moneylenders. Assuming that 'crowding out' occurs evenly across types of informal lending (a proposition that cannot be tested given the relatively small sample size), this has policy implications. The introduction of microfinance institutions may indeed drive out predatory moneylenders, but some displacement of traditional sources of credit within the extended family or among informal institutions such as rotating credit associations may occur. Whether this is desirable depends on whether such informal credit channels are seen as desirable in themselves, or are exposed to the various non-diversifiable risks that make insurance against income risk incomplete in rural areas.

Finally the paper subjects the key results to a battery of sensitivity analyses. We test the crowding out hypothesis using several treatments and outcome definitions as well as adopting alternative matching strategies. These analyses do not change our key conclusions. A contribution to existing studies focusing on realised transfers (borrowing) rather than potential transfers (membership of group lending institutions) is that we have identified and quantified a more significant crowding out effect on informal borrowing of membership of microfinance programmes.

References

P. Albarran and O. P. Attanasio. Do private transfers crowd out private transfers? Evidence from a randomized experiment in Mexico. Technical report, Discussion Paper N. 2002-6. World Institute for Development Economics, Helsinki, Finland, 2002.

B. Armendariz and J. Morduch. The Economics of Microfinance. The MIT Press Cambridge, Massachusetts and London, England, 2005.

O. Ashenfelter. Estimating the effect of training program on earnings. Review of Economics and Statistics, 60:47–57, 1978.

O. P. Attanasio and J. V. Rios-Rull. Consumption smoothing and extended families. Technical report, University College London, mimeo, 2000.

P.C. Austin and M. M. Mandami. A comparison of propensity score methods: a case-study estimating the effectiveness of post-ami statin use. Statistics in Medicine, 25 (12): 2084–2106, 2006.

T. Besley and S. Coate. Group lending, repayment incentives and social collateral. Journal of Development Economics, 46:1–18, 1995.

T. Brodaty, B. Crepon, and D. Fougere. Using matching estimators to evaluate alternative youth employment programs: evidence from France, 1986-1988. In M. Lechner and M. Pfeiffer (eds) Econometric evaluation of labour market policies. Heidelberg: Physica-Vertlag, 2001.

A. Bryson, R. Dorsett, and S. Purdon. The use of propensity score matching in the evaluation of active labour market policies. Technical report, Policy Studies Institute and National Centre for Social Research Working Paper No.4, 2002.

D. Cox, Z. Eser, and E. Jimenez. Motives for private transfers over the life cycle: an analytical framework and evidence for Peru. Journal of Development Economics, 55: 57–80, 1998.

D. Cox and M. Fafchamps. Extended family and kinship networks: economic insights and evolutionary directions. In T.P. Schulz (ed) Handbook of Development Economics Vol. 4. Elsevier B. V., 2008.

D. Cox, B. E. Hansen, and E. Jimenez. How responsive are private transfers to income? Evidence from a laissez-faire economy. Journal of Public Economics, 88: 2193–2219, 2004.

D. Cox and E. Jimenez. Social security and private transfers in developing countries: the case of Peru. The World Bank Economic Review, 6(1):155–169, 1992.

R. B. Jr. D'Agostino. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Statistics in Medicine, 17: 2265–2281, 1998.

A. Deaton. Household saving in LDCs: credit markets, insurance and welfare. Scandinavian Journal of Economics, 94: 253-273, 1992a.

A. Deaton. Saving and income smoothing in Cöte d'Ivoire. Journal of African Economies, 1: 1-24, 1992b

R. H. Dehejia and S. Wahba. Propensity score matching methods for non-experimental causal studies. The Review of Economics and Statistics, 84 (1):151–161, 2002.

A. Diagne. Determinants of household access to and participation in formal and informal credit markets in Malawi. Food Consumption and Nutrition Division Discussion Paper No. 67. International Food Policy Research Institute, Wasington DC. 1999.

J. J. Diaz and S. Handa. An assessment of propensity score matching as a non-experimental impact estimator. Evidence from Mexico's PROGRESA program. Journal of Human Resources, 41(2):319–345, 2006.

J. DiNardo and J. Tobias. Nonparametric density and regression estimation. Journal of Economic Perspectives, 15 (4):11–28, 2001.

R. Dorsett. The new deal for young people: relative effectiveness of the options in reducing male unemployment. Technical report, PSI Discussion Paper N. 7, 2001.

P. Dubois, B. Jullien and T. Magnac Formal and informal risk sharing in LDCS: Theory and empirical evidence. Econometrica. 76: July, 679-725, 2008.

B. Fitzenberger and H. Prey. Evaluating public sector sponsored training in East Germany. Oxford Economic Papers, 52:497–520, 2000.

M. Frölich, A. Heshmati, and M. Lechner. A microeconometric evaluation of rehabilitation of long-term sickness in Sweden. Journal of Applied Econometrics, 19:375–396, 2004.

M. Ghatak. Group lending, local information and peer selection. Journal of Development Economics, 60:27–50, 1999.

J. J. Heckman, H. Ichimura, and P. E. Todd. Matching as an econometric evaluation estimator: evidence from evaluating a job training program. Review of Economic Studies, 64: 605–654., 1997.

J. J. Heckman, R. J. LaLonde, and J. A. Smith. The economics and econometrics of active labor market programs. In O. Ashenfelter and D. Card (eds) Handbook of Labor Economics, Vol.3A, 1865–2097. Amsterdam: Elsevier Science, 1999.

J. J. Heckman and J. A. Smith. The pre-program earnings dip and the determinants of participation in a social program: implications for simple program evaluation strategies. Economic Journal, 108:313–348, 1999.

J. J. Heckman and P.E. Todd. A note on adapting propensity score matching and selection models to choice based samples. Journal of Econometrics, 12, 1: S230-S234, 2009.

G. W. Imbens. The role of propensity score in estimating dose-response functions. Biometrika, 87 (3):706–710, 2000.

R. T. Jensen. Do private transfers displace the benefits of public transfers? Evidence from South Africa. Journal of Public Economics, 88:89–112, 2003.

J. P. Kaboski and R. M Townsend. The impacts of credit on village economies. Technical report, mimeo University of Chicago, 2006.

L. Larsson. Evaluation of Swedish youth labour market programmes. Technical report, Discussion paper 2000(1) Office for Labour Market Policy Evaluation, Uppsala, 2000.

M. Lechner. Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. Technical report, Discussion paper 9908, University of St. Gallen, 1999a.

M. Lechner. Earnings and employment effects of continuous off-the-job training in East Germany after unification. Journal of Business Economics and Statistics, 17:74–90, 1999b.

M. Lechner. An evaluation of public sector sponsored continuous vocational training programs in East Germany. Journal of Human Resources, 35:347–375, 2000.

M. Lechner. Programme heterogeneity and propensity score matching: an application to the evaluation of active labour market policies. Review of Economics and Statistics, 84, 2: 205-220, 2002..

A. Manca and P.C. Austin. Using propensity score methods to analyse individual patient level cost effectiveness data from observational studies. Technical report, Health, Econometrics and Data Group, The University of York, Working Paper 08/20, 2008.

C. F. Manski and S. R. Lerman. The estimation of choice probabilities from choice based samples. Econometrica, 45 (8):1977–1988, 1977.

S. M. McKernan, M. Pitt, and M. Moskowitz. Use of the formal and informal financial sectors: does gender matter? Empirical Evidence from rural Bangladesh. Technical report, World Bank Policy Research Working Paper, 3491, 2005.

J. Morduch. Does microfinance really help the poor? New evidence from flagship programs in Bangladesh. Technical report, Unpublished Working paper. Available at http://www.wws.princeton.edu/ rpds/macarthur/downloads/avgimp 6.pdf, 1998.

J. Morduch. The microfinance schism. World Development, 28 (4):617–629, 2000.

S. Navajas, J. Conning, and C. Gonzalez-Vega. Lending technologies, competition and consolidation in the market for microfinance in Bolivia. Journal of International Development, 15:747–770, 2003.

A. Pagan and A.Ullah. Nonparametric econometrics. Cambridge University Press, Cam- bridge, 1999.

M. Pitt and S. Khandker. The impact of group-based credit programs on poor households in Bangladesh: Does the gender of participants matter? Journal of Political Economy, 106. N. 5:958–995, 1998.

D. Ray. Development economics. Princeton University Press, 1997.

P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. Biometrika, 70:41–50, 1983.

M. Rosenzweig and K. Wolpin. Credit market constraints, consumption smoothing and the accumulation of durable production assets in low-income countries. Journal of Political Economy, 101 (2):223–244, 1993.

A. D. Roy. Some thoughts on the distribution of earnings. Oxford Economic Papers, 3:135–146, 1951.

D. B. Rubin. Estimating causal effects of treatments in randomized and non-randomized studies. Journal of Educational Psychology, 66:688–701, 1974.

B. Silverman. Density estimation for statistics and data analysis. Chapman and Hall, London, 1986.

J. Smith and P.E. Todd. Does matching overcome Llalonde's critique of non-experimental estimators? Journal of Econometrics, 125(1-2):305–353, 2005.

J. Stiglitz. Peer monitoring and credit markets. World Bank Economic Review, 4 (3):351–366, 1990.

C. Udry. Risk and insurance in a rural credit market: An empirical investigation in northern Nigeria. Review of Economic Studies, 61, 495-52, 1994.

B. Wydick. The effect of microenterprise lending on child schooling in Guatemala. Economic Development and Cultural Change, 47 (4):853–869, 1999.