Targeting Credit through Community Members

Department of Research and
Chief Economist

Diego A. Vera-Cossio

DISCUSSION
PAPER Nº
IDB-DP-609

August 2018

Targeting Credit through Community Members

Diego A. Vera-Cossio

Inter-American Development Bank

August 2018

Targeting credit through community members

Diego A. Vera-Cossio *

Inter-American Development Bank

August 6, 2018

Abstract

Decentralizing the allocation of resources to community members is an increasingly popular form of delivering public resources in developing countries. However, this approach is associated with the tradeoff between improved information about potential beneficiaries and favoritism towards local elites, which could be strengthened in the context of credit. Unlike targeting cash transfers at poor households, allocating publicly-provided credit is a more complex problem involving issues of risk, neediness, productivity, and market responses: This paper analyzes this problem using a large-scale lending program, the Thai Million Baht Credit Fund, which decentralizes the allocation of loans to an elected group of community members, and provides three main results. First, exploiting a long and detailed panel, I recover pre-program structural estimates of household productivity and find that neither repayment, nor poverty, nor productivity explains program participation. Second, using socioeconomic networks data, I show that actual targeting is strongly driven by connections to village elites and is related to lower program profitability, which suggests favoritism as a reason for mistargeting. Finally, I exploit quasi-experimental variation in the rollout of the program and uncover evidence that, in general equilibrium, informal credit markets partially compensate for targeting distortions by redirecting credit towards unconnected households, albeit at higher interest rates than those provided by the program. The results highlight the limitations of community-driven approaches to program delivery and the role of markets in attenuating potential targeting errors.

Keywords: Microcredit, decentralization, entrepreneurship, targeting.

JEL: D14, G21, O12, O16, O17, L14, Z13

*I would like to thank Prashant Bharadwaj, Gordon Dahl, Craig McIntosh, Karthik Muralidharan, Krislert Samphantharak and Robert Townsend for their support and guidance during this project. I would also like to thank Mauricio Romero, Claudio Labanca, Desmond Ang, and attendees at the UCSD Applied Economics seminar series,PACDEV 2018 and ABCDE 2018 for their insightful comments. All errors are my own. E-mail: diegove@iadb.org

1 Introduction

Community-driven approaches to delivering public resources are increasingly popular in developing countries. A number of policy efforts such as public works or cash transfer programs rely on community members for their implementation, monitoring and targeting (Casey, 2018; Mansuri and Rao, 2004; Alatas et al., 2012). One of the foundations of this approach is the idea that community members, as opposed to traditional policy makers, may have better information to identify local needs. In the context of credit, delegating the allocation of loans to community members could lead to more accurate identification of potential borrowers, and the delivery of capital to those who would benefit the most: poor, high-productivity households.

One important class of credit expansion policies is that of government infusions of resources into villages for the establishment of community-managed local credit funds.[1] The economic rationales for this approach include lower implementation costs as well as the benefit from information available to community members, which is costly to obtain by banks or government officers. In contrast, community-based approaches are prone to favoritism (Bardhan and Mookherjee, 2005). This tension is particularly salient in the case of publicly funded credit programs: Soft information provided by peers may improve borrower screening Iyer et al. (2016) and can be used to identify high-return entrepreneurs Hussam et al. (2017), but relying on information that is hard to measure may prevent accountability. Thus, whether the allocation of credit is consistent with risk, poverty, productive efficiency or favoritism as targeting criteria is an empirical question.

While previous studies have analyzed this question in similar settings, such as production kits and loans (Bardhan and Mookherjee, 2006b), and fertilizer subsidies (Basurto et al., 2017), their analysis is based on contemporary or post-program proxies for returns which are likely to be affected by the program itself. Moreover, the analysis of different targeting schemes, in a variety of settings, have mostly focused their empirical analysis on understanding the direct delivery of resources to the target population (Alatas et al., 2012, 2016;

[1] Examples of this include the Million Baht program in Thailand (Kaboski and Townsend, 2012) and the Integrated Rural Development Program in India (Bardhan and Mookherjee, 2006b), and the Rural Financial Institutions Programme in Uttar Pradesh, India. Additionally, self-help groups are prominent in developing countries (see Deininger (2013); Greaney et al. (2016); Ksoll et al. (2016); Karlan et al. (2017), among others).

Coleman, 2006; Stoeffler et al., 2016; Karlan and Thuysbaert, 2016), but have ignored the role of markets in reallocating resources, which could attenuate potential targeting errors.

This paper empirically assesses these isues in the context of one of the largest credit-expansion programs, the Thai Million Baht Village Fund (MBVF). Between 2001 and 2002, the government donated resources to over 90% of rural villages for the creation of local credit funds, which represented, on average, a 25% increase in the available funds for credit in each village. These funds were fully managed by elected village committees made up of community members, who decided who obtained credit and under what loan conditions.[2] The context of the MBVF program coincides with the availability of The Townsed-Thai Monthly Survey (Townsend, 2014), which includes comprehensive information regarding household characteristics and financial activities corresponding to three pre-program years as well as several post-program waves.

This paper reports results from three empirical exercises: first, I exploit 14 years of panel data to estimate a household production function following Blundell and Bond (2000)'s approach. I then use the estimated factor elasticities to recover *pre-program* estimates of household total factor productivity.[3] I combine these estimates with baseline repayment and per-capita consumption data to test: *(i)* whether baseline repayment, poverty and productivity drive selection into the program, and *(ii)* whether these characteristics differentially predict eligibility under a hypothetical targeting criterion, which is based on a repayment-score model. Second, I combine detailed data on pre-program socioeconomic networks with data about ex-post returns to the lender to test for favoritism towards elite-connected households. Third, I use quasi-experimental variation in the rollout of the program to test for within-village redistribution through informal credit markets, which could lead to program spillovers to households with limited access to credit from the program.

First, I find that program participation was neither explained by credit history, nor poverty, nor productivity. Among households with pre-program credit history, committee

[2]Kaboski and Townsend (2012) documents increases in income growth and consumption due to the program. Despite positive impacts Kaboski and Townsend (2011) argue that a counter-factual cash transfer program would have been more cost-effective.

[3]Concretely, I exploit data on households' financial statements, in particular balance sheets, to measure capital as the value of the stock of total fixed assets for each household. The financial accounts data was compiled by Samphantharak and Townsend (2010).

members did not penalize households who had missed payments or required loan extensions in the past. Among the entire pool of potential borrowers, I find that baseline per-capita consumption is positively correlated with program participation, suggesting that poverty was not the main targeting criterion for program committee members. In addition, I find that baseline TFP is not correlated with program participation, which is suggests misallocation. In contrast, households of Village Council members–i.e., the village government– and households with direct connections to council members were 30 and 16 percentage points more likely to obtained program resources than unconnected households. The results are robust to controlling for health and production shocks that could trigger demand for credit.

While understanding program selection is intrinsically important, it is insightful to analyze the extent to which program borrowers differ from the pool of households that would have been eligible under an alternative policy-relevant way of delivering credit, namely a credit score criterion.[4] Exploiting pre-program data from 3,800 loans and household financial and demographic characteristics, I estimate a repayment probability model, which I then use to identify the set of households that would have been eligible for a loan, holding the village coverage constant. I find that 34% of program borrowers wouldn't be eligible by the repayment-based criterion (over-inclusion error), and that 47% of households who would have been eligible by the repayment-based criterion ended up not borrowing from the program (over-exclusion). Elite-connectedness was not correlated with repayment-based eligibility, and relative to committee members, a scoring model would have targeted less risky and more productive households, suggesting gains from re-allocation. One explanation is that a scoring model, while imperfect, provides an objective rule to allocate resources which is less vulnerable to elite influence.

Second, I analyze two non-exclusive mechanisms through which connections may explain program participation: information and favoritism. I find that while elite-connected households were not poorer or more productive relative to unconnected households, they were more likely to have a history of missed payments. One interpretation, is that while, on average, elite-connected households were not better borrowers, their location in the net-

[4]The state-owned Bank of Agriculture and Agricultural Cooperatives BAAC operates in all the sample villages.

work allowed them to better transmit information to program committee members. After controlling for total number of connections in the village, the correlation between program participation and connection to local elites falls sharply. While it is no longer significant in the case of households with business links to elite households, Village Council members are still 20 percentage points more likely to obtain program credit, which raises concerns regarding favoritism.

I find evidence of favoritism towards connected households with implications for program profitability. While both, lower monitoring costs associated to elite-connected households and favoritism would predict higher program participation for connected households, favoritism should be costly for the program. A cross-section sample of loans corresponding to 335 households who borrowed both from the program and member-funded local credit groups allows me to compare loan returns across different lenders for the same household and control for unobserved borrower characteristics.[5] I test for favoritism by analyzing whether loans to connected households yield lower returns to the program compared to loans from private credit groups and comparing these differences to those for unconnected households. The results show that the *ex post* internal rate of return on program loans to connected households was 2.7 percentage points lower than the return on privately funded loans (on average 7%), relative to similar comparisons in the case of unconnected households. These results are a consequence of committee members favoring elite-connected households with low initial interest rates and larger loans for a similar level of risk.

Third, while committee members favored connected households and the program might not have directly reached unconnected households, the program's rollout indirectly benefited unconnected households through informal credit markets. Using high-frequency data, I exploit cross-village variation in the monthly rollout of the program to identify the short-term effects of the program on credit use for unconnected households. Event-study estimates reveal that borrowing from informal lenders increased by 30% in the case of unconnected households. These loans were mostly from relatives, albeit at an average annual interest rate of 14%, twice as high than that of program loans. Overall, spillovers mildly offset the dif-

[5]These groups constitute quasi-formal sources of credit. They include production credit groups and women's, groups among others. See Kaboski and Townsend (2005) for an in-depth assessment of these type of lenders.

ference in program borrowing between connected and unconnected households: back of the envelop calculations suggest that these effects only account for 10% of program-borrowing gap between connected and unconnected households.

This paper makes two contributions to the literature studying community-based approaches to distributing public resources. First, it highlights the limitations of these approaches when attributes of program beneficiaries are not easily observable by most community members. Unlike the context of poverty targeting, in the context of credit, soft information may only be observable by direct economic interactions, strengthening the tension between information and favoritism. The results from this paper are at odds with studies that find little or no influence of connections with the local elites in the context of cash transfer programs (Alatas et al., 2012; Galasso and Ravallion, 2005). However, they are well-aligned with evidence of rent-seeking behavior in financial markets in the context of banks and firms (Khwaja and Mian, 2005; Haselmann et al., 2017). In addition, this paper shows that the accurate use of information may depend on social connections. While community members can identify high-return entrepreneurs Hussam et al. (2017), in practice, both information transmission costsand favoritism can impose higher program-participation costs to households without the relevant connections, with consequences for poverty targeting, productive efficiency, and program sustainability. These losses should be considered whenever policy makers choose among alternative approaches to program delivery.

Second, by studying a context in which active credit markets interact with the implementation of a large-scale program, this paper examines the targeting problem both directly through the program, and indirectly through credit markets. The literature has generally focused only on the targeting or screening process. This paper expands the analysis beyond the program and tests the consequences of the *de facto* targeting criterion on village credit markets. By providing novel evidence on the role of informal credit markets in attenuating targeting errors, this paper contributes to the literature documenting general equilibrium effects and spillovers from large-scale programs (Angelucci and De Giorgi, 2009; Muralidharan et al., 2017; Kaboski and Townsend, 2012). In particular, the results show that economic connections and political economy factors can affect not only the distribution of public resources in the village economy, but also the redistribution of these resources through markets

(Kinnan and Townsend, 2012; Acemoglu, 2010). Overall, the results suggest that a complete understanding of targeting problems should involve an analysis of how resources are redistributed across agents.

The results from this paper also build on the literature studying the introduction of micro-credit products in developing countries. A core concern in the development economics literature is that of delivering affordable credit to poor, high-productivity households in order to enable them to escape from poverty traps (Banerjee and Duflo, 2010; Morduch, 1999). While the literature has mostly focused on studying the effects of the introduction of credit products on several household outcomes (Kaboski and Townsend, 2012; Banerjee et al., 2015; Crépon et al., 2015) empirical assessments of different methods for targeting credit is rather scarce. This paper provides one explanation for the lack of transformative effects from micro-credit interventions (Banerjee et al., 2015): targeting. The results of this paper are consistent with evidence of heterogeneity in the returns to credit based on household productivity in the context of the MBVF program (Breza et al., 2018). Finally, a comparison of the results from this study with those from studies analyzing selection into credit highlights the importance of different screening mechanisms in credit markets. For instance, Beaman et al. (2014) show that high-return households select into credit in a context in which the screening mechanism is price.[6] This paper documents a less efficient result in a context in which the *de facto* screening mechanisms are social connections with local elites.

2 Context

2.1 The Million Baht Village Fund program

I exploit the context of the Million Baht Village Fund (MBVF) program in Thailand as a laboratory to evaluate the extent to which community members target loans at poor, high-productivity, low-risk households. The program was one of the largest credit-expansion programs of its kind: It delivered resources to over 77,000 Thai villages for the establishment

[6] They do so in the context of a micro-credit program in Mali, managed by an NGO with no government intervention at all.

of village revolving loan funds, and by 2004 its gorss lending portfolio exceeded USD 3 billion (Haughton et al., 2014).

The central government donated THB 1 million (USD 22,500 in 1999 values) to each participating rural and peri-urban villages.[7] The program funds were used as seed capital for the creation of village credit funds which would provide loans to community members. Any villager was allowed to apply for a loan,[8] and borrowers were expected to repay with interest. Once a loan was repaid, both the principal and revenues from interest were re-invested in the village funds and were allocated to other local borrowers. In this sense, the program could be interpreted as a sustained increase in the supply of credit in the village economies.

The program represented a large unexpected increase in the supply of credit in the village financial system. For instance, it was announced following a change in government in January 2001 and was initially rolled out in June 2001. On average, the village gross lending portfolio increased by 24% in the sample villages during the year following the program rollout, and the program was able to reach 62% of households in the study sample during its first two years of operation. While the program lead to increases in income growth and consumption (Kaboski and Townsend, 2012), it is still unclear whether the programs' community-based organization contributed to a better targeting of the resources.

A unique feature of the program is its community-based management. Each village elected a village fund committee (VFC) made up of 9-12 community members. Committee members received a small compensation for their services and were elected for a two-year term in community meetings, however most of them continued in the position for several years (Haughton et al., 2014). The committee was responsible for evaluating loan applications and monitoring repayment. Committee members generally met once or twice a year to review loan applications and authorize disbursements which were typically deposited in bank accounts in the state-owned Bank of Agriculture and Agricultural Cooperatives (BAAC). No specific

[7] Around 95% of all Thai villages participated in the program, including all villages in the study sample. A detailed discussion of the application and disbursement processes is provided by Kaboski and Townsend (2012), Boonperm et al. (2013), Menkhoff and Rungruxsirivorn (2011) and Haughton et al. (2014).

[8] In order to apply, households were required to purchase a share of the fund, at a very low costs which was mainly symbolic.

training was provided to committee members.[9]

The program offered individual liability loans, which did not required a collateral but required one or two cosigners. The program delivered medium size loans at an average annual interest rate of 7% with an average repayment period of 12 months.[10] With respect to other pre-existing sources of credit in the villages, program loans exhibited the lowest interest rate in the market: the second lowest interest rate corresponded to bank loans were (11% per annum, see Table 1). In terms of maturity, the program offered loan terms which were similar to quasi-formal lenders,[11] shorter than banks, but longer than informal lenders.

A crucial concern for the central government was the sustainability of the program. A set of incentives for sustainable management and sanctions in case of mismanagement were established. For instance, if repayment was high, villages would be rewarded with further infusions of resources. In contrast, if default was high, the government would sanction the village by cutting other transfers or funding of other programs. However, there were no direct incentives or sanctions to committee members. While committee decisions were subject to a set of restrictions regarding loan size and term,[12] committee members had full discretion to approve or deny applications and to set interest rates.

2.2 The program and the local political elite

The program was implemented in villages with well-established local political elites. Each village is governed by a Village Council (head and a group of advisers), hereinafter the "local elite". Village Council members are generally elected by villagers, appointed by district authorities, and usually serve in office until retirement.[13] The Village Council represents the

[9] Committee members were supposed to be educated and well known in the community, but the elections were conducted in rural settings in which the majority of household heads had barely completed primary education.

[10] Average loan size is TBH 15,000 (approximately USD 450) which represents roughly 25% of total household annual income.

[11] Quasi-formal institutions include organizations that have a set of procedures for recording their operations, but do not have a physical location. Examples of these are production credit groups (PCGs), women's groups and other village saving and loan associations. See Kaboski and Townsend (2005) for a detailed description of these quasi-formal organizations in the Thai context.

[12] Loans could not exceed THB 20,000, a positive interest rate had to be imposed on all loans, and loan terms could not exceed one year.

[13] This was the case during the study period. However, a reform in 2011 established 5 year terms, but allowed Village Heads to run for reelection.

main link between community members and higher-level authorities. For instance, village council members attend district meetings, collect resources from villagers for religious celebrations or public works, and oversee resolution of disputes between villagers (Moerman, 1969; Mabry, 1979). In the study sample, while households of Village Council members do not seem more educated than their fellow villagers (on average, both groups have 5 years of schooling), they are richer and hold almost twice as much land than their fellow villagers.

While the village fund committee was *de jure* an independent entity, it is possible that the local elite, had enough *de facto* authority to influence committee decisions. Although the election of village fund committee members is intended to induce accountability in the allocation of loans, committee members may have incentives to favor their political supporters or households with connections to the local elite. For instance, when elections could not take place, the committee members were appointed by the Village Head.[14] The local elites could indirectly influence committee members through their economic or family connections: On average, 46% of households in the sample reported transacting with village council members during the two years preceding the program and 13% were firs-degree relatives of elite members. In such a context, the potential gains in information from decentralizing the allocation of resources to community members could be undermined by rent-seeking behavior (Bardhan and Mookherjee, 2005).

3 Theoretical framework

The program's stated objectives were to expand access to institutional credit, promote career development and income generation (Government of Thailand, 2004), which suggest that poverty, productivity and repayment are key dimensions to understand program participation. For instance, access to institutional credit was low among poorer households,[15] the government claimed publicly that resources were allocated to productive activities (Pasuk and Baker, 2004), and program sustainability relies on repayment. However, there were not clear program guidelines regarding how these criteria would be balanced. In this section, I

[14] Haughton et al. (2014) document that 15% of village fund committee members were appointed directly by either the Village Head or the Village Council

[15] Per-capita consumption was 16% lower for households without access to institutional credit at baseline.

first present a simple theoretical framework that characterizes the optimal allocation of loans by the Village Fund Committee. Second, I show that although the committee has incentives to target high-return households, concerns regarding repayment and the committee's own preferences towards certain households may limit its ability to deliver credit to the households that would benefit the most. Finally, I compare the optimal committee's allocations to the benchmark allocation that would be achieved by a profit-maximizing bank in charge of managing the credit funds.

3.1 The Village Fund Committee's allocation

Committee members decide the amount of credit (b_i) that each of their N_v fellow villagers obtains from the program. They do so in order to maximize a weighted sum of utilities corresponding to all the community members, subject to a sustainability constraint:

$$\max_{\{b_1,\ldots,b_{N_v}\}} \sum_{i=1}^{i=N_v} \psi_i V(b_i)$$

$$\text{s.t.}$$

$$\sum_{i=1}^{i=N_v} b_i \leq \sum_{i=1}^{i=N_v} q_i(1+r)b_i \tag{1}$$

Political favoritism, social norms, and preferences or non-pecuniary costs may determine the weights associated to each village member (ψ_i), which are exogeneous with respect to the allocation problem. V_i denotes household i's indirect utility function–i.e., the value function from the corresponding household optimization problem– which is increasing and concave in b_i (see Appendix B for a micro foundation).

For simplicity, suppose that households repay their loans with an exogenous probability q_i which is known to the committee, and that loans are provided at a government-imposed interest rate r. Thus, the sustainability constraint (1) dictates that the expected revenues from lending should be at least as high as the total amount lent by the village funds. The optimal allocation satisfies:

$$\left(\frac{\psi_i}{1-q_i(1+r)}\right)\frac{\partial V_i}{\partial b_i}=\left(\frac{\psi_j}{1-q_j(1+r)}\right)\frac{\partial V_j}{\partial b_j} \tag{2}$$

$$\tag{3}$$

In words, MBVF committee members will allocate resources such that the weighted marginal utilities from receiving extra-liquidity are equal across all villagers.

How do committee members balance risk, returns and weights? First, consider the case in which both repayment probability and weights are constant across villagers. In such a case, given the concavity of V, it would be optimal to provide more credit to households with higher marginal returns. Moreover, I argue that it would be optimal for the committee members to target poor, high-productivity households as they may derive higher marginal utility from relaxing their budget constraints (Breza et al., 2018). Appendix section B illustrates this argument more formally.[16] Intuitively, poorer and productive households may not be able to self-finance profitable projects and would benefit substantially from additional credit.

Second, heterogeneity in the probability of repayment (q_i) may generate a trade-off for committee members. For instance, the village committee would be willing to give up higher marginal returns ($\frac{\partial V_i}{\partial b_i}$) in order to obtain higher repayment rates. Thus, the delivery of program credit to high-return households will depend on the extent to which high-return households are also low-risk households. Finally, an extra tension may arise when committee members weight households differently. Suppose household i has higher marginal utility, but lower weights than j. Because i's utility is poorly weighted in the objective function, it is optimal for the committee to allocate more loans to j as the weighted marginal utility may still be higher for j. As a result, committee members would be willing to trade-off risk and return in order to deliver resources to preferred households. In practice, the empirical salience of this trade-off will depend on the extent to which risk and marginal returns correlate with committee member's weights.

[16] Empirically, Breza et al. (2018) show that high-productivity households exhibit higher marginal returns to credit and are best able to expand their businesses in the aftermath of the MBVF, highlighting the importance of understanding the targeting mechanisms in micro-credit programs.

3.2 Benchmark: a profit-maximizing bank

It is insightful to think of alternative ways in which resources could have been delivered. For instance, the central government could have donated the resources to banks instead of grassroots organizations with the aim of providing low interest credit to community members. Would a profit-maximizing bank allocate resources differently? Typically, a bank would try to maximize the expected revenues from lending ($\sum_{i=1}^{i=N_v} q_i(1+r)b_i$) net of the cost of managing the portfolio of loans ($F(\sum_{i=1}^{i=N_v} b_i), F' > 0, F'' > 0$). With interest rates externally set –i.e., by the government–, banks would allocate loans such that the marginal revenue equals the marginal cost of operation. As higher probability of repayment q_i increases the expected return, it would be optimal for a bank to target credit at less risky households.

The optimal bank allocation contrasts sharply with the committee's allocation. Because the bank is unable to capture the returns of the projects it decides to fund, productivity does not affect the allocation of credit, directly. However, relative to the case of the committee, it is possible that a profit-maximizing bank allocates more credit to higher return households as the Village Fund Committee balances risk, return, and different household weights. Thus, the relative efficiency of a community-based approach to allocating loans is an important empirical question which I explore in the following sections.

4 Data

The context of the MBVF program coincides with the availability of a high-frequency and detailed dataset: The Townsed-Thai Monthly Survey (Townsend, 2014) which follows 710 households on a monthly basis for over 14 years. Starting in September 1998, the survey includes three years preceding the program's implementation, which is essential to provide a full characterization of potential borrowers at baseline. While the survey covers only 16 villages,[17] the number of surveyed households per village is high, averaging 44 households per village, and representing a sampling rate of 42%. This feature makes the dataset ideal for the analysis of the distribution of resources within each village.

[17] The 16 sample villages were selected randomly from four provinces in Central and Northeast Thailand: Chachoengsao, Lop Buri, Buri Ram, and Si Sa Ket.

The level of detail of the information is unusual for developing country settings: The dataset provides high-frequency information regarding transactions among households, the portfolio of loans held by each household, purchase, sales and use of inputs as well as the destination of final output. Additionally, it is possible to link the survey with detailed information regarding households' financial statements (cash flows, income and balance sheets), which were constructed by Samphantharak and Townsend (2010). Such information allows to study households as corporate firms.

Table 2 reports summary statistics. Over 80% of households are landowners, and over one-third of household revenues are obtained from agricultural production. However, the average household obtains revenues from 4 different economic activities such as livestock, fishing and shrimping, and wage labor provision. Off-farm business ownership is not rare either: 15% of sample households obtained income from these businesses. Thus, in order to consider potential interactions across occupations within a household, the analysis in this study focuses on total revenues and input usage from all household operations.

In terms of household finances, during the year preceding the implementation of the program, 50% of the households obtained a loan from any source, and 40% of them obtained a loan from institutional lenders. Although borrowing is common, total debt represented only around 10% of household assets. In contrast, cash represented over 30% of household assets, which suggests that households were likely to self-finance their projects. Finally, among households with credit history, the average share of pre-program loans with missing payments is low (6%), but low delinquency coincides with high shares of loan term extensions (36%).

5 Measuring pre-program characteristics

5.1 Repayment

To assess repayment history, I exploit self-reported information regarding all new and outstanding loans from all sources of credit. Concretely, enumerators record information regarding loan characteristics for all new loans in a given survey wave, and follow up with with information regarding repayment until a loan is either fully paid or defaulted on. For each borrower, I recover the share of pre-program loans which exhibited at least one delinquent payment, and the share of loans with extension of the maturity period.[18]

5.2 Neediness

In order to proxy for baseline poverty, I compute averages of total per-capita consumption corresponding to the years preceding the program. In order to proxy for transitory neediness, I use the pre-program periods to compute several measures of shocks that may trigger demand for credit: the number of health symptoms reported by family members, and the number of operation problems for agricultural, livestock and off-farm household businesses, which include problems with production, input delivery and costumer delays in payment.

5.3 Productivity

The ideal assessment of the productive efficiency of the program's allocation requires information regarding the distribution of marginal returns to credit for all potential borrowers. While the estimation of the distribution of baseline returns to credit is beyond the scope of this paper, I focus on estimates of baseline household total factor productivity as a proxy. Intuitively, by capturing variation in output which is unexplained by input use, household TFP captures the ability of a household to generate revenues, holding input usage fixed. From a theoretical perspective, higher TFP may induce higher demand for inputs which may lead to binding credit constraints for households with limited access to credit. While imperfect, the ability to come up with a baseline proxy for productivity to evaluate targeting performance is unique in the targeting literature as other studies use contemporary or post-program proxies of yields or beliefs about average returns to inputs (Bardhan and Mookherjee, 2006b; Basurto et al., 2017).

Consider the following log production function in which output $y_{i,t}$, corresponding to household i during period t, is a function of labor $l_{i,t}$, non-labor variable inputs $m_{i,t}$, fixed capital $k_{i,t}$, productivity shocks $\omega_{i,t}$, and shocks to production $\epsilon_{i,t}$. Productivity shocks are

[18] I use delinquency rates rather than repayment rates as default is not very common in this context: recovery rates are on average over 97% (see Table 1)

known to the household but are unobserved by the researcher. They capture managerial ability as well as household-specific economic conditions that may affect the choice of inputs (e.g., rainfall or contract enforcement with costumers and providers). Production shocks $\epsilon_{i,t}$ are unforeseen shocks to production (e.g., production loss due to theft).

$$y_{i,t} = \beta_l l_{i,t} + \beta_k k_{i,t} + \beta_m m_{i,t} + \omega_{i,t} + \epsilon_{i,t} \tag{4}$$

The aim of this section is to back out estimates of pre-program $\omega_{i,t}$ for all potential borrowers. In order to do so, I first use 14 years of panel data to estimate the factor elasticities corresponding to (4). Second, I combine the estimated factor elasticities with 3 years of pre-program data regarding output and input use to back out estimates of baseline household productivity $\hat{\omega}_{i,t} = y_{i,t} - \hat{\beta_l} l_{i,t} - \hat{\beta_k} k_{i,t} - \hat{\beta_m} m_{i,t}$. Finally, I obtain averages corresponding to the pre-program years and use these estimates to study selection into program credit.

I focus the empirical analysis on total output and input usage across all household economic activities which include agriculture, livestock farming, fishing and shrimping, off-farm family businesses and wage work outside the household. I do so as households may simultaneously optimize resources across all economic activities. In the data, all households have at least two sources of revenues and, on average, derive income from four different sources (see Table 2).[19] However, I allow factor elasticities to vary across households with different business orientations; I separately estimate factor shares for households which are mainly involved in farm activities (agriculture, livestock, fishing and shrimping) and households households who mostly obtain revenues from off-farm activities (businesses or wage labor).

In order to estimate (4), I construct an annual panel by aggregating the balances of monthly income statements, and time and labor use survey data over each Thai economic year.[20] I do so to prevent seasonality from driving the results and to capture household

[19] As production functions are product-specific, aggregating across sources of revenues comes at the cost of interpretation of the factor shares. This issue may not be a first-order concern for this paper as its main aim is to recover a measure of variation in output that is not explained by input use, and not the analysis or comparison of factor shares themselves.

[20] Thai economic years begin in April and end in March. The beginning of the year coincides with the beginning of the rainy season and capture two full rice production cycles.

behavior over the production cycle. I then merge this dataset with household balance sheets measured at the beginning of each economic year (April). The resulting dataset is fundamental for the analysis of households as corporate firms as it allows the use of household balance sheets and variations in inventories to measure capital and non-labor inputs. Moreover, the availability of time-use data by both external workers and household members is essential to obtain a good approximation of household labor.

I proxy total output with gross revenues from all household activities in a given year. Non-labor inputs include fertilizer, seeds, feed, merchandise, and other tools required for non-farm family businesses. I focus on the total value of their usage in revenue-generating activities. One important feature of the dataset is that it allows the distinction between input purchases and input use, which is important in context in which households may store intermediates as inventories.[21] Consistent with models of time-to-build (Kydland and Prescott, 1982), capital is measured as the value of the stock of fixed assets for each household at the beginning of each year, and includes land, agricultural equipment, the value of livestock, real-state, non-farm business assets, machinery and other household assets. Finally, labor is measured as total hours per year devoted to all household operations, which includes labor provided by household members (on average 85% of total labor), and by workers outside the household. Note that I don't focus on labor expenditure as it will ignore the usage of household members' labor.

5.3.1 Identification of the gross revenue function

The empirical challenge is to consistently estimate the parameters from equation (4) in a context in which households choose inputs (l, m, k) in response to productivity shocks (ω). Ideally, one would like to rely on exogenous variation in capital, labor and inputs to estimate the production function. While such a rich setting is not available in the Thai context, the availability of a long panel dataset provides an opportunity to exploit the panel structure of the data to deal with endogeneity. For instance, credit constraints or other frictions may generate persistence in input decisions which, combined with assumptions regarding the law of motion of productivity, can be exploited to achieve identification of factor elasticities.

[21] See Appendix Section D for a more detailed discussion of measurement.

In particular, I take advantage of such rigidities to implement the dynamic panel approach proposed by Blundell and Bond (2000).[22]

Identification requires two types of assumptions: *i)* Assumptions regarding the law of motion of productivity, and *ii)* assumptions regarding the timing and available information for input choice. I discuss both below.

Linearity: Following Blundell and Bond (2000), I assume that $\omega_{i,t}$ follows a first-order autoregressive process:

$$\omega_{i,t} = \alpha_i + \rho\omega_{i,t-1} + \zeta_{i,t} \tag{5}$$

where α_i denotes cross-household, time-invariant heterogeneity and $\zeta_{i,t}$ denotes unforeseen productivity shocks. Note while $\zeta_{i,t}$ is not observed by the researcher, it is known to the household which can respond by adjusting inputs. Plugging in (5) into (4), and subtracting $\rho y_{i,t-1}$ from both sides of the equation (i.e., ρ-differencing), it is possible to obtain the following reduced-form equation:

$$y_{i,t} = \tilde{\alpha}_i + \gamma_{l1}l_{i,t} + \gamma_{l2}l_{i,t-1} + \gamma_{k1}k_{i,t} + \gamma_{k2}k_{i,t-1} + \gamma_{m1}m_{i,t} + \gamma_{m2}m_{i,t-1} + \gamma_y y_{i,t-1} + \zeta_{i,t} + \tilde{\epsilon}_{i,t} \tag{6}$$

where $\tilde{\alpha}_i = \alpha_i(1-\rho)$, $\gamma_{j1} = \beta_j$, $\gamma_{j2} = -\rho\beta_j$ (for $j = \{l, k, m\}$), $\gamma_y = -\rho$ and $\tilde{\epsilon}_{i,t} = \epsilon_{i,t} - \rho\epsilon_{i,t-1}$.

A direct consequence of equation (5) is that the error term in the reduced-form equation is no longer a function of past realizations of productivity ($\omega_{i,t-1}$) and is only a function of the unforeseen part of productivity $\zeta_{i,t}$ and the fixed effect α_i.

Equation (5) could be restrictive as it imposes linearity and only allows productivity shocks in $t-1$ to directly feedback into productivity in t. Despite imposing linearity,[23]

[22] I use the dynamic-panel method suggested by Blundell and Bond (2000) over choice-based methods which are common in the production function literature (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015) based on the insights from Shenoy (2017b) and Shenoy (2017a): when households or firms face credit constraints, they may not be able to adjust their demand of inputs to accommodate productivity shocks, which is the main assumption in choice-based methods. In contrast, the existence of constraints or rigidities makes autoregressive estimators viable as lagged input choices may be relevant instruments for current input choices.

[23] At the costs of imposing structural assumptions, namely that households freely adjust the demand of

equation (5) is flexible enough as it allows the presence of a fixed effect and a time-variant component with persistence, which is consistent with models of learning. Moreover, equation (5) is less restrictive than assuming that TFP is a fixed-effect or than assuming that there is no persistence in productivity shocks.

Timing: Identification of the reduced-form parameters requires imposing assumptions regarding how households adjust inputs in order to accommodate unforeseen productivity shocks. Note that since the reduced form equation includes a fixed effect, equation (6) should be estimated in first differences (Arellano and Bond, 1991),[24] yielding the following error term: $\zeta_{i,t} - \zeta_{i,t-1} + \epsilon_{i,t} - (1+\rho)\epsilon_{i,t-1} - \epsilon_{i,t-2}$. Thus output and input choices from $t-3$ backwards are valid candidates for instruments. In this case, the associated exclusion restriction would be that while households may change input choices as a response to current changes in productivity, input choices made in $t-3$ or earlier are not correlated with future unforeseen productivity shocks. In terms of economic behavior, this assumption allows households to accommodate future productivity shocks up to two periods ahead. In terms of instrument relevance, this assumption demands that third lags of inputs and output in levels are predictive of first differences of inputs and outputs. Appendix Table A1 shows that indeed lagged levels have predictive power.

5.3.2 Estimation and validation of productivity measures

Estimating household productivity involves three steps. First, I estimate equation (6) using the system-GMM (Blundell and Bond, 1998) approach with lags 3 to 5 as instruments. In addition, I included rainfall and operation shocks as well as year and village fixed effects as controls. An econometric discussion of this specification are detailed in Appendix ection C.[25]

inputs in response to productivity, the process in equation (5) could be generalized to a first-order Markov process which can be non-perametrically estimated. However, such structural assumptions are unlikely to hold in developing countries (Shenoy, 2017a).

[24]The model in first differences is:

$$\Delta y_{i,t} = \gamma_{l1}\Delta l_{i,t} + \gamma_{l2}\Delta l_{i,t-1} + \gamma_{k1}\Delta k_{i,t} + \gamma_{k2}\Delta k_{i,t-1} + \gamma_{m1}\Delta m_{i,t} + \gamma_{m2}\Delta m_{i,t-1} + \gamma_y \Delta y_{i,t-1} + \Delta\zeta_{i,t} + \Delta\tilde{\epsilon}_{i,t}$$

[25]I use the system-GMM estimator over the difference-GMM estimator (Arellano and Bond, 1991) as the latter only exploits within household variation to estimate elasticities. In contrast, system-GMM uses a richer set of moment conditions that also exploit cross-household variation. This distinction is important as variation in the stock of capital is mostly explained by cross-household variation since investment is rather

This process yields 7 reduced-form parameters ($\hat{\gamma}_y, \hat{\gamma}_{j,p}$ with $j = \{k, m, l\}, p = \{1, 2\}$). Second, I use the reduced-form parameters to back out the structural parameters ($\hat{\beta}_k, \hat{\beta}_m, \hat{\beta}_l, \hat{\rho}$) through optimal minimum distance (OMD). I implemented this process separately for two sectors: *i)* households whose revenues from farm activities (cultivation, livestock and fishing) account for more than 50% of their baseline revenues, and for *ii)* households whose revenues mostly come from off-farm activities.

Table 3 reports the estimated factor elasticities as well as the reduced-form coefficients for households in the farm and off-farm sectors. Panel A shows that the reduced-form specifications are likely to pass the Hansen test for overidentifying restrictions highlighting the validity of the instruments. Turning into the structural estimates, factor elasticities are estimated with precision and, in both cases, pass the common factor restriction test. The point estimates suggest decreasing returns to scale in the case of farm-oriented households, while it is not possible to reject the constant-returns-to-scale hypothesis in the case of households who derive income mostly from off-farm activities. Importantly, in both cases there is evidence of strong persistence suggesting that baseline productivity captures important information for targeting analysis.

Once the estimated factor elasticities are recovered, it is possible to compute baseline TFP for all potential borrowers. Appendix Table A2 shows correlates of baseline TFP with demographic characteristics and productivity shifters after controlling for village fixed effects. As expected, in the case of farm-oriented households, rainfall interacted with the share of agricultural revenues is strongly and positively correlated with TFP. Moreover, productivity is negatively related with household age, which is consistent with a sector with more physically demanding activities. In contrast, for off-farm-oriented households, household productivity is correlated with the household head's years of schooling, which is consistent with the idea that better educated households may have comparative advantages into off-farm businesses.

lumpy. For instance, Samphantharak and Townsend (2010) find that only 11% out of 55,000 household-month observations during the first 84 months of the Townsend Thai survey recorded positive investments. However, while more precise, the system-GMM estimator requires an extra identification assumption, mainly that first-differences of k, m, l and y are not correlated with their initial levels.

5.4 Measuring connections with the village government

To proxy for connections with the local leaders I follow a two-step approach. First, I use pre-program information regarding participation in the local government to identify households with at least one member participating in Village Council. Second, I use baseline data regarding kinship relations across households in each village to elicit kinship networks. In a similar way, I use detailed information regarding several economic interactions to elicit a baseline undirected, unvalued transaction network for each village.[26]

I then use these socioeconomic networks to define connectedness with the elites. Concretely, a household is defined as connected with the local political elite if any of its members reports either being a member of the village council, or a first-degree kin of a council member, or having engaged in at least one transaction, of any type, with any village council member during the baseline periods.[27] One limitation of this approach is that, by using the extensive margin of transactions to define connections, it is possible that a household is identified as connected because of one isolated interaction. Since the relative salience of each interaction cannot be identified nor valued, when pertinent I provide robustness checks using an alternative definition of connectedness based on Principal Component analysis across the different types of transactions.

6 Selection

Equipped with a comprehensive set of baseline characteristics, I proceed to analyze selection into the program. While the program currently operates in several villages, I focus on the first two years of the program for two reasons.[28] First, I investigate the extent to which

[26]The transactions can be roughly categorized in seven groups: output sales/purchases, asset purchases/relinquishments, transfers (gifts), borrowing/lending, paid labor provision/demand, unpaid labor exchange, and other inputs, which include materials purchases/sales as well as advising. Consistent with Banerjee et al. (2013), I consider all possible transactions as information may be transferred through different type of interactions. See Appendix Section D for a detailed explanation of the construction of the variables.

[27]While other measures–such as geodesic distance (shortest path)–might provide a better approximation of the distance between a household (node) and the elites in the network, these measures are subject to potentially high biases arising from the sampled nature of the transaction data (Chandrasekhar and Lewis, 2017).

[28]I choose two years in order to capture households that may not have needed credit during the first year but obtained credit during the second year.

baseline characteristics predict program participation, and since they are time variant and could potentially respond to the program, baseline characteristics are more representative of the context around the rollout of the program. Second, modifications were made to the program in 2004, three years after its initial rollout.[29]

Panel A from Table 5 reports-within village correlates of baseline characteristics with program participation. Column (1) shows that households with high pre-program access to institutional credit were 32 percentage points more likely to obtain program resources than households without access to institutional credit. While this result suggests that the committee members were unable to reach unbanked households, it is also possible that committee members used experience with formal credit as a proxy for credit-worthiness. If that was the case, then program borrowers should exhibit better credit history. Column (2) shows that, among households with baseline credit history (76% of total sample households), households who missed loan payments at baseline were not penalized when it comes to program participation. Moreover, Panel A from Appendix Table A3 shows that program borrowers' repayment history was rather poor. Among household with credit history, program borrowers were 12 percentage points more likely to have requested term extensions (FWER-adjusted $p - val = 0.02$).

Selection into the program was not consistent with a poverty-targeting criterion. I find that better-off households obtained significantly more program resources. Column (3) in Panel A from Table 5 shows that a one-percent increase in baseline per-capita consumption is related to a 7 percentage points increase in program participation. Figure 1 plots the distribution of per-capita consumption for program borrowers and non-borrowers. At each point in the distribution, consumption is higher for program borrowers. These results suggest that the allocation of credit was fairly regressive, and that redistribution was not a driving force in the allocation of resources. However, it is possible that loans were allocated to households who suffered negative shocks in the periods preceding the program. As suggested by Alatas et al. (2012) and Basurto et al. (2017), a decentralized approach to targeting resources performs particularly well at identifying transitory components of neediness, which are likely

[29]For instance, there were changes in the orientation of the funds to community improvement projects, sanctions for poorly managed funds, and rewards for successful ones.

to be ignored by per-capita consumption. Panel B from Appendix Table A3 suggest that program borrowers were more exposed to negative shocks during baseline periods. Nevertheless, it is unlikely that negative shocks account for most of the differences in program participation: the differences are exclusively driven by shocks to livestock business which represent only 8% of household revenues (see Table 2).

One possible explanation is that committee members did not allocate credit to the needy, but allocated credit to those who are best able to convert credit into profits. Column (4) in Table 5 shows that baseline productivity is not a good predictor of program participation. Panel C from Appendix Table A3 shows that the results are similar using a TFP based on elasticities estimated through an alternative method (difference-GMM $p - val = 0.097$) and are qualitatively similar to using the baseline Return over Assets (RoA) as an alternative measure of household profitability. While this result suggests that resources were inefficiently allocated on average, an analysis of the differences between program borrowers and non borrowers along the distribution of productivity provides insights about the extent to which misallocation is explained by the inclusion of low-productivity households, or the exclusion of high-productivity households. Figure 2 compares the cumulative distribution functions of program borrowers and non-borrowers, it shows that committee members were indeed able to screen out lower productivity households but missed high-productivity households.

Interestingly, the committee's ability to screen out low productivity households varies by sector. Figure 3 presents differences in the distribution of tfp between program borrowers and non borrowers by main business orientation. In the case of farm-oriented households, the bottom quantiles of the tfp distribution were higher in the case of program borrowers. After controlling for shocks to household operations, quantile regressions show that these differences were significant for the bottom third of the distribution. No significant differences were found in the case of households from the non-farm sector. One explanation could be that community members may have a better understanding of traditional agricultural projects but not so in the case of less traditional off-farm businesses.

The results suggest that neither repayment, nor poverty, nor productive efficiency were relevant targeting criteria. Given that these important dimensions are not related to the allocation of resources, the theoretical framework discussed in Section 3 suggests that committee members may have applied different weighted to households. I examine this hypothesis by comparing program participation between households with and without connections to the local elite.

I find strong evidence that being connected to the village government is associated with higher chances of obtaining program credit. Figure 4 shows how access to program credit and loan size vary with the type of relation with local leaders. As the resources from the program were rolled out, households with a member in the village council or with baseline connections to council members were more likely to be obtain program credit sooner than unconnected households. After a year of the rollout of the program, elite members and households with connections to elite members are 30 and 20 percentage points more likely to hold program credit than unconnected households, respectively. Moreover, elite members obtain twice as much program resources that non-elite members with business connections to elite members, and almost three-times as much resources than unconnected households. Column (5) from Panel A in Table 5 shows that households who are either connected to the village council or have a direct link in the socioeconomic network are 16 percentage points more likely to obtain credit after controlling for village fixed effects. Panel D in Appendix Table A3 shows that is this correlation is robust across several measures of elite-connectedness.[30] The results suggest that committee members applied higher weights on elite-connectedness than in risk, poverty or productivity. Column (6) from Panel A in Table 5 shows that even after controlling for risk, neediness and productivity, being connected with the elite is a key predictor of program participation.

6.1 Comparison with hypothetical eligibility criterion based on repayment scores

While the results from the previous section suggest that program resources were poorly targeted, it is not clear to which extent a community-based approach to allocating credit outperforms other policy-relevant ways of dispensing resources. In this section, I use the set of households who would have been eligible under a repayment-scoring model as a hypo-

[30] In particular, an extra link with a member of the village council increases the chances of obtaining program credit by 10 percentage points.

thetical comparison group. While a scoring model would only identify eligible households as opposed to households that would actually borrow, it is still a policy-relevant comparison for two reasons. First, several micro-credit institutions in developing countries have relied on scoring models to screen applicants for a couple of decades (Schreiner, 2000). Second, given the fast rates of data digitalization in developing countries, the use of scoring models could be the building block for technology innovation in credit delivery (Björkegren and Grissen, 2018). Intuitively, it is not clear whether credit score models which rely on hard information and objective targeting rules outperform alternative ways of screening applicants such as community-based approaches which rely in soft information and are likely to be more discretionary.

In order to estimate delinquency risk, I exploit a rich dataset corresponding to over 3,800 loans obtained by surveyed households from different types of lenders (formal and informal) during the pre-program periods. I complement the loan data with several household financial and demographic characteristics. I then estimate models of the probability of having a delinquent payment, for a given loan, as a function of loan, lender and borrower characteristics, and village and year fixed effects. To obtain a parsimonious model and minimize the risk of over-fitting, I use the least absolute shrinkage and selection operator LASSO with a penalty parameter chosen through 10-fold cross validation.[31] I use the LASSO coefficients to estimate delinquency risk for all potential borrowers, and use the delinquency risk estimates to create within village risk rankings. For each village, the households with the $k-th$ lowest positions (lower risk) are classified as the hypothetical target group. k is picked such that the number of targeted households in the hypothetical target group coincides with the number of program borrowers in each village. Thus, it is possible to think of this exercise as an hypothetical change in the targeting criterion holding the number of program beneficiaries constant.

This process classifies households into four groups: households that would have been

[31] Consistent with other scoring models in Thailand (Limsombunc et al., 2005), household debt-to-assets ratio predicts higher risk of delinquent payments. In contrast, education and the number of previous loans with the same lender reduce the delinquency rink, suggesting that the scoring model reasonably captures repayment behavior. Appendix Table A4 reports estimates of linear probability models of the probability of exhibiting a delinquent payment. Column (1) reports OLS coefficients using the full set of covariates. Column (2) presents OLS coefficients of a more-parsimonious model using only the variables selected by the LASSO.

targeted by both the program and the respective alternative criterion, households that would have been excluded from both allocations, households that were reached by the MBVF program but would have been excluded by the alternative criterion, and households that were excluded from the VF program but would have been targeted by the alternative criterion. Panel B of see Table 4 shows that 34% of program borrowers would be ineligible under a repayment-scoring criterion (over-inclusion error), and that 47% of non-borrowers would have been eligible under the scoring model (over-exclusion error). These results confirm the findings in the previous section: repayment risk was not a relevant dimension considered by committee members.

How do different baseline characteristics relate to the probability of eligibility? Table 5 reports correlates of program selection (Panel A) and hypothetical eligibility (Panel B) with baseline characteristics. While both criteria target households with baseline experience with institutional credit and higher per-capita consumption, the hypothetical criterion achieves this by punishing households with a history of missed payments and giving priority to productive households. Having loans with missed payments before the program significantly reduces the probability of being eligible by 14 percentage points under the hypothetical criterion, while it was not predictive of higher uptake in the case of the allocation achieved by committee members (difference p-value = 0.002).[32] Turning to productivity, household TFP is significantly correlated with eligibility in the case of the hypothetical criterion but not so in the case of the committee's allocation, however the difference across criteria is not precisely estimated (difference p-val=0.27).

Interestingly, connections with the local elite do not explain eligibility in the case of the hypothetical targeting criterion, but they are highly correlated with actual program participation (difference $p - val < 0.01$), which suggest that elite-connected households, on average, were not more reliable borrowers. One possible explanation is that scoring models, while imperfect, provide an objective rule for the allocation of resources that may be less prone to resource capture. In contrast, while community members may use soft information to allocate resources more efficiently, it is not clear that the set of incentives or committee's

[32]Inference across regressions is performed through jointly estimating both models through seemingly unrelated regressions (SUR).

preferences are well aligned with program objectives.

Overall, a repayment-based targeting criterion seems to dominate the actual community-based allocation in terms of repayment, productivity without relying on connections. This pattern persists even after controlling for demographic characteristics and exposure to pre-program shocks (see Column (6)), and suggests potential gains from re-allocating the resources.

Figure 5 depicts the distribution of TFP based on program participation and eligibility based on the hypothetical criterion and shows that there are important differences by groups. Household TFP corresponding to households who borrowed but would not be eligible under the hypothetical criterion is lower than that of households who would have been eligible by the hypothetical criterion. This pattern is similar in the case of TFP estimated through difference-GMM (see Appendix table A1). Table 4 provides back-of-the-envelop calculations to assess potential gains/losses in terms terms of risk, per-capita consumption and productivity. It reports relative differences from reallocating resources from program borrowers who would be ineligible by the repayment-based criterion to non-borrowers who would have been eligible by this criterion.[33] It shows that reallocating resources towards less risky households would have reduced the delinquency risk of the pool of borrowers by 6 percentage points (25% of the average risk of actual borrowers) and would have increased average productivity by 6%. These risk and productive efficiency gains would have been achieved at the cost of delivering credit to even richer households. This result highlights the complexity of allocating credit: targeting the poor might be at odds with repayment or efficiency.

7 The role of connections with the village council

The presence of asymmetries in power within local communities may undermine the success of efforts to decentralize the allocation of resources, exposing these efforts to resource capture and favoritism. However, the appeal of decentralized approaches to delivering pub-

[33] I obtain the estimates by subtracting the difference between program borrowers with low-repayment probabilities and non-borrowers with high-repayment probability, and then scaling down this difference by 0.34 (the over-inclusion error rate) in order to obtain gains representative of the pool of borrowers.

lic resources relies on the idea that social connections may transmit information regarding program beneficiaries which might be costly to obtain by traditional policy makers. Both mechanisms may not be mutually exclusive in the context of allocating credit; community members may use information to select credit receivers, but may provide more favorable credit conditions to households elite-connected households, which could be costly for the program. The next section analyzes these two mechanisms.

7.1 Do connections with local elites transmit information?

If being connected to village council was used as a signal of credit worthiness or efficiency, elite-connected households should be, on average, better potential borrowers. If that was true, the observed committee's allocation could be a result of statistical discrimination. However, Appendix Table A5 shows that elite-connected households are if anything, riskier and not more productive than unconnected households. For instance, among households with baseline credit history, elite-connected households are more likely to have had delinquent payments and expand the term of their loans at baseline. Such patterns are stronger for Village Council members, but still present for households with direct connections in the transaction network.

An alternative explanation is that while elite-connected households may not be better borrowers, their location in the network may allow them to better transmit soft information that is relevant for the allocation of credit. An empirical implication of this mechanism is that the correlation between program participation and elite-connectedness should vanish after controlling for location in the network. Table 6 reports within village correlations of connectedness with and without controlling for network degree centrality (number of links in the network). Interestingly, adding degree centrality as a control reduces the difference in program participation based on connections with the elite from 16 to 6 percentage points, which is no longer significant.

This set of results suggests that connections ease the transmission of information. In principle, such finding is expected and encouraging as community-based targeting should exploit connections. A more pessimistic interpretation suggests that community-based approaches, by relying on networks, impose higher costs of obtaining resources to households not very well located in the network, which could lead to mistargeting. In this regard, lack of connections may operate as targeting frictions imposing higher costs of transmitting information for unconnected households. However, Columns (4) to (6) from Table 6 show that there are differences by type of connections. Even after accounting for network position, elite members are still almost 20 percentage points more likely to obtain program credit, raising suspicions regarding resource capture or favoritism.

7.2 Was there favoritism?

Committee members were also in charge of deciding the size, term and price of program loans, and connections could play an important role. On the one side, committee members may use connections to achieve better enforcement. In contrast, committee members may provide more favorable loans to elite-connected households. Empirically, both motives would predict the provision of larger and cheaper loans to connected households. However, unlike the enforcement motive which is efficient for the lender, favoritism should be costly for the program. Similar insights have been used to across different monitoring models in agriculture (Shaban, 1987), to study the role of comparative advantages and taste-based discrimination in agricultural tasks (Foster and Rosenzweig, 1996), and to test for favoritism towards firms with connections to the central government in Pakistan (Khwaja and Mian, 2005). This rationale provides the theoretical foundation to empirically test for favoritism.

Testing for favoritism requires two important elements: First, it requires a measure of the *ex-post* returns to the lender as the main outcome variable. Second, it requires a credible way of computing differences in loan profitability controlling for unobserved borrower characteristics.

I tackle the first issue by exploiting detailed data regarding the full stream of payments for each loan, which includes all the periods in which they were reported as active.[34] I use this information to compute the *ex-post* internal rate of return for each loan (IRR) as a measure of returns to the lender, which considers the combined potential effects of loan size, interest rates, loan term and repayment behavior along the life of each loan.

[34] The sample includes fully repaid loans and loans which were defaulted on. It excludes loans for which there was no payment ever reported

In order to test for favoritism accounting for unobserved borrower characteristics, I exploit the following insight: While connections with the elite may be salient in the case of a publicly funded credit program, elite-connectedness should be of lower relevance in the case of privately funded sources of credit. A testable implication is that while, regardless of the borrower, loans from the program may exhibit lower IRRs than those from private lenders, larger differences in the case of loans to elite-connected borrowers relative to loans to unconnected borrowers should be indicative of favoritism.

The Thai context offers an ideal setting to implement this test as the program overlapped with the existence of other community-based sources of credit. I focus on loans from production credit groups (PCGs), women's groups, and other village organizations to construct a comparison group for program loans.[35] These credit groups and the MBVF program are both managed by community members. However, their funding source is different: The MBVF program is fully funded by the Central Government while local credit groups are funded through contributions from group members. Thus, it is possible to exploit two sources of variation: variation in borrower's connection status, for a given lender, which captures the potential for political influence; and variation in the origin of the funds, for a given borrower, which captures the ability of borrowers to take advantage of their connections.[36]

I bring this idea to the data by exploiting a sub-sample of 6,700 loans which were obtained after the program was introduced. These loans correspond to 335 households who obtained credit from both the program and other local credit sources. While using a selected sample limits the extrapolation of the results to the entire village financial system, observing the same household borrowing from both sources of credit is essential to control for unobserved borrower characteristics that are invariant with respect to the lender. In addition, as each type of lender lends to connected and unconnected borrowers, it is possible to control for unobserved lender characteristics. Consider the following specification:

[35] These sources of credit, sometimes labeled as quasi-formal, have been shown to be helpful in promoting asset growth, consumption smoothing, and occupational mobility through the provision of cash credit to community members in the context of Thailand (Kaboski and Townsend, 2005).

[36] See Table 1 for comparative summary statistics for different sources of credit.

$$IRR_{kijt} = \alpha_i + \theta_j + \beta \text{Connected}_i \times \text{MBVF}_j + \delta_{vt} + \epsilon_{kijt} \quad (7)$$

The unit of observation is a loan k obtained by household i from lender j in year t. α_i and θ_j denote households and lender fixed effects. In order to account for potential differences in the local financial conditions when loans are obtained, I include village-year fixed effects (δ_{vt}). $Connected_i$ and $MBVF_j$ are indicators of whether a borrower had pre-program connections with the elites and whether the loan was obtained from the MBVF program.

The parameter of interest is β which measures relative returns of lending to connected households for the MBVF program, with respect to other privately-funded community based lenders. Under the assumption that there were no unobserved factors disproportionately affecting program loans corresponding to connected households, $\beta < 0$ will be supportive of favoritism. In contrast, $\beta > 0$ is consistent with better monitoring based on connections.

Column (7) from Table 7 reports estimates of β corresponding to the specification in equation (7). Panel A shows that there is a 2.7-percentage-points decrease in the *ex-post* internal rate of return to the lender for MBVF loans to connected households, which accounts for over one-third of the average IRR for program loans. This result is neither driven by loans from borrowers that rarely borrow nor by smaller loans: Column (8) shows that the results are robust to weighting observations by the number of loans corresponding to each borrower in the sample, and Column (9) shows that the results are robust to weighting each loan by loan size.

The differences in returns are likely to be driven by preferential interest rates for elite-connected borrowers. Panel B shows that there are not significant differences in repayment behavior,[37] suggesting that elite-connected households at least complied with their payments. However, Panel C shows that elite-connected households obtained lower initial interest rates and larger amounts. Put together, the results suggest that, for a similar level of risk, program committee members delivered cheaper credit to elite-connected borrowers which implied forgone returns to the lender of the order of 2.7 percentage points per TBH

[37] I omit default as both the program and local credit groups have almost null default rates.(see Table 1).

lent to elite-connected borrowers, and a premium of 1 percentage point in interest rates for connected borrowers. Note however, that these estimates are only valid for the subset of loans corresponding to households who borrowed from both sources, which might be different than those who only borrowed from the program. In that regard, these results suggest the existence of favoritism, but are unable to explain the potential consequences for the full portfolio of program loans.

The results are not surprising as the relation between the incentives provided by the Central Government and the village fund committees is similar to that of a principal-agent problem of multiple tasks and incomplete contracts (Holmstrom and Milgrom, 1991; Hart et al., 1997). The government threatened to eliminate other transfers to villages with poor repayment performance. In contrast, other than demanding a positive interest rate, the pricing decision was completely let to the committee's discretion. Empirically, committee members don't seem to be compromising repayment which could trigger government sanctions, but seem to be using interest rates to favor elite-connected households. These results is consistent with other studies analyzing favoritism in credit markets in the context of state-owned banks and politically-aligned firms (Khwaja and Mian, 2005), and private firms and membership to elite clubs (Haselmann et al., 2017).

The results could be interpreted as favoritism under the assumption that there were not unobserved characteristics that particularly affected program loans to elite-connected households. One violation to this assumption would arise if projects from connected households were more likely to have spillovers and spur aggregate growth. I argue that while it is possible, this mechanism is unlikely. First, loans were not particularly allocated to the most productive households (Section 6), and elite-connected households were not more productive (see Appendix Table A5). It is also possible that committee's choices reflect simply social norms or community preferences which would imply that the estimates provided in this section represent the financial cost of such preferences rather than the cost of favoritism. Given that loans from the comparison group are also likely to be exposed to such preferences, concerns regarding social norms or community preferences seem unlikely to drive the results.

8 Redistribution through informal credit markets

The results from the previous sections suggest that connections with the local elites may create targeting frictions that result in unconnected households obtaining substantially less resources from the program. If these connection-based frictions prevent credit-worthy unconnected households from financing profitable projects through program loans, then other well-informed lenders in the village should be willing to lend to unconnected households.

I test the empirical relevance of this argument by analyzing whether the credit supply shock generated by the program indirectly increased credit use by unconnected households. Two important features make the Thai context ideal for this test. First, the program represented a sudden increase in total lending in the village economy: Within one year from the rollout of the program, aggregate borrowing increased by 24% in the sample villages. Second, the presence of active informal credit markets in the study villages provides a potential mechanism for redistribution.[38] Testing for connection-based program spillovers is important as it contributes to understanding the extent to which program participation was driven by unconnected households self-excluding from the program or by connection-based targeting frictions. Moreover, testing for program spillovers is informative about the role of markets in offsetting targeting errors and about the extent to which devoting resources to improve targeting of social programs may be a first order concern for policy makers.

I exploit monthly variation in the differential rollout of the program across villages to identify the effect of an increase in the aggregate supply of credit in the local economy on borrowing from informal lenders. The resources were released in June 2001 in the first village in the study sample, and the rollout continued until February 2002 for the last village in the dataset. I combine this source of variation with pre-program measures of elite-connectedness to test for heterogeneity in borrowing from informal lenders.

Identification of the treatment effects from the rollout of the program is achieved under the assumption that, conditional on household time-invariant characteristics, the rollout of the program was not related to unobserved shocks that determined household decisions to

[38]For instance, using the first 88 waves of the Townsend monthly survey, (Kinnan and Townsend, 2012) documented that among households without access to formal credit, being connected to a households with access contributes to consumption smoothing.

obtain credit. This assumption seems plausible as the timing of the program was mostly generated by differences in the timing of the establishment of village fund committees, which is arguably orthogonal to the village economic environment. In order to examine the presence of pre-program trends and the dynamics of the effect of the program, I focus on the 18 months preceding and following the implementation of the program to compute flexible difference-in-differences estimates of the effect of the rollout of the program on credit:

$$Y_{ivt} = \alpha_i + \delta_t + \sum_{j=-18, j\neq -1}^{j=18} \beta_j \mathbb{I}[\tau_{vt} = j] + \epsilon_{ivt} \tag{8}$$

where Y denotes total gross borrowing from local informal lenders by household i, in village v, at month t. τ_{vt} denotes time to treatment, for each village in a given month. Household fixed effects are denoted by α_i, and δ_t denotes a set of month and year indicators. The coefficients of interest are $\{\beta_j\}_{j=-18}^{18}$, which capture the difference between borrowing in period $\tau_{vt} = j$ relative to the month preceding release of the funds ($\tau_{vt} = -1$) compared to the difference in borrowing in villages where funds were not released by that month. To approximate the average treatment corresponding to the post-rollout periods, I also estimate:

$$Y_{ivt} = \alpha_i + \delta_t + \beta Post_{vt} + \epsilon_{ivt} \tag{9}$$

In this equation, $Post_{vt}$ is an indicator that takes the value of 1 in the months following the rollout of the program in each village.

To account for within village correlation of error terms in the context of a small number of clusters (16 villages), regression tables report p-values from the wild bootstrap-t procedure suggested by Cameron and Miller (2015) imposing the null hypothesis of no effect. However, this approach tends to have low power and lead to conservative inference.[39] For comparison, standard errors clustered at the household level are also reported.

Figure 6 presents estimates of the impact of the rollout of the program on program credit

[39] As discussed by Cameron and Miller (2015) most available corrections for small number of clusters lead to appropriate acceptance rates, but they have reduced power. This is a concern in this paper as the number of cross-section observations is small.

and total credit. There are sharp differences in program credit between elite-connected and unconnected households. The situation is the opposite in the informal credit market. Figure 7 shows that while elite-connected households do not seem to change their behavior in the informal credit market, unconnected households respond by borrowing more from informal lenders. Table 8 presents average treatment effects by elite-connectedness and shows that the program rollout lead to a 30% increase in informal debt in the case of unconnected households, mostly driven by loans from relatives (see Figure 8). Overall, the results show that informal credit markets can offset targeting distortions based on connections. However, this correction is rather partial as the effect on informal credit markets only accounts for 15% of the program borrowing gap among connected and unconnected households.[40] Moreover, reallocation comes at a cost: the average interest rate corresponding to loans from relatives during the 18 months following the program implementation is 14% per year, which is twice as large as the average rate for program loans (7% per year).

The results reported in the previous paragraph are consistent with evidence of re-lending. Appendix table A7 shows that the probability of lending to other households increased by 2 percentage points in the case of connected households (12% of pre-program average), as a result of the rollout of the program. Event-study estimates show that there was a surge in total lending for connected households within two months of the rollout of the program (see Appendix figure A2), yet these effects are imprecisely estimated.

9 Concluding remarks

Community-based approaches to targeting public resources are increasingly popular in the policy world. Despite that, little is known regarding the performance of these approaches in market-driven environments such as credit. This paper brings together two central debates in development economics: the delivery of public resources through local democratic organizations and the provision of affordable credit to poor, high-productivity households. The results in this paper highlight the limitations of a subsidized community-based credit program to deliver credit to poor, high productivity households. Consistent with the tra-

[40] Appendix Table A6 shows that the program credit gap is around TBH 5,000.

ditional concern of resource capture in the literature that studies the decentralization of public programs to community members (Bardhan and Mookherjee, 2005), resources from the program were disproportionately allocated to elite-connected households.

These results are partially explained by information transmission frictions. While connected households are not better borrowers, their location in the village network may better allow them to transmit soft information to committee members. This has important policy implications in contexts in which attributes for beneficiaries are hard to observe. The extent to which community-based targeting approaches lead to better targeting will depend on how well connected are potential beneficiaries. Concretely, if poor,high-productivity households are socioeconomically isolated, even in the absence of rent-seeking behavior they may be less likely to be targeted. This result complements evidence showing how village network characteristics explain heterogeneity in targeting errors from a community-based cash transfer program(Alatas et al., 2012).

This paper also documents evidence of favoritism in a context of transparent elections of village fund committee members and speaks to the debate regarding the delivery of public resources through local democratic organizations. While the expectation was that transparent elections would ensure accountability, the results in this paper suggests that elections politicized the allocation of resources. The results are consistent with the theoretical prediction that decentralization may lead to regressive allocations when policies are financed through government grants instead of user contributions (Bardhan and Mookherjee, 2006a), as is the case of the MBVF program, and with cross-village studies documenting favoritism and clientelism (Asher and Novosad, 2017; Anderson et al., 2015). Overall, the results suggest that differences in connections to the local elite across households capture different in costs of accessing to public resources. These costs are related to information transmission but also to favoritism and are consequential in terms of equity, productive efficiency and program sustainability.

The results contrasts sharply with evidence in the context of community-based targeting of cash-transfer programs (Alatas et al., 2012) but is consistent with evidence of favoritism towards politically connected firms and credit from state-owned banks (Khwaja and Mian, 2005). The intuition for this result is that, as opposed to targeting cash transfers, the allocation of credit not only involves information regarding poverty but also productivity and repayment. Information regarding poverty is more likely to be objective and common knowledge to the community as a whole. Community members may use observable characteristics that describe a poor household and may not need to interact directly in order to figure out who is poor. In contrast, information regarding productivity and repayment requires direct economic interactions and thus may increase the incentives for moral hazard behavior.

A first order concern is that of how to effectively use the information available to community members and simultaneously prevent rent-seeking behavior in community-based approaches. One way could be by fostering self-funded credit groups, as opposed to creating village funds with subsidized resources. This is already a popular policy approach backed with encouraging evidence of its effects both on household productive behavior (Kaboski and Townsend, 2005; Deininger, 2013) and in relieving households from high-interest money lenders (Hoffmann et al., 2017). Research testing whether there are social barriers preventing poor, high-productivity households from participating in these groups would shed light regarding the effectiveness of this approach to alleviate poverty. Moreover a more careful comparison of the mechanisms driving selection into credit across different policy-relevant implementation approaches –i.e., CBT, self-help groups and traditional microfinance– would provide insights towards future policy directions. An alternative way is to provide monetary incentives for accurate information (Hussam et al., 2017), however the implementation of these incentives may require bureaucracy which is precisely what CBT approaches are trying to avoid.

This study also speaks to the importance of understanding the interactions of public policy efforts with markets, and political economy factors in a general equilibrium framework. In particular, this paper contributes with novel evidence showing that credit markets may offset potential targeting errors. While evidence of spillovers from large scale programs towards mistargeted households may suggest that targeting should not be a first order concern as markets may deliver resources to the intended destination, the relevant question is the price mistargeted households have to pay in order to benefit from public resources. This study finds that other lenders in the village financial system and kinship networks are important in indirectly delivering results to households lacking of connections with local leaders. While the former involved higher interest rates than those from loans from the program, the latter may imply interlinked transactions which may be costly for either the borrower or the lender. These costs may ultimately determine if targeting should be a first or second order issue in public policy.

Finally, this paper provides evidence that aids in interpreting the results from the impact evaluation of the MBVF program. First, Kaboski and Townsend (2012) find increases in consumption and income growth with no effect on investment. Ongoing work by Breza et al. (2018) document heterogeneous effects of credit from the MBVF on investment, driven by heterogeneity in productivity. My results provide a bridge between these studies by showing that credit was inefficiently allocated and documenting the mechanisms leading to that allocation. Second, other studies analyzing whether the program reached poor households suggest that resources were directed towards the poor, based on inter-village comparisons (Haughton et al., 2014; Menkhoff and Rungruxsirivorn, 2011). By using socioeconomic networks data, the results from this paper suggest that cross-village comparisons hide substantial asymmetries in access to resources from the program, which only a detailed intra-village analysis is able to capture.

References

Acemoglu, D. (2010). Theory, general equilibrium, and political economy in development economics. *The Journal of Economic Perspectives 24*(3), 17–32.

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6), 2411–2451.

Alatas, V., A. Banerjee, A. G. Chandrasekhar, R. Hanna, and B. A. Olken (2012, August). Network structure and the aggregation of information: Theory and evidence from indonesia. Working Paper 18351, National Bureau of Economic Research.

Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, R. Purnamasari, and M. Wai-Poi (2016). Self-targeting: Evidence from a field experiment in indonesia. *Journal of Political Economy 124*(2), 371–427.

Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, and J. Tobias (2012, June). Targeting the poor: Evidence from a field experiment in indonesia. *American Economic Review 102*(4), 1206–40.

Anderson, S., P. Francois, and A. Kotwal (2015, June). Clientelism in indian villages. *American Economic Review 105*(6), 1780–1816.

Angelucci, M. and G. De Giorgi (2009). Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption? *American Economic Review 99*(1), 486–508.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies 58*(2), 277–297.

Asher, S. and P. Novosad (2017, January). Politics and local economic growth: Evidence from india. *American Economic Journal: Applied Economics 9*(1), 229–73.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013). The diffusion of microfinance. *Science 341*(6144).

Banerjee, A., E. Duflo, R. Glennerster, and C. Kinnan (2015, January). The miracle of microfinance? evidence from a randomized evaluation. *American Economic Journal: Applied Economics 7*(1), 22–53.

Banerjee, A., D. Karlan, and J. Zinman (2015, January). Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics 7*(1), 1–21.

Banerjee, A. V. and E. Duflo (2010, Summer). Giving Credit Where It Is Due. *Journal of Economic Perspectives 24*(3), 61–80.

Bardhan, P. and D. Mookherjee (2005). Decentralizing antipoverty program delivery in developing countries. *Journal of Public Economics 89*(4), 675 – 704. Cornell - ISPE Conference on Public Finance and Development.

Bardhan, P. and D. Mookherjee (2006a). Decentralisation and accountability in infrastructure delivery in developing countries*. *The Economic Journal 116*(508), 101–127.

Bardhan, P. and D. Mookherjee (2006b). Pro-poor targeting and accountability of local governments in west bengal. *Journal of Development Economics 79*(2), 303 – 327. Special Issue in honor of Pranab Bardhan.

Basurto, P. M., P. Dupas, and J. Robinson (2017, May). Decentralization and Efficiency of Subsidy Targeting: Evidence from Chiefs in Rural Malawi. NBER Working Papers 23383, National Bureau of Economic Research, Inc.

Beaman, L., D. Karlan, B. Thuysbaert, and C. Udry (2014, August). Self-selection into credit markets: Evidence from agriculture in mali. Working Paper 20387, National Bureau of Economic Research.

Björkegren, D. and D. Grissen (2018). The potential of digital credit to bank the poor. *AEA Papers and Proceedings 108*, 68–71.

Blundell, R. and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics 87*(1), 115 – 143.

Blundell, R. and S. Bond (2000). Gmm estimation with persistent panel data: an application to production functions. *Econometric Reviews 19*(3), 321–340.

Boonperm, J., J. Haughton, and S. R. Khandker (2013). Does the village fund matter in thailand? evaluating the impact on incomes and spending. *Journal of Asian Economics 25*, 3 – 16.

Breza, E., R. Townsend, and D. Vera-Cossio (2018). Access to credit and productivity: Evidence from thai villages. Technical report.

Cameron, C. A. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–372.

Casey, K. (2018). Radical decentralization: Does community-driven development work? *Annual Review of Economics 10*(1), null.

Chandrasekhar, A. G. and R. Lewis (2017). Econometrics of sampled networks. Technical report.

Coleman, B. E. (2006). Microfinance in northeast thailand: Who benefits and how much? *World Development 34*(9), 1612–1638.

Crépon, B., F. Devoto, E. Duflo, and W. Parienté (2015, January). Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in morocco. *American Economic Journal: Applied Economics 7*(1), 123–50.

Deininger, K. (2013). Evaluating program impacts on mature self-help groups in india. *World Bank Economic Review 27*(2), 272–296.

Foster, A. D. and M. R. Rosenzweig (1996). Comparative advantage, information and the allocation of workers to tasks: Evidence from an agricultural labour market. *The Review of Economic Studies 63*(3), 347–374.

Galasso, E. and M. Ravallion (2005). Decentralized targeting of an antipoverty program. *Journal of Public Economics 89*(4), 705 – 727. Cornell - ISPE Conference on Public Finance and Development.

Government of Thailand (2004). Act on national village and urban community fund (b.e. 2547). *Royal Thai Government Gazette 59*(9), 442–455.

Greaney, B. P., J. P. Kaboski, and E. V. Leemput (2016). Can Self-Help Groups Really Be "Self-Help"? *Review of Economic Studies 83*(4), 1614–1644.

Hart, O., A. Shleifer, and R. W. Vishny (1997). The proper scope of government: Theory and an application to prisons. *Quarterly Journal of Economics 112*(4), 1127–1161. Reprinted in Michael A. Crew and David Parker, eds., Developments in the Economics of Privatization and Regulation, Edward Elgar Publishing Inc., 2008.

Haselmann, R., D. Schoenherr, and V. Vig (2017). Rent-seeking in elite networks. Technical report, Forthcoming, Journal of Political Economy.

Haughton, J., S. R. Khandker, and P. Rukumnuaykit (2014). Microcredit on a large scale: Appraising the thailand village fund. *Asian Economic Journal 28*(4), 363–388.

HOCHBERG, Y. (1988). A sharper bonferroni procedure for multiple tests of significance. *Biometrika 75*(4), 800–802.

Hoffmann, V., V. Rao, V. Surendra, and U. Datta (2017). Relief from usury: Impact of a community-based microcredit program in rural india. Technical Report 8021, The World Bank.

Holmstrom, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization 7*, 24–52.

Hussam, R., N. Rigol, and B. Roth (2017). Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design In The Field. Technical report.

Iyer, R., A. I. Khwaja, E. F. P. Luttmer, and K. Shue (2016). Screening peers softly: Inferring the quality of small borrowers. *Management Science 62*(6), 1554–1577.

Kaboski, J. P. and R. M. Townsend (2005). Policies and impact: An analysis of village-level microfinance institutions. *Journal of the European Economic Association 3*(1), 1–50.

Kaboski, J. P. and R. M. Townsend (2011). A structural evaluation of a large-scale quasi-experimental microfinance initiative. *Econometrica 79*(5), 1357–1406.

Kaboski, J. P. and R. M. Townsend (2012, April). The impact of credit on village economies. *American Economic Journal: Applied Economics 4*(2), 98–133.

Karlan, D., B. Savonitto, B. Thuysbaert, and C. Udry (2017). Impact of savings groups on the lives of the poor. *Proceedings of the National Academy of Sciences 114*(12), 3079–3084.

Karlan, D. and B. Thuysbaert (2016). Targeting ultra-poor households in honduras and peru. *The World Bank Economic Review*, lhw036.

Khwaja, A. I. and A. Mian (2005). Do lenders favor politically connected firms? rent provision in an emerging financial market. *The Quarterly Journal of Economics 120*(4), 1371–1411.

Kinnan, C. and R. Townsend (2012). Kinship and financial networks, formal financial access, and risk reduction. *The American Economic Review*, 289–293.

Ksoll, C., H. B. Lilleør, J. H. Lønborg, and O. D. Rasmussen (2016). Impact of village savings and loan associations: Evidence from a cluster randomized trial. *Journal of Development Economics 120*(Supplement C), 70 – 85.

Kydland, F. E. and E. C. Prescott (1982). Time to build and aggregate fluctuations. *Econometrica 50*(6), 1345–1370.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies 70*(2), 317.

Limsombunc, V., C. Gan, and M. Lee (2005, 08). An analysis of credit scoring for agricultural loans in thailand. *2*, 1198–1205.

Mabry, B. D. (1979). Peasant economic behaviour in thailand. *Journal of Southeast Asian Studies 10*(2), 400–419.

Mansuri, G. and V. Rao (2004). Community-based and -driven development: A critical review. *The World Bank Research Observer 19*(1), 1–39.

Menkhoff, L. and O. Rungruxsirivorn (2011). Do village funds improve access to finance? evidence from thailand. *World Development 39*(1), 110 – 122.

Moerman, M. (1969). A thai village headman as a synaptic leader. *The Journal of Asian Studies 28*(3), 535–549.

Morduch, J. (1999, December). The microfinance promise. *Journal of Economic Literature 37*(4), 1569–1614.

Muralidharan, K., P. Niehaus, and S. Sukhtankar (2017). General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India. Technical report, University of California, San Diego.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Pasuk, P. and C. J. Baker (2004). *Thaksin : the business of politics in Thailand* (1st ed. ed.). Silkworm Books Bangkok.

Samphantharak, K. and R. M. Townsend (2010, December). *Households as Corporate Firms*. Number 9780521195829 in Cambridge Books. Cambridge University Press.

Schreiner, M. (2000). Credit scoring for microfinance: Can it work? *Journal of Microfinance 2*(2).

Shaban, R. A. (1987). Testing between competing models of sharecropping. *Journal of Political Economy 95*(5), 893–920.

Shenoy, A. (2017a). Estimating the production function when firms are constrained. Working paper, University of California, Santa Cruz.

Shenoy, A. (2017b). Market failures and misallocation. *Journal of Development Economics 128*(Supplement C), 65 – 80.

Stoeffler, Q., B. Mills, and C. del Ninno (2016). Reaching the poor: Cash transfer program targeting in cameroon. *World Development 83*, 244 – 263.

Townsend, R. M. (2014). Townsend thai project monthly survey (1-172) initial release.

10 Figures

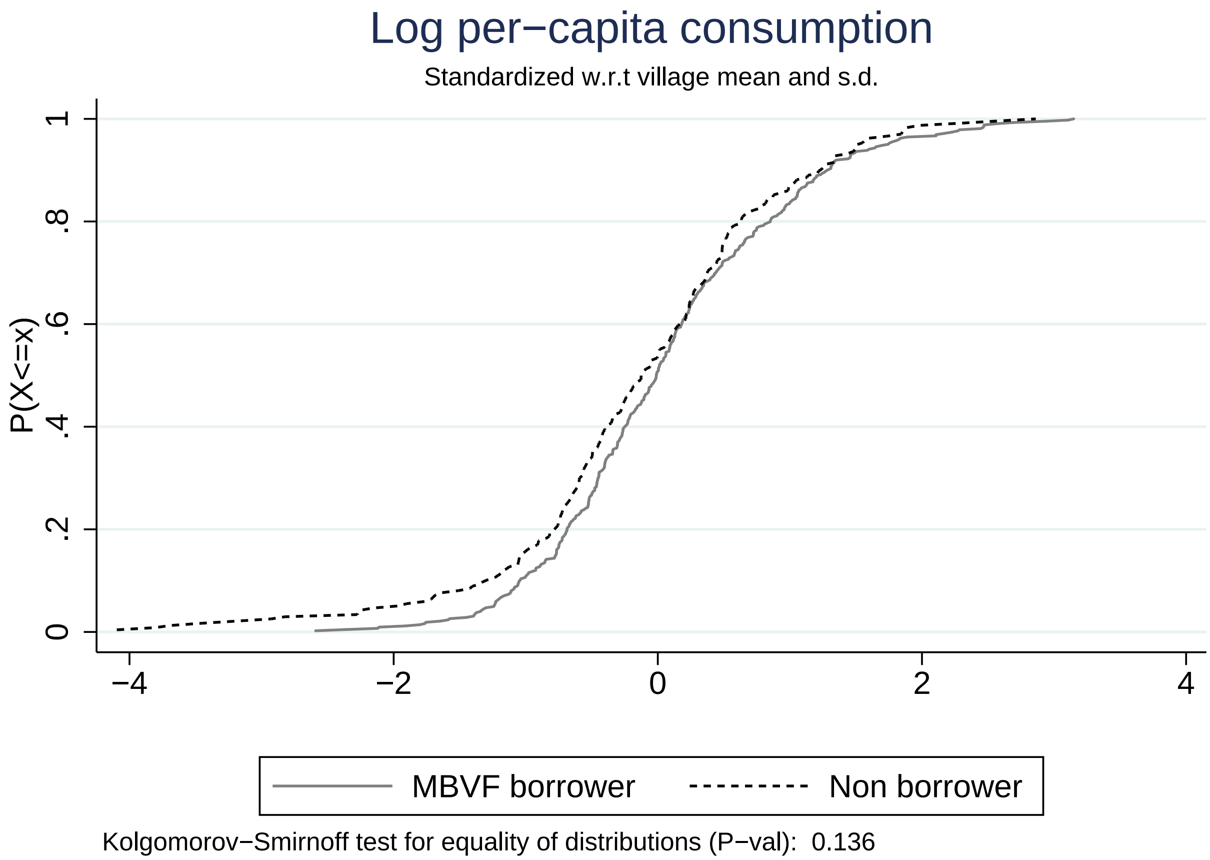


Figure 1: Cumulative distribution function of baseline log per-capita consumption

Note: The figure plots the cumulative distribution function (CDF) of log per-capita consumption, measured at baseline, for households with access to credit from the program (59%) and households who didn't obtain credit from the program (41%) during the first two years of its implementation. Per-capita consumption is measured as the total per-capita expenditure in consumption goods purchased outside the household plus the sales value of self-consumption items. The measure corresponds to the 12 months preceding the introduction of the program,and is standardized with respect to the village mean in order to perform within village comparisons.

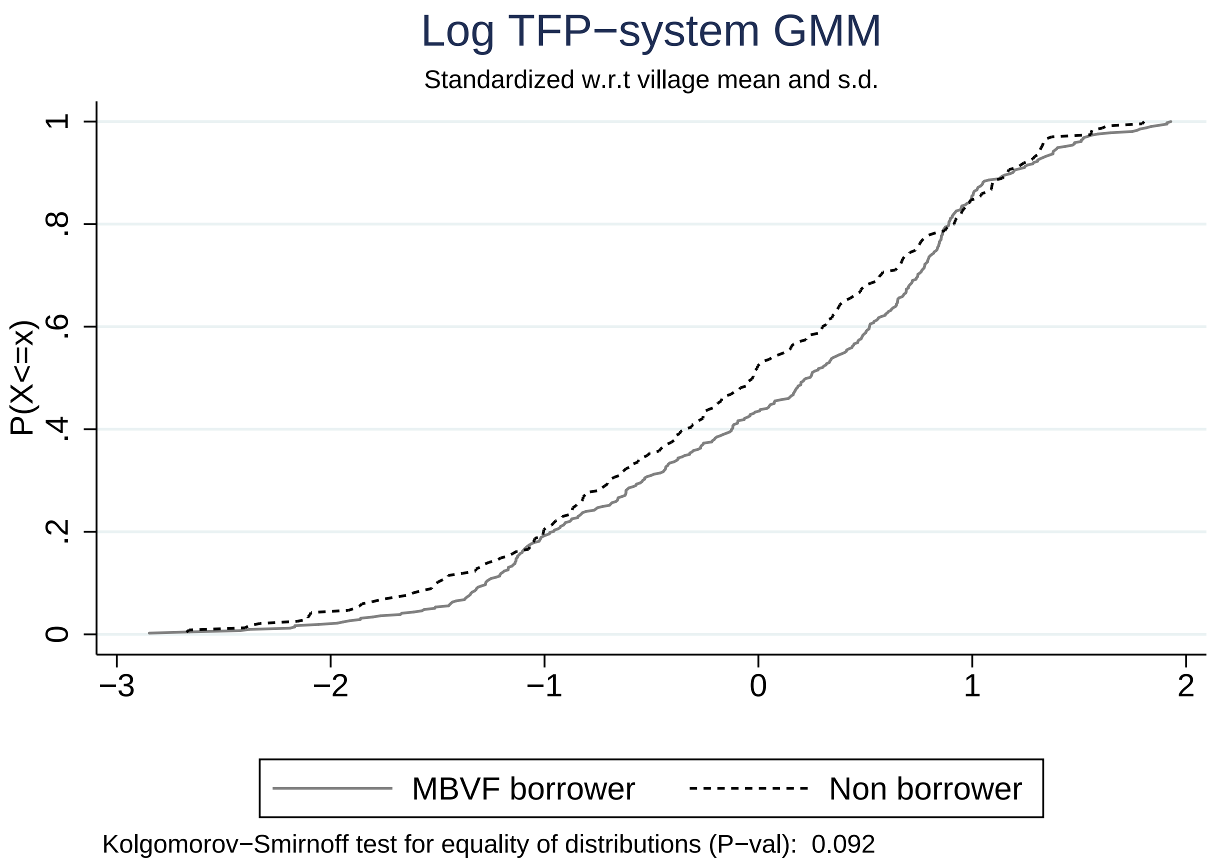
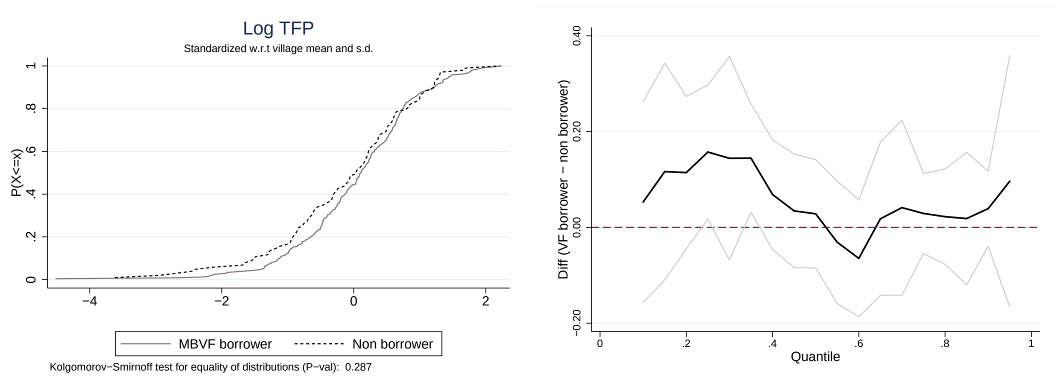
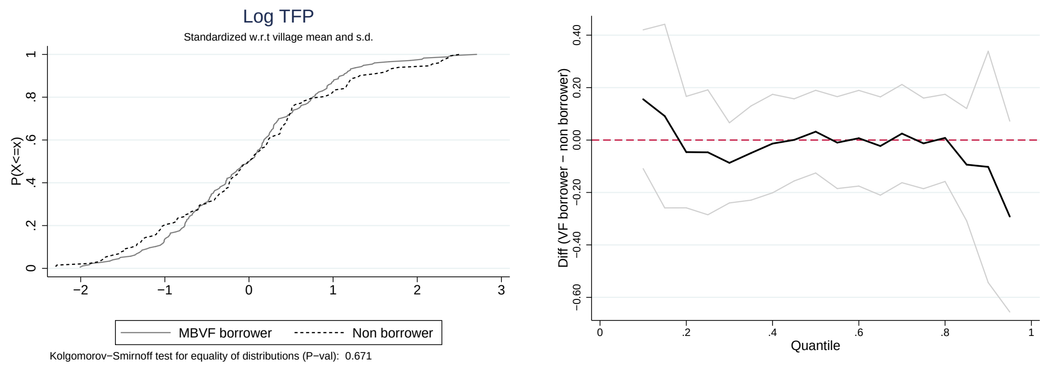


Figure 2: Cumulative distribution function of baseline household TFP by program participation

Note: The figure plots the cumulative distribution function (CDF) of baseline TFP, standardized with respect to the village mean and standard deviation, for households with access to credit from the program (59%) and households who didn't obtain credit from the program (41%) during the first two years of its implementation.



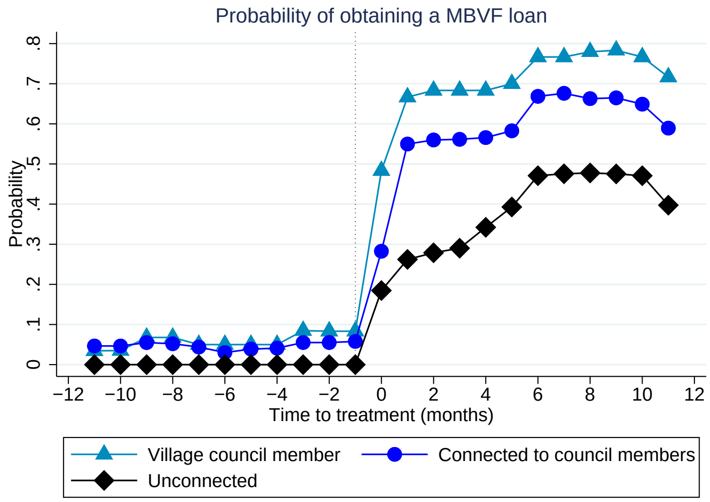
(a) Farm



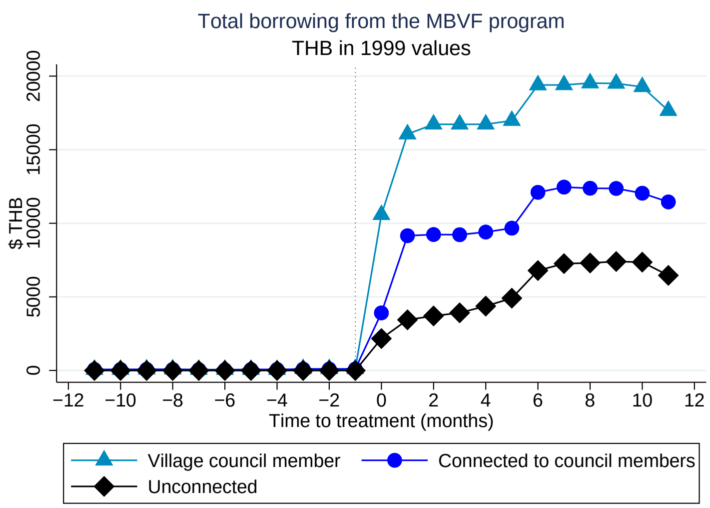
(b) Non Farm

Figure 3: Cumulative distribution function of baseline household TFP (left) and differences in quantiles (right) by program participation

Note: The figures on the left plot the cumulative distribution function (CDF) of baseline TFP for households with access to credit from the program (59%) and households who didn't obtain credit from the program (41%) during the first two years of its implementation. The figures on the right present the associated quantile differences estimated through quantile regressions after controlling for pre-program shocks and rainfall. Both measures are standardized with respect to the village mean in order to perform within village comparisons. Panels a) and b) report results for households with farm oriented operations and households with off-farm oriented operations. Farm sector: households who exhibited a pre-program average share of farm income larger than 0.5. Farm income includes cultivation, livestock and fishing and shrimping. Off-farm households are households whose baseline share of off-farm income exceeded 0.5. Off-farm activities include labor provision and off-farm family businesses.



(a) Probability of borrowing



(b) Total credit

Figure 4: Access to credit from the MBVF program and connections with the Village Council

Note: The figure depicts the probability of holding an outstanding loan from the the Village Fund program (top panel), and the average gross stock of debt from the program (bottom panel) for the 12 months preceding and following the implementation of the program. Each symbol denotes the mean for each category in a given month. The dotted line denotes the period preceding the release of the program's funds $\tau_{v,t} = -1$. Village council member: Households in which at least one member is either the village head or on the village council during pre-program periods. Connected to council members: households who reported having any socioeconomic interaction or direct kin relations with council members during the survey waves preceding the release of the funds from the program. Unconnected: households without any direct connection with members of the village council.

.

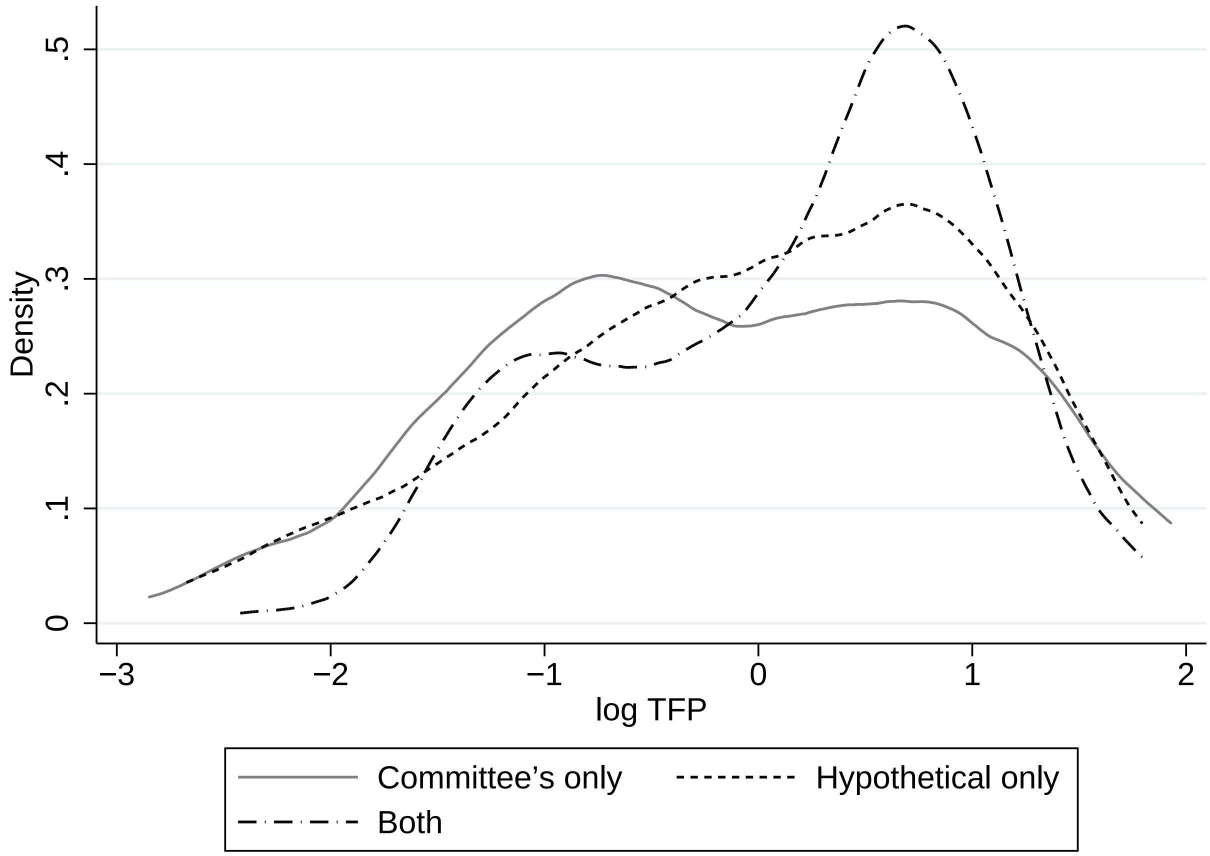


Figure 5: Distribution of baseline household TFP by program participation and eligibility based on the misspayment-risk hypothetical criterion

Note: The figure density estimates corresponding to log TFP, standardized with respect to the village mean and standard deviation, for households with access to credit from the program but would be ineligible under the hypothetical criterion, households that did not borrow but would be eligible under the hypothetical criterion, and households who borrowed from the program and would have been eligible by the hypothetical criterion. Program participation is measured as an indicator of whether a household borrowed from the program within the first two years of program implementation.

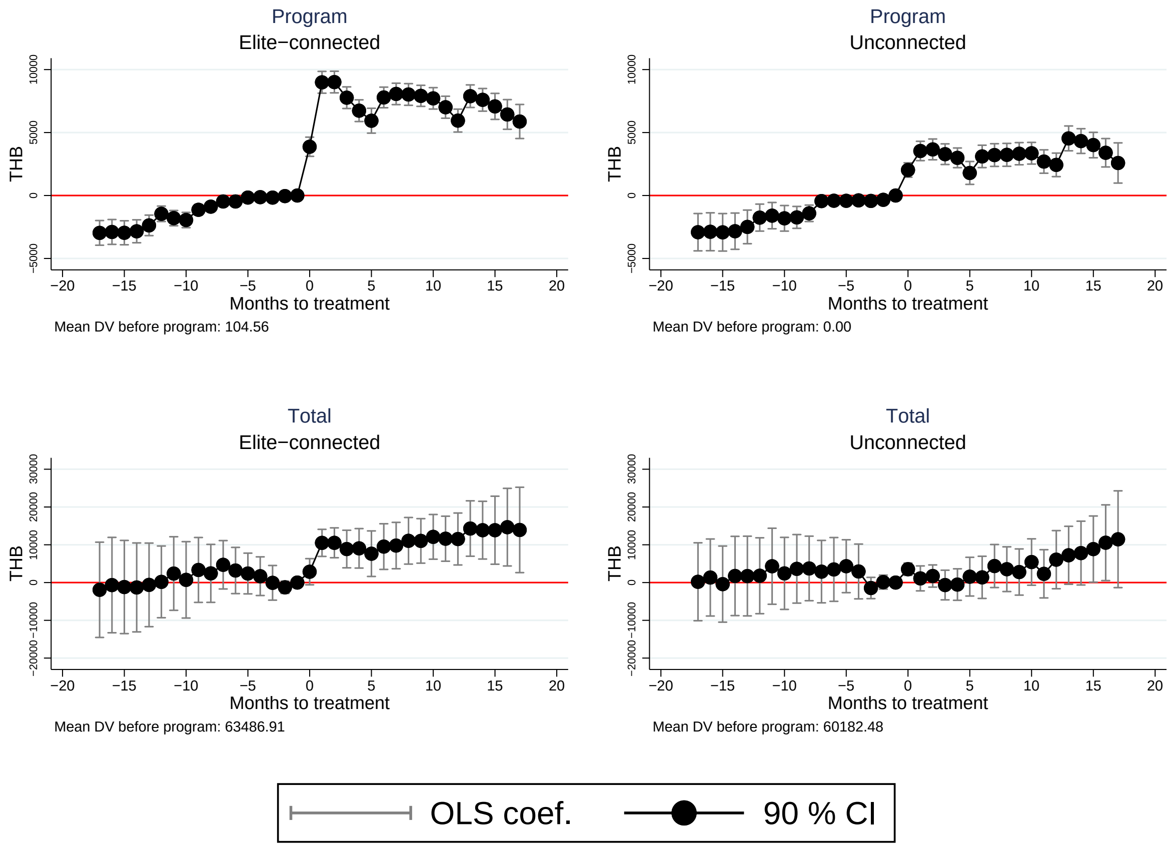


Figure 6: Short-term effects on program and total credit by elite-connectedness

Note: The figure depicts OLS point estimates from a flexible difference-in-difference model following equation (8). Each dependent variable was regressed on household fixed effects, calendar month and year fixed effects, an a set of indicators that denote time to treatment. Each dot represents the coefficient associated with each of these indicators. The base category corresponds to the period preceding the first month of operation of the fund: $\tau_{vt} = -1$. Estimations were performed using all the available observations for the 18 months before and after the rollout of the program in each village. Confidence intervals are constructed using standard errors clustered at the household level, to account for serial correlation.

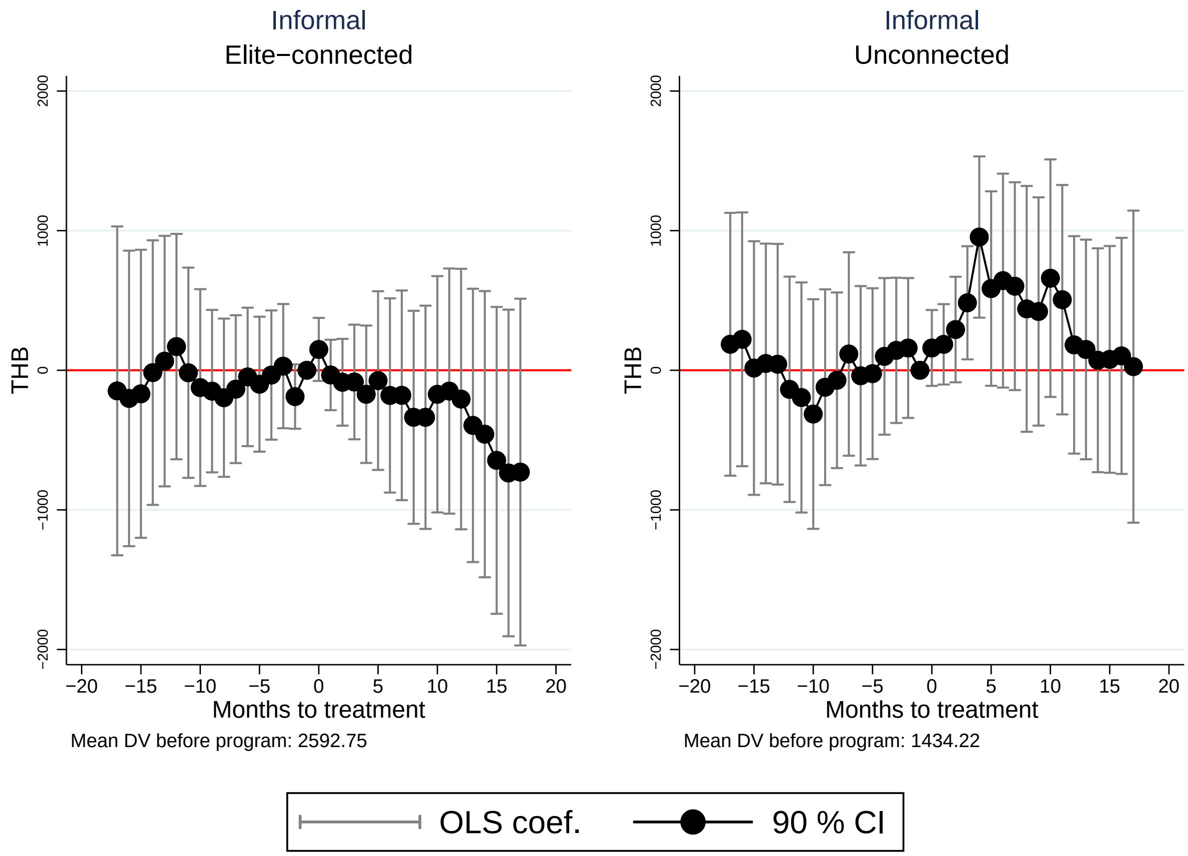


Figure 7: Short-term effects on credit from informal leners by elite-connectedness

Note: The figure depicts OLS point estimates from a flexible difference-in-difference model following equation (8). The left-hand panel presents estimates for loans from relatives, while the right-hand panel shows estimates for loans from local non-relative lenders. Each dependent variable was regressed on household fixed effects, calendar month and year fixed effects, and a set of indicators that denote time to treatment. Each dot represents the coefficient associated with each of these indicators. The base category corresponds to the period preceding the first month of operation of the fund: $\tau_{vt} = -1$. Estimations were performed using all the available observations for the 18 months before and after the rollout of the program in each village. Confidence intervals are constructed using standard errors clustered at the household level, to account for serial correlation. Informal lenders include personal lenders and relatives in the villages.

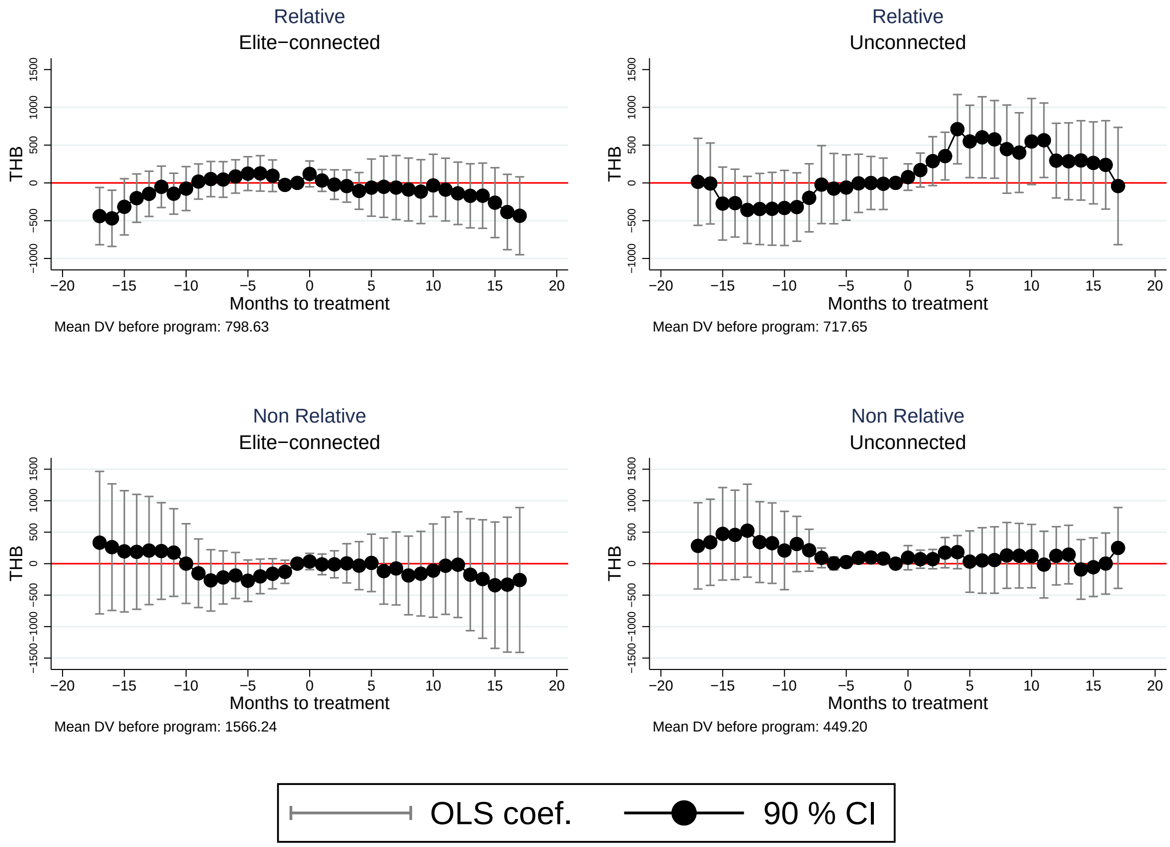


Figure 8: Short-term effects on borrowing from relatives and non-relatives

Note: The figure depicts OLS point estimates from a flexible difference-in-difference model following equation (8). The left-hand panel presents estimates for loans from relatives, while the right-hand panel shows estimates for loans from local non-relative lenders. Each dependent variable was regressed on household fixed effects, calendar month and year fixed effects, and a set of indicators that denote time to treatment. Each dot represents the coefficient associated with each of these indicators. The base category corresponds to the period preceding the first month of operation of the fund: $\tau_{vt} = -1$. Estimations were performed using all the available observations for the 18 months before and after the rollout of the program in each village. Confidence intervals are constructed using standard errors clustered at the household level, to account for serial correlation. Informal lenders include personal lenders and relatives in the villages.

11 Tables

Table 1: Summary statistics: Loan characteristics in the village financial system

	Formal		Quasi-formal		Personal lender		Relative		Village Fund	
Panel A: Distribution of loan by type of lender										
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Number of loans (Share)	0.28	0.12	0.22	0.33	0.32	0.15	0.16	0.07	0.02	0.32
Total amount (Share)	0.45	0.31	0.28	0.22	0.17	0.15	0.09	0.05	0.00	0.26
Panel B: Loan Characteristics										
	Formal		Quasi-formal		Personal lender		Relative		Village Fund	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Cosigner (Indicator)	0.63	0.48	0.34	0.48	0.01	0.09	0.00	0.05	0.95	0.21
Collateral (Indicator)	0.29	0.45	0.18	0.38	0.06	0.24	0.04	0.19	0.00	0.04
Group loan (Indicator)	0.63	0.48	0.01	0.07	0.00	0.04	0.00	0.00	0.00	0.04
Size (TBH)	40614	52595	31584	60850	13627	30759	14287	34157	14579	8493
Term (months)	17	16	13	13	8	7	8	7	12	3
Interest rate (% annual)	11%	22%	22%	79%	30%	34%	19%	28%	7%	24%
Panel C: Loan Performance										
	Formal		Quasi-formal		Personal lender		Relative		Village Fund	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Delinquency (share)	0.04	0.15	0.02	0.10	0.02	0.09	0.01	0.05	0.01	0.06
Recovery rate	0.97	0.14	0.93	0.20	0.97	0.15	0.98	0.11	1.00	0.02
Loan required a term extension (share)	0.40	0.49	0.54	0.50	0.43	0.50	0.51	0.50	0.35	0.48
Ex-post Internal rate of return (% annual)	11.69%	23.82%	15.06%	54.46%	43.30%	112.63%	18.61%	63.85%	6.80%	9.80%

Source: Author's own calculations using the the Townsend-Thai monthly dataset. *Note:* Panel A presents portfolio shares by lender for the two-years preceding and following the rollout of the program. Panels B and C present summary statistics for a sample of all loans that have reached maturity in the dataset and were obtained from January 1999 to December 2001, with the exemption of loans from the Village Fund program (loans obtained between 2001-2003). Loans that reached maturity include loans that were fully repaid and defaulted loans. Statistics are presented by type of lender for comparison. Formal loans: Includes loans from commercial banks and the Bank for Agriculture and Agricultural Cooperatives (BAAC), loans from the latter source represent 98% of formal loans. Quasi-formal lenders include production credit groups, cooperatives, womens's group and other loans from village organizations that keep records of their operations but do not have a physical location. Interest rates are nominal. Initial interest rates are self reported and converted to annual values by multiplying them by 12 or 52, in the case of monthly and weekly rates, respectively. Internal rates of return are computed using the entire stream of payments over the life of the loan.

Table 2: Summary statistics: household characteristics

Variable	N	Mean	S.D.	10th percentile	90th percentile
Panel A: Demographic characteristics					
Age (household head)	656	52.77	13.80	35.13	71.88
Household head is a male	673	0.76	0.43	0	1
Education (household head)	656	4.29	2.39	2	7
Number of adults	673	2.46	1.24	1	4
Number of elderly (> 65)	673	0.44	0.67	0.00	1.42
Number of children (< 15)	673	1.23	1.07	0.00	2.67
Age (household average)	673	35.89	13.81	20.65	55.71
Education (household average)	673	4.58	1.95	2.50	7.06
Panel B: Borrowing and liquidity					
Household borrowed (any source)	710	0.53	0.50	0	1
Household borrowed from institutional lender	710	0.41	0.49	0	1
Share of loans with delinquent payments	544	0.06	0.16	0	0.25
Household ever missed a payment	544	0.22	0.42	0	1
Share of loans with term extensions	544	0.36	0.31	0	0.80
Household ever extended a loan term	544	0.74	0.44	0	1
Cash holdings as a share of total assets	688	0.28	0.26	0.03	0.70
Total liabilities as a share of total assets	688	0.11	0.19	0.00	0.26
Panel C: Household productive operations					
Household does not own land	710	0.18	0.38	0	1
Household owns non-farm business	710	0.14	0.34	0	1
Number of sources of revenues	710	3.54	1.51	1.00	5.00
Revenues from cultivation (share of total)	673	0.34	0.36	0.00	0.89
Revenues from livestock (share of total)	673	0.08	0.21	0.00	0.17
Revenues from fishing-shrimping (share of total)	673	0.06	0.18	0.00	0.09
Revenues from non-farm businesses (share of total)	673	0.11	0.26	0.00	0.50
Revenues from wage labor (share of total)	673	0.34	0.43	0.00	0.96
Panel D: Household monthly income and consumption					
Per-capita income (TBH)	673	2056.08	7009.84	-5.52	5117.01
Per-capita consumption (TBH)	673	1311.32	1203.75	541.67	2423.83
Per-capita food consumption (TBH)	673	594.00	321.71	321.55	976.30
Panel E: Social connections					
Number of direct relatives in the village	710	2.35	2.19	0	6
Number of links with other households (transactions)	710	11.59	8.94	2	24
Household is part of the Village Council	710	0.09	0.28	0	0
Number of links with Village Council members (transactions)	710	1.26	1.51	0	3
Any link with Village Council members	710	0.49	0.50	0	1

Note: The table reports summary statistics regarding household characteristics measured at baseline. The period of reference for the variables in Panels A-D is 2000, the year preceding the initial implementation of the program in the sample villages. The variables from Panel E as well as the indicators of whether a household ever missed a payment, and whether a household extended the repayment period of a loan are measured with respect to all the survey waves preceding the program implementation. The latter information is only available for a subset of households who reported borrowing from any source during the baseline periods.

Table 3: Estimates of factor elasticities

Panel A: Reduced-form estimates

	(1) Farm	(2) Non Farm
y_{t-1}	0.67***	0.75***
	(0.097)	(0.048)
k_t	0.32*	0.17
	(0.177)	(0.175)
k_{t-1}	-0.28	-0.08
	(0.175)	(0.171)
m_t	0.39***	0.33***
	(0.068)	(0.055)
m_{t-1}	-0.27***	-0.23***
	(0.049)	(0.045)
l_t	0.12*	0.31***
	(0.071)	(0.118)
l_{t-1}	-0.06	-0.30***
	(0.067)	(0.112)
Hansen stat	185.2	196.3
DF	170	170
P-val(Hansen)	0.201	0.0813

Panel B: Estimates of factor elasticities

	Farm	Non Farm
ρ	0.66***	0.72***
	(0.06)	(0.03)
β_k	0.15**	0.26**
	(0.06)	(0.09)
β_m	0.40***	0.33***
	(0.04)	(0.04)
β_l	0.14**	0.28***
	(0.04)	(0.08)
Obs	3584	2586
Returns to scale	0.69	0.87
Chi2 (RTS=1)	17.26	1.31
P-Val (RTS=1)	0.00	0.25
J-stat CFR-OMD	1.69	1.70
P-val (CFR-OMD)	0.64	0.64

Panel C: Household baseline TFP

	Farm	Non Farm
Mean	4.58	2.94
SD	0.62	0.84

Standard errors in parentheses

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: Panel A reports reduced-form "system - GMM" estimates using 3 to 5 lags of output, capital, non-labor inputs and labor in levels as instruments for current and lagged first differences. Standard errors are clustered at the household level. The variance-covariance matrix corresponding to the two-step GMM approach is corrected for small sample bias. Columns (2) and (3) report separate estimates for households in the farm sector and off-farm sector, respectively. Farm sector: households whose baseline farm-to-total revenues ratio exceeds 0.5. Off-farm sector: households whose baseline farm-to-total revenues ratio is below 0.5. Panel B reports factor elasticities corresponding to the gross-revenue function which were estimated through Optimal Minimum Distance (OMD) using the reduced-form variance-covariance matrix as weighting matrix. Panel C reports summary statistics corresponding to the predicted baseline household productivity.

Table 4: Targeting errors and gains from reallocation with respect to hypothetical repayment-score criterion

Panel A: Distribution of households by access to program credit and hypothetical eligibility based on delinquency risk

	N	Percentage
Group A: hhs who borrowed and would be eligible by the repayment-score criterion	281	39.6%
Group B: HHs who did not borrow and would be eligible by the repayment-score criterion	136	19.2%
Group C: HHs who borrowed but would be ineligible by the repayment-score criterion	144	20.3%
Group D: HHs who did not borrow and would be ineligible	149	21.0%

Panel B: Targeting errors with respect to hypothetical eligibility criterion

% of ineligible hh who borrowed (C / (A+ C))	33.9%
% of eligible hh who did not borrow (B/ (B+D))	47.7%

Panel C: Gains/losses from reallocation (From group C to group B)

	Delinquency risk*	PC consumption (logs)	TFP (logs)
Difference (Group B-C)	-0.172***	0.122*	0.140
	(0.055)	(0.070)	(0.129)
Implied reduction in delinquency risk - percentage points	-0.06		
Implied poverty targeting gain (% change in per-capita consumption)		-0.04	
Implied productivity gain (% change in TFP)			0.05

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: Panel A reports the distribution of households by program borrowing and eligibility under the hypothetical repayment score. Panel B reports ratios corresponding to the share of program borrowers that would have been ineligible under the hypothetical criterion (over-inclusion error), and the share of non borrowers who would have been eligible under the hypothetical repayment-based criterion (over-exclusion error). Panel C reports back-of-the-envelop calculations of the average gain or loss from reallocation resources from ineligible households to eligible households who did not borrow. The calculations imply scaling down the difference between ineligible borrowers and eligible non-borrowers by the inclusion error. Standard errors are clustered at the household level and are reported in parentheses.

Table 5: Baseline characteristics and program participation and eligibility under alternative criterion

Panel A: Correlates of probability of borrowing from Village Fund and baseline characteristics							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Access to institutional credit	0.342***					0.222***	0.174***
	(0.041)					(0.051)	(0.067)
Ever missed a payment		0.067					0.022
		(0.048)					(0.049)
Per-capita consupmtion (logs)			0.072**			0.051	0.042
			(0.028)			(0.032)	(0.037)
TFP (logs)				0.030		-0.012	-0.006
				(0.019)		(0.021)	(0.023)
Connected with Village Council					0.163***	0.103**	0.109**
					(0.043)	(0.046)	(0.053)
Observations	710	544	660	648	710	622	514
R-squared	0.180	0.082	0.109	0.097	0.111	0.175	0.138
Panel B: Correlates of probability of eligibility based on misspayment risk and baseline characteristics							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Access to institutional credit	0.133***					-0.002	0.019
	(0.043)					(0.045)	(0.062)
Pval (VF-RS)	0.000					0.001	0.041
Ever missed a payment		-0.146***					-0.132**
		(0.053)					(0.052)
Pval (VF-RS)		0.002					0.013
Per-capita consupmtion (logs)			0.138***			0.118***	0.082**
			(0.027)			(0.028)	(0.034)
Pval (VF-RS)			0.079			0.112	0.217
TFP (logs)				0.061***		0.043**	0.057***
				(0.020)		(0.020)	(0.022)
Pval (VF-RS)				0.271		0.058	0.052
Connected with Village Council					0.014	0.017	0.042
					(0.042)	(0.041)	(0.047)
Pval (VF-RS)					0.008	0.155	0.224
Observations	710	544	660	648	710	622	514
R-squared	0.099	0.083	0.126	0.104	0.086	0.280	0.260
Controls (shocks + demographics)	No	No	No	No	No	Yes	Yes

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: Panel A reports OLS coefficients of a regression of the probability of borrowing from the program during the first two years of its implementation on several baseline characteristics. Panel B reports OLS coefficients of a regression of the probability of being eligible under the hypothetical repayment-based criterion on several baseline characteristics. Standard errors are clustered at the household level. All regressions control for village fixed effects. Columns (2) and (7) report estimates using only the households who had pre-program credit history with any lender, either formal or informal. Panel B also reports P-values corresponding to the differences, within each column, between Panel A and Panel B. They are computed by jointly estimating both models using seemingly unrelated regressions.

Table 6: Probability of borrowing and connections with village council

	Borrowed from the program (dummy)					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Social Connections						
Connectedness with Village Council	0.163***	0.067	0.071			
	(0.043)	(0.046)	(0.056)			
Village Council member				0.330***	0.195***	0.176**
				(0.060)	(0.066)	(0.074)
Directly transacted with council member				0.166***	0.076	0.068
				(0.043)	(0.047)	(0.056)
First-degree relative to council member				-0.004	0.004	0.051
				(0.054)	(0.052)	(0.059)
Degree (count of links)		0.014***	0.006***		0.013***	0.005**
		(0.002)	(0.002)		(0.002)	(0.002)
Other baseline characteristics						
Access to institutional credit			0.151**			0.144**
			(0.068)			(0.068)
Ever missed a payment			0.016			0.009
			(0.050)			(0.050)
Log per-capita consumption			0.036			0.032
			(0.039)			(0.039)
log TFP			-0.003			-0.002
			(0.023)			(0.023)
Operation and health shocks						
Cultivation			0.009			0.010
			(0.015)			(0.015)
Livestock			0.019*			0.021*
			(0.012)			(0.012)
Non-agricultural business			-0.007			-0.005
			(0.011)			(0.011)
Health symptoms			-0.006***			-0.006***
			(0.002)			(0.002)
Observations	710	710	514	710	710	514
Control for demographics	NO	NO	YES	NO	NO	YES
R-squared	0.111	0.156	0.157	0.126	0.162	0.163

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: The table reports coefficients from regressions of the probability of obtaining a loan from the program on an indicator of whether a household is connected to the village council (Columns (1) to (3)) and on indicators of membership to the council and connectedness through transaction networks (Columns (4) to (6)). All regressions include village fixed effects. Columns (2),(3), (5) and (6) control for the number of links in the baseline transaction network. Columns (3) and (6) control fro demographic characteristics including average household age, average household years of schooling, number of working-age household members, and household head age, gender and schooling. Columns (3) and (6) report estimates only for the sub sample of 544 households with baseline credit history. Standard errors are clustered at the household level, and are reported in parentheses.

Table 7: Differences in loan outcomes by connections with village council member and lender

	Means				Difference (MBVF-CG)		Double difference		
	Connected (N=260)		Unconnected (N=75)		Connected (N=260)	Unconnected (N=75)	All borrowers (N=335)		
	MBVF	Credit groups (CG)	MBVF	Credit groups (CG)	(1)-(2)	(3)-(4)	(5)-(6)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Returns to the lender									
Ex post IRR (annual)	0.061	0.078	0.068	0.057	-0.018***	0.011	-0.027***	-0.021***	-0.027***
					(0.004)	(0.007)	(0.008)	(0.007)	(0.008)
Panel B: Loan outcomes									
Any delinquent payment	0.008	0.016	0.002	0.000	-0.004	0.003	-0.007	-0.005	-0.007
					(0.004)	(0.002)	(0.005)	(0.004)	(0.005)
Delinquent payments as a share of due payments	0.006	0.010	0.001	0.000	-0.002	0.001	-0.004	-0.003	-0.004
					(0.003)	(0.001)	(0.003)	(0.003)	(0.003)
Any loan extension	0.468	0.401	0.374	0.338	0.009	0.031	-0.023	-0.064	-0.021
					(0.022)	(0.041)	(0.043)	(0.057)	(0.043)
Panel C: Loan characteristics									
Initial interest rate (annual)	0.054	0.078	0.059	0.068	-0.019***	-0.007	-0.011***	-0.010**	-0.011**
					(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
Term (months)	11	12	11	12	-0.189	-1.134**	0.907	0.892	0.878
					(0.251)	(0.537)	(0.592)	(0.574)	(0.592)
Loan size (TBH-1999 prices)	15,118	4,108	11,670	3,664	10,778***	9,070***	1,587**	1,201.648*	n.a
					(405)	(652)	(654)	(688)	
Borrower fixed effect					YES	YES	YES	YES	YES
Lender fixed effect					NO	NO	YES	YES	YES
Village -year trends					YES	YES	YES	YES	YES
Weights for number of loans					NO	NO	NO	YES	NO
Weights for loan size					NO	NO	NO	NO	YES
Observations (loans)					5,274	1,404	6,741	6,741	6,741

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: The sample correspond to loans obtained after the rollout of the program that were fully repaid, reached maturity or were declared as defaulted on. The sample includes only loans belonging to households that borrowed both from the program and other community-based sources of credit. Columns (1) to (4) report means of loan characteristics and outcomes by type of borrower and lender. Columns (5) and (6) report differences in returns to loans from the program with respect to the comparison group by type of borrower. Columns (7) to (9) report double difference estimates following several specifications. Standard errors are two-way clustered by borrower and lender. Lenders include the 16 village funds in the sample as well as each of the local community-based lenders (production credit groups, women's groups, and other similar lenders.)

Table 8: Short-term effects of the program on informal credit

Panel A: Gross borrowing from informal lenders

	Connected			**Unconnected**		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Any informal	Relatives	Non-relatives	Any informal	Relatives	Non-relatives
$Post_{vt}$	176.622	-49.509	290.470*	584.114	483.738*	94.496
	(286.919)	(150.579)	(164.473)	(363.398)	(283.524)	(104.353)
	[0.544]	[0.808]	[0.080]	[0.280]	[0.088]	[0.848]
Observations	13,428	13,428	13,428	6,732	6,732	6,732
R-squared	0.770	0.689	0.782	0.636	0.587	0.592
P-val (Connected-Unconnected)	0.264	0.061	0.084			
Baseline DV mean	2328	816.7	1223	1544	800.1	450
# of households	373	373	373	187	187	187

Panel B: Probability of borrowing from informal lenders

	Connected			**Unconnected**		
	(1)	(2)	(3)	(4)	(5)	(6)
	Any informal	Relatives	Non-relatives	Any informal	Relatives	Non-relatives
$Post_{vt}$	-0.008	-0.002	-0.005	0.046	0.050**	0.002
	(0.014)	(0.011)	(0.012)	(0.023)	(0.020)	(0.013)
	[0.676]	[0.824]	[0.688]	[0.288]	[0.048]	[0.968]
Observations	13,428	13,428	13,428	6,732	6,732	6,732
R-squared	0.702	0.663	0.645	0.562	0.539	0.556
P-val (Connected-Unconnected)	0.160	0.052	0.700			
Baseline DV mean	0.167	0.0678	0.112	0.0915	0.0505	0.0460
# of households	373	373	373	187	187	187

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: The table presents difference-in-difference estimates of the short-run effect of the rollout of the program on borrowing from informal lenders, by connectedness with the local elites. Informal lenders include personal money lenders and relatives in the village. The reported coefficients correspond to OLS regressions of the respective dependent variables on whether the resources from the program were released in village v in month t, controlling for household fixed effects and calendar month and year fixed effects (see equation (9)). Estimations were performed using all the available observations for the 18 months before and after the rollout of the program in each village. Panel A shows results for total gross borrowing (winsorizing the top 1% of observations), and Panel B reports results for probability of holding a loan. Standard errors, presented in parentheses, are clustered at the household level to allow for flexible serial correlation. P-values that account for potential within village correlation are presented in brackets; they are computed using a wild bootstrap t-procedure to account for a reduced number of clusters (16) as in Cameron and Miller (2015). Connected: households who reported having any socioeconomic interaction or direct kin relations with council members during the survey waves preceding the release of the funds from the program. Unconnected: households without any direct connection with members of the village council.

A Appendix

A.1 Supplementary figures

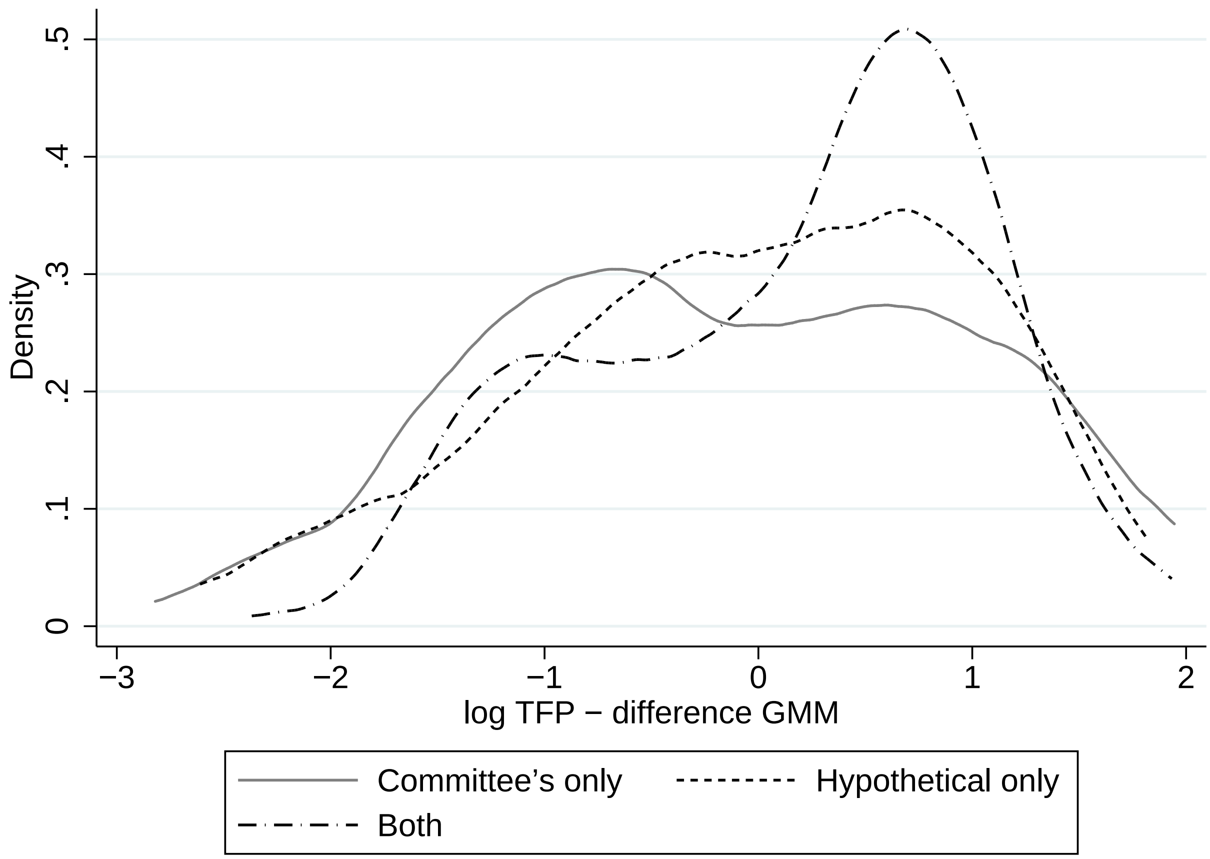


Figure A1: Distribution of baseline household TFP (difference-GMM) by program participation and eligibility based on the misspayment-risk hypothetical criterion

Note: The figure density estimates corresponding to log TFP, standardized with respect to the village mean and standard deviation, for households with access to credit from the program but would be ineligible under the hypothetical criterion, households that did not borrow but would be eligible under the hypothetical criterion, and households who borrowed from the program and would have been eligible by the hypothetical criterion. Program participation is measured as an indicator of whether a household borrowed from the program within the first two years of program implementation.

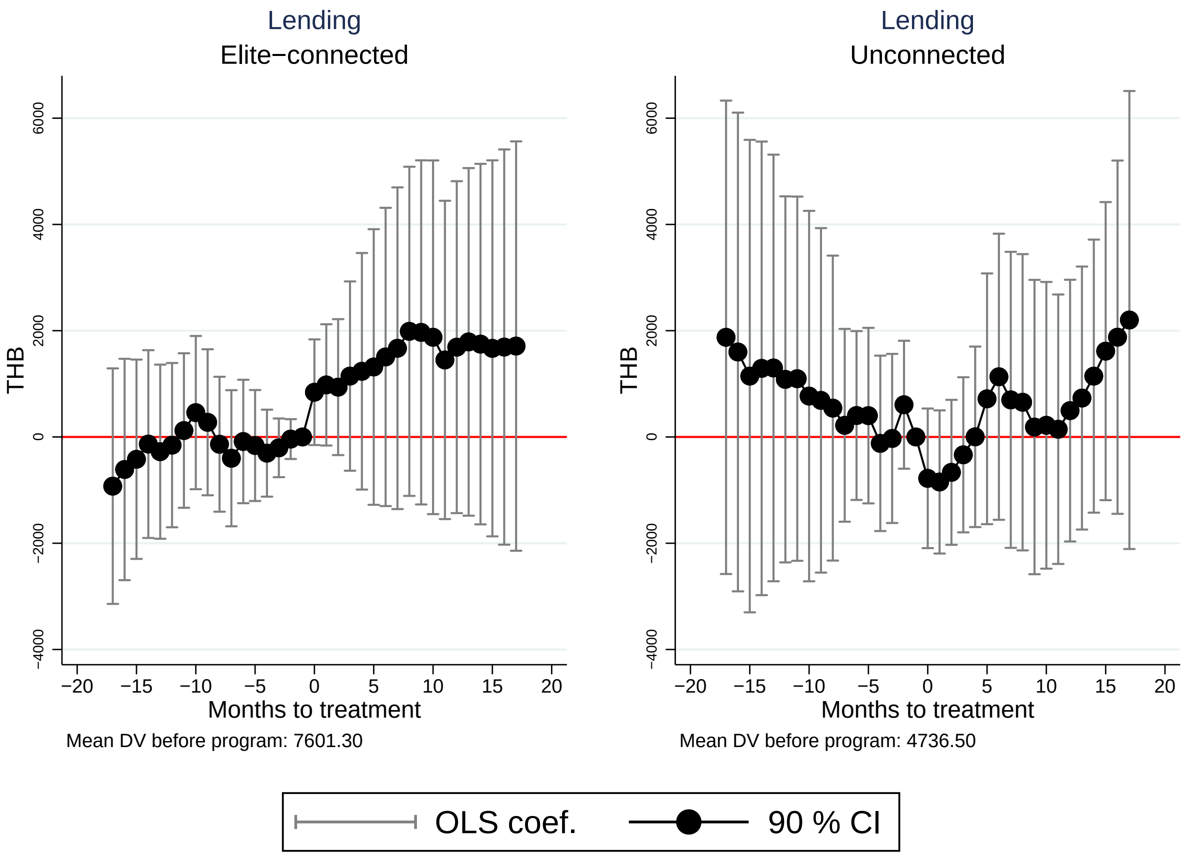


Figure A2: Short-term effects on lending to other households

Note: The figure depicts OLS point estimates from a flexible difference-in-difference model following equation (8). The left-hand panel presents estimates for loans from relatives, while the right-hand panel shows estimates for loans from local non-relative lenders. Each dependent variable was regressed on household fixed effects, calendar month and year fixed effects, and a set of indicators that denote time to treatment. Each dot represents the coefficient associated with each of these indicators. The base category corresponds to the period preceding the first month of operation of the fund: $\tau_{vt} = -1$. Estimations were performed using all the available observations for the 18 months before and after the rollout of the program in each village. Confidence intervals are constructed using standard errors clustered at the household level, to account for serial correlation.

A.2 Supplementary tables

Table A1: Correlates between current first-differences and lagged levels

VARIABLES	(1) $\Delta y_{i,t}$	(2) $\Delta k_{i,t}$	(3) $\Delta m_{i,t}$	(4) $\Delta l_{i,t}$
$y_{i,t-3}$	**-0.047***	0.016***	0.037	0.023**
	(0.016)	(0.005)	(0.024)	(0.011)
$k_{i,t-3}$	-0.001	**-0.026***	-0.005	0.009*
	(0.006)	**(0.004)**	(0.010)	(0.005)
$m_{i,t-3}$	0.029***	-0.004	**-0.062***	-0.009
	(0.009)	(0.003)	**(0.015)**	(0.006)
$l_{i,t-3}$	-0.025*	0.002	-0.004	**-0.058***
	(0.013)	(0.004)	(0.020)	**(0.011)**
Observations	6,275	6,307	6,211	6,272
R-squared	0.042	0.024	0.021	0.051
F-stat	5.214	13.60	9.455	10.29
p-val	0.000	0.000	0.000	0.000

*** p<0.01, ** p<0.05, * p<0.1

Note: The table reports OLS estimates of a regression of first differences of output, capital, non-labor inputs and labor on the 3rd lags of their respective levels. Regressions include 14 years of data for the households who are always observed in all the survey waves. All regressions control for year and village fixed effects, rainfall, the count of days in which any household member reported suffering health symptoms as well as the number of episodes of issues with household business operations. Standard errors are clustered at the household level to allow for flexible serial correlation and are reported in parenthesis.

Table A2: Baseline correlates of TFP estimates and demographic characteristics

	(1) All household	(2) Farm	(3) Non Farm
Age of household's head	0.000	-0.003	0.002
	(0.003)	(0.003)	(0.005)
Education of household's head	0.012	0.010	0.046**
	(0.019)	(0.019)	(0.020)
Head of household gender (male)	0.077	0.081	0.136
	(0.076)	(0.092)	(0.088)
Number of adults	0.089**	0.062**	0.073
	(0.035)	(0.031)	(0.049)
Number of elder	-0.184	-0.143	-0.229
	(0.171)	(0.178)	(0.236)
#of children under 5 at baseline	-0.008	0.020	-0.073
	(0.057)	(0.050)	(0.085)
Share of females in the household	-0.016	0.019	-0.009
	(0.053)	(0.053)	(0.067)
Average age in household	-0.005	-0.007*	0.001
	(0.003)	(0.004)	(0.005)
Average education level in household	-0.016	0.001	-0.017
	(0.024)	(0.023)	(0.028)
Count of health symptoms	0.006**	0.004*	0.005
	(0.003)	(0.003)	(0.004)
Count of shocks to non farm business	-0.022	-0.017	-0.014
	(0.015)	(0.014)	(0.016)
Count of shocks to livestock business	0.015	0.012	0.036
	(0.018)	(0.014)	(0.027)
Count of shocks to agriculture	-0.053**	-0.032	-0.057**
	(0.026)	(0.021)	(0.028)
Share of agricultural revenues	2.540***	0.638**	-0.367
	(0.326)	(0.280)	(1.032)
Share of agricultural revenues X rainfall	1.313	0.853**	-0.234
	(0.813)	(0.331)	(1.780)
Observations	637	333	304
R-Squared	0.571	0.401	0.465
Adjusted R-squared	0.548	0.334	0.400

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Note: The table presents baseline correlates of household TFP and baseline characteristics estimated through OLS. Column(1) includes all the households for which TFP was recovered. Columns (2) and (3) report correlates for the farm and non-farm sectors respectively. A household belongs to the farming sector if the baseline share of income from farm activities (agriculture, livestock, fishing and shrimping) exceeds 0.5. A household belongs to the non-farming sector if most of its baseline income comes from off-farm operations such as wage labor provision and off-farm family businesses. Standard errors are clustered at the individual level.

Table A3: Baseline characteristics and program participation

	OLS Coefficient	S.E. (Diff)	P-val	Hochberg Adj-Pval
Panel A - Program participation, risk and credit history				
	(1)	(2)	(3)	(4)
Leverage rate	-0.09	(0.07)	0.18	0.55
Income volatility (coef. of variation)	0.02**	(0.01)	0.02	0.09
Share of loans with delinquent payments	-0.00	(0.14)	0.99	0.99
Missed a payment (dummy)	0.07	(0.05)	0.16	0.65
Share of loans with term extensions	0.09	(0.07)	0.19	0.38
Extended loan (dummy)	0.15***	(0.05)	0.00	0.03
Panel B - Program participation and shocks				
	OLS Coefficient	S.E. (Diff)	P-val	Hochberg Adj-Pval
Health symptoms	0.00	(0.00)	0.48	0.96
Problems with cultivation operations	0.02	(0.01)	0.15	0.44
Problems with livestock operations	0.04***	(0.01)	0.00	0.00
Problems with business operations	0.00	(0.01)	0.70	0.70
Panel C - Program participation and alternative measures of TFP				
	OLS Coefficient	S.E. (Diff)	P-val	Hochberg Adj-Pval
log TFP (Difference GMM)	0.03*	(0.02)	0.09590724	0.19
Return over Assets	0.08	(0.07)	0.21601465	0.22
Panel D - Program participation and connections with local leaders				
	OLS Coefficient	S.E. (Diff)	P-val	Hochberg Adj-Pval
Inverse distance to council members	0.48***	(0.09)	0.00	0.00
# of links with council members	0.11***	(0.02)	0.00	0.00
Connectedness PCA index	0.11***	(0.02)	0.00	0.00

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: The table presents OLS coefficients from a regression of program participation on several baseline characteristics. Each row represents the coefficient of a separate regression. Column (4) reports p-values which are adjusted following the HOCHBERG (1988) step-up method to control the FWER across variables within each panel.

Table A4: Predictive model for loan delinquency

	(1) Coefficients-OLS		(2) Coefficients - Selected regressors	
VARIABLES	Coef.	SE	Coef.	SE
Household Financial characteristics				
Leverage (Total Liabilities/Assets)	0.055**	(0.022)	0.063***	(0.022)
Wealth (TBH M)	0.001	(0.001)	0.001	(0.001)
Asset turnover ratio	-0.001*	(0.000)	-0.000	(0.000)
Returns over asset ratio	0.001	(0.000)		
Previously borrowed from lender	0.008	(0.008)	0.010	(0.008)
# of outstanding loans	0.002	(0.002)	0.002	(0.002)
Houseold Demographic characteristics				
Avg hh age	-0.001**	(0.000)	-0.001**	(0.000)
Ave hh years of schooling	-0.002	(0.003)	-0.005*	(0.003)
Head's age	0.000	(0.000)		
Head's years of schooling	-0.004*	(0.002)		
Number of working age adults	-0.002	(0.004)		
Household head is a male	-0.012	(0.011)	-0.013	(0.010)
Loan Characteristics				
Cosigner	0.004	(0.013)	0.006	(0.012)
Collateral	-0.031*	(0.017)	-0.027*	(0.016)
Group loan	-0.046*	(0.024)	-0.044*	(0.025)
Loan term (base category: unsettled term)				
Very short term loan (less than 6 months)	0.036	(0.036)		
Short term loan (6-12 months)	0.023	(0.036)		
Long term (more than 12 months)	0.037	(0.037)	0.006	(0.009)
Loan size (base category: small loans)				
Midsize loan (TBH 10-20K)	0.051***	(0.014)	0.049***	(0.014)
Large loan (> TBH 20K)	0.076***	(0.014)	0.073***	(0.014)
Interest rate (base category: 0)				
< 5 % annual	-0.003	(0.036)	0.031***	(0.012)
5-10% annual	-0.045	(0.038)		
10-20% annual	-0.070*	(0.036)	-0.035***	(0.012)
>20% annual	-0.020	(0.036)	0.012	(0.011)
Lender type (base category: personal lenders)				
BAAC	0.075***	(0.022)	0.064***	(0.019)
PCG	0.001	(0.027)		
Commercial Bank	0.216***	(0.083)	0.202**	(0.083)
Cooperatives	0.028	(0.021)		
Other quasi-formal	0.014	(0.011)	0.009	(0.011)
Relatives	-0.017*	(0.010)	-0.018*	(0.009)
Constant	0.009	(0.078)	0.009	(0.077)
Observations	3,878		3,878	
Adjusted R-squared	0.070		0.069	

Robust standard errors in parentheses

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: Column (1) presents OLS estimates of the likelihood that a loan experienced at least one delinquent payment over its maturity period using a comprehensive set of borrower and lender characteristics as well as village and year fixed effects. Column (2) reports OLS estimates of a more parsimonious model for which the regressors were selected through a LASSO model of all the variables included in Column (1). The penalty parameter for the LASSO model was picked through 10-fold cross validation in order to minimize the out of sample mean squared error. The estimating sample includes all the loans that were active before the rollout of the program.

Table A5: Baseline correlates between connections with the Village Council and creditworthiness, and profitability

Panel A: Credit history, productivity and connections with Village Council

VARIABLES	(1) Institutional credit	(2) Income volatility	(3) Leverage ratio	(4) Delinquent loans (share)	(5) Missed a payment	(6) Term extensions (share)	(7) Loan term extension	(8) TFP
Connected with Village Council	0.193***	0.034	-0.013	-0.009	0.064*	-0.030	0.081*	0.072
	(0.041)	(0.029)	(0.020)	(0.018)	(0.038)	(0.035)	(0.047)	(0.091)
Observations	710	692	710	544	544	544	544	648
R-squared	0.225	0.020	0.095	0.095	0.119	0.084	0.143	0.193

Panel B: Credit history and productivity by type of connection with Village Council

VARIABLES	(1) Institutional credit			(2) Delinquent loans (share)	(3) Missed a payment	(4) Term extensions (share)	(5) Loan term extension	(6) TFP
Village council member	0.365***	0.033	-0.031	0.028	0.175**	-0.005	0.132**	0.066
	(0.050)	(0.041)	(0.027)	(0.029)	(0.068)	(0.047)	(0.064)	(0.152)
Direct transactions with council member	0.179***	0.045	-0.009	-0.013	0.056	-0.033	0.084*	0.064
	(0.042)	(0.045)	(0.020)	(0.019)	(0.040)	(0.036)	(0.049)	(0.097)
First-degree relative with council member	-0.051	-0.060	-0.008	-0.012	-0.049	-0.007	-0.059	0.051
	(0.050)	(0.065)	(0.027)	(0.016)	(0.054)	(0.036)	(0.053)	(0.117)
Base category mean: unconnected	0.45	0.02	0.12	0.06	0.13	0.33	0.61	3.71
Observations	710	692	710	544	544	544	544	648
R-squared	0.239	0.021	0.095	0.102	0.130	0.085	0.147	0.193

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: The table presents OLS coefficients from a regression of baseline characteristics on different measures of connectedness with the village council.

Table A6: Short-term effects on program and total credit

Panel A: Effects on credit from the program

	Connected		Unconnected	
VARIABLES	(1) Any loan	(2) Total borrowing	(3) Any loan	(4) Total borrowing
$Post_{vt}$	0.400***	7,112.051***	0.239***	2,690.714***
	(0.027)	(553.383)	(0.032)	(498.112)
	[0.004]	[0.000]	[0.000]	[0.000]
Observations	13,428	13,428	6,732	6,732
R-squared	0.629	0.627	0.561	0.525
P-val (Connected-Unconnected)	0.000	0.000		
# households	373	373	187	187

Panel B: Effects on total credit

	Connected		Unconnected	
VARIABLES	(1) Any loan	(2) Total borrowing	(3) Any loan	(4) Total borrowing
$Post_{vt}$	0.065***	4,262.577	0.101**	-3,690.733
	(0.015)	(2,760.086)	(0.029)	(3,581.866)
	[0.000]	[0.212]	[0.040]	[0.388]
Observations	13,428	13,428	6,732	6,732
R-squared	0.627	0.842	0.645	0.866
P-val (Connected-Unconnected)	0.300	0.160		
Baseline DV mean	0.743	62683		
# households	373	373	187	187

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: The table presents difference-in-difference estimates of the short-run effect of the rollout of the program on program borrowing (Panel A) and total borrowing (Panel B), by connectedness with the local elites. The reported coefficients correspond to OLS regressions of the respective dependent variables on whether the resources from the program were released in village v in month t, controlling for household fixed effects and calendar month and year fixed effects (see equation (9)). Estimations were performed using all the available observations for the 18 months before and after the rollout of the program in each village. Standard errors, presented in parentheses, are clustered at the household level to allow for flexible serial correlation. P-values that account for potential within village correlation are presented in brackets; they are computed using a wild bootstrap t-procedure to account for a reduced number of clusters (16) as in Cameron and Miller (2015). Connected: households who reported having any socioeconomic interaction or direct kin relations with council members during the survey waves preceding the release of the funds from the program. Unconnected: households without any direct connection with members of the village council.

Table A7: Short-term effects of the program on lending to other households

	Connected		**Unconnected**	
	(1)	(2)	(3)	(4)
VARIABLES	Any lending	Total lending	Any lending	Total lending
$Post_{vt}$	0.032**	709.325	0.013	-891.495
	(0.014)	(891.324)	(0.017)	(1,061.214)
	[0.032]	[0.416]	[0.420]	[0.216]
Observations	13,428	13,428	6,732	6,732
R-squared	0.783	0.900	0.783	0.739
Baseline DV mean	0.203	7377	0.145	5118
# households	373	373	187	187

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: The table presents difference-in-difference estimates of the short-run effect of the rollout of the program on lending to other households, by connectedness with the local elites. The reported coefficients correspond to OLS regressions of the respective dependent variables on whether the resources from the program were released in village v in month t, controlling for household fixed effects and calendar month and year fixed effects (see equation (9)). Estimations were performed using all the available observations for the 18 months before and after the rollout of the program in each village. Standard errors, presented in parentheses, are clustered at the household level to allow for flexible serial correlation. P-values that account for potential within village correlation are presented in brackets; they are computed using a wild bootstrap t-procedure to account for a reduced number of clusters (16) as in Cameron and Miller (2015). Connected: households who reported having any socioeconomic interaction or direct kin relations with council members during the survey waves preceding the release of the funds from the program. Unconnected: households without any direct connection with members of the village council.

B Microfoundations of the marginal returns to credit

Which households exhibit high marginal utilities from program credit? The identity of these households depends on the economic context in which households make their optimal decisions regarding consumption and input use. For instance, in a context of complete markets, optimal input choice should not depend on household characteristics (i.e., wealth) as households behave as unconstrained profit-maximizer firms. In that context well functioning credit markets will deliver resources to all profitable projects, and the marginal utility from a program loan should not be a function of poverty. However, in contexts of incomplete credit markets, input use will be a function of household's characteristics, and the marginal utility of a household from obtaining a loan from the program will depend on the type of frictions that characterize rural credit markets.

Consider the case of a rural household which chooses the optimal amount of inputs to be used for the family business or farm at the beginning of the year ($t = 0$) and uses the profits and other government transfers to finance consumption in the rest of the year ($t = 1$). These households may finance the only input in this economy (k_{0i}) using their initial exogenous wealth (w_i) or borrow (d_{0i}) at an interest rate of r. However, they may be liquidity constrained and only be able to borrow up to $\bar{d}$, which is exogenously determined and can be expanded by receiving loans from by the MBVF committee (b_i). Households maximize the following simplified problem:

$$\max_{c_{1i},k_{0i},d_{0i}} U(c_{1i}) \tag{B1}$$

$$\text{s.t.}$$

$$c_{1i} + (1+r)q_i d_{0i} = A_i f(k_{0i}) \tag{B2}$$

$$p_k k_{0i} \leq w_i + d_{0i} \tag{B3}$$

$$d_{0i} \leq \bar{d} + b_i \tag{B4}$$

where U denotes an increasing and concave utility function of consumption in period $t = 1$ (c_{1i}), A_i denotes household total factor productivity associated to the production function $f(k_{0i})$ which is increasing and concave in k.

For simplicity, assume $U = c_{1i}$. f is a production function that transforms the only input (k) into units of consumption goods and is increasing in k and concave ($f'' < 0$). Let $\lambda_1, \lambda_2, \lambda_3$ be the Lagrange multipliers associated to constraints (B2)-(B4), respectively. The lagrangian function associated to the optimization problem solved by household i is:

$$\mathbf{L} = c_{1i} + \lambda_1(A_i f(k_{0i}) - c_{1i} - (1+r)q_i d_{0i}) + \lambda_2(w_i + d_{0i} - p_k k_{0i}) + \lambda_3(\bar{d} + b_i - d_{0i}) \quad \text{(B5)}$$

The first order conditions imply:

$$1 = \lambda_1 \quad \text{(B6)}$$

$$\frac{1}{p_k}(A_i f'(k_{0i})) = \lambda_2 \quad \text{(B7)}$$

$$\frac{1}{p_k}(A_i f'(k_{0i}) - p_k(1+r)) = \lambda_3 \quad \text{(B8)}$$

Proposition 1. *If households face borrowing constraints, the marginal utility of relaxing this constraint is decreasing in initial wealth. Moreover, the marginal utility of relaxing a household's liquidity constraint is an increasing function of household productivity.*

Proof. The previous proposition is simply a re-statement of the argument discussed by Breza et al. (2018). In the context of binding liquidity constraints, each households only borrows up to $d^*_{0i} = \bar{d}$ and purchases inputs such that $k^*_{01} = \frac{w_i + \bar{d} + b_i}{p_k}$. Without loss of generality assume $q_i = 1$. Optimal consumption in this case is $c_{1i} = A_i f(\frac{w_i + \bar{d} + b_i}{p_k}) - (1+r)(\bar{d} + b_i)$. As a consequence of the envelop theorem, the marginal utility of an extra TBH of credit equals the marginal utility of relaxing the household's liquidity constraint ($\frac{\partial V}{\partial b} = \lambda_3$). Note that λ_3 is positive for constrained households as these households are unable equalize marginal product to marginal costs ($A_i f'(k_{0i}) > p_k(1+r)$). Moreover, because f is concave and optimal input use k^* is increasing in b_i, the marginal utility of relaxing the credit constraint will be decreasing in b_i.

Using equation (B8) to take partial derivatives of λ_3 with respect to w and A, I obtain:

$$\frac{\partial \lambda_{3i}}{\partial w_i} = \frac{A_i f''}{p_k} < 0 \tag{B9}$$

$$\frac{\partial \lambda_{3i}}{\partial A_i} = \frac{f'}{p_k} > 0 \tag{B10}$$

Equation (B9) implies that the marginal utility from borrowing from the program is decreasing in wealth. Equations (B10) and (2) imply that households with a higher utility derived from the program are high-productivity households. The intuition is simple: conditional on A, households with lower wealth can't access the resources to finance their projects. Moreover, conditional on initial wealth, the presence of borrowing constraints generates a larger misallocation of resources for high-A households (i.e., $A_i f'(k_{0i}) > p_k(1+r)$ is higher). $\square$

A similar result arises in a case in which households can borrow, but have to pay high interest rates. If high intermediation costs or information rents generate a gap between the borrowing and saving interest rates, self-financing would be cheaper option than borrowing. Poorer households would be more likely obtain more high-interest loans as self-financing may not be possible due to lack of resources. By providing loans at the lowest interest rate in the village, the program would leave poorer households better off as they can finance their operations at a cheaper cost. Moreover, conditional on budget size, a lower interest rate may boost investments generating higher marginal returns for high-productivity households.

Proposition 2. *If households do not face borrowing constraints but face high borrowing interest rates, the marginal utility from a reduction in the interest rate is a decreasing function of initial wealth and an increasing function of household productivity*

Proof. If the liquidity constraints are not binding ($\lambda_3 = 0$ in (B5)), then households choose inputs equalizing the marginal product to marginal cost and optimal input choice is a function of prices, interest rates and household productivity ($k_{0i}^{**} = k(A_i, r, p_k)$). In this environment, household debt accounts for ($d_{0i}^{**} = p_k k(A_i, r, p_k) - w_i$), and the marginal utility of decreasing interest rates is: $\frac{\delta V}{\delta r} = \lambda_1 d_{0i}^{**}$ and is positive if households are net borrowers ($d_{0i}^{**} > 0$). Taking derivatives with respect to w_i and A_i:

$$\frac{\partial \lambda_1 d_{0i}^{**}}{\partial w_i} = -1 < 0 \tag{B11}$$

$$\frac{\partial \lambda_1 d_{0i}^{**}}{\partial A_i} = p_k \frac{\partial k^{**}}{\partial A_i} > 0 \tag{B12}$$

Equation (B12) is positive as optimal input use is an increasing function of productivity as $f'' < 0$. The intuition again is simple, poorer households are more likely to be net borrowers conditional on productivity and lower interest rates are welfare increasing. Moreover, conditional on wealth, high productivity households would experience higher increases in output after adjusting their demand for inputs. □

C Econometric Appendix: Production function estimation

This section discusses the technical details regarding the estimation of the gross revenue functions. In a nutshell, estimating the gross-revenue function involves four important decisions. First, exploring the extent to which several timing restrictions related to input choice are possibly valid instruments. Second, balancing the trade-off between increased precision from adding more lags as instruments against the risk of over-fitting in the case of adding too many instruments. Third, deciding which type of estimator should be used: difference-GMM as in Arellano and Bond (1991) or system-GMM as in Blundell and Bond (1998). Finally, because I use estimates based on a 15-year long panel excluding attritors, it is important to evaluate the sensitivity of the productivity estimates to attrition.

C.1 Timing restrictions and lag selection

The reduced form specification in (6) includes variables that are not strictly exogenous (y,k,l,m), thus a fixed-effects approach would lead to dynamic panel bias (i.e., Nickell bias). An alternative way of dealing with the presence of a fixed effect is by estimating the former equation in first differences as in:

$$\begin{aligned}\Delta y_{i,t} &= \gamma_{l1}\Delta l_{i,t} + \gamma_{l2}\Delta l_{i,t-1} + \gamma_{k1}\Delta k_{i,t} + \gamma_{k2}\Delta k_{i,t-1} + \gamma_{m1}\Delta m_{i,t} + \gamma_{m2}\Delta m_{i,t-1} + \gamma_{y}\Delta y_{i,t-1} \\ &\quad + \zeta_{i,t} + \zeta_{i,t-1} + \epsilon_{i,t} - (1+\rho)\epsilon_{i,t-1} + \rho\epsilon_{i,t-2}\end{aligned}$$

Note that the structure of the error term includes up to two lags of $\epsilon_{i,t}$, which suggests that input choices made before $t-2$ are potentially valid instruments for the regressors in the first-differences equation.

The first set of potential instruments includes $k_{i,t-2}$ and $k_{i,t-3}$, as k is measured at the beginning of each period, and the third and fourth lags of m, l and y, which are predetermined with respect to $\epsilon_{i,t-2}$. Columns (1) and (3) from Panel A in Table C1 present reduced-form estimates estimated using Arellano and Bond (1991)'s "Difference-GMM" approach for households in the farm and the non-farm sectors. The Hansen test of over-identifying restrictions is strongly rejected in both cases and suggests that the timing restrictions might not be valid: p-val<0.02 in the case of farm-sector households and p-val<0.05 in the case of households from the off-farm sector. This could mean that while capital is measured at the beginning of each period, households may invest based on the expectations of business conditions in period $t + 1$.

A less restrictive model that allows households to invest based on one-year ahead expectations is presented in Columns (2) and (4). In this case, the set of instruments includes lags 3 and 4 for all inputs and the lagged dependent variable. Note that in this case, the Hansen-statistic decreases and also the number of instruments used. As a consequence, the over-identifying restrictions are only rejected at a 10% level in the case of the farm-sector households and are not rejected in the case of the off-farm households (p-val>0.1). While it is likely that the validity of the instruments increases with the lag length, this could also lead to the problem of weak instruments. Thus, I choose to use lags 3 to 4 as a starting point for subsequent model selection.

Table C1: Estimates of factor elasticities using different specifications

Panel A: Reduced-form estimates

	Farm		Non-Farm	
	(1)	**(2)**	(3)	(4)
y_{t-1}	0.21***	**0.17**	0.38***	**0.23****
	(0.071)	**(0.114)**	(0.075)	**(0.108)**
k_t	0.24	**0.26**	-0.30	**-0.18**
	(0.211)	**(0.280)**	(0.281)	**(0.329)**
k_{t-1}	-0.31	**-0.24**	0.28	**0.14**
	(0.205)	**(0.181)**	(0.179)	**(0.230)**
m_t	0.44***	**0.37*****	0.19***	**0.18*****
	(0.061)	**(0.074)**	(0.067)	**(0.060)**
m_{t-1}	-0.19***	**-0.14****	-0.20***	**-0.21*****
	(0.062)	**(0.063)**	(0.047)	**(0.043)**
l_t	-0.00	**0.08**	0.19	**0.34*****
	(0.084)	**(0.070)**	(0.159)	**(0.131)**
l_{t-1}	0.02	**-0.05**	-0.12	**-0.10**
	(0.073)	**(0.070)**	(0.154)	**(0.152)**
IV lags y & k	t-2,t-3	**t-3,t-4**	t-2,t-3	**t-3,t-4**
IV lags l & m	t-3,t-4	**t-3,t-4**	t-3,t-4	**t-3,t-4**
Observations	3,283	**3,283**	2,279	**2,279**
AR(1) p	5.77e-08	**4.83e-05**	1.89e-09	**4.96e-05**
AR(2) p	0.373	**0.614**	0.0581	**0.431**
Hansen stat	114.2	**103.5**	108.7	**98.47**
DF	86	**82**	86	**82**
P-val(Hansen)	0.023	**0.054**	0.049	**0.104**
Panel B: Common factor estimates - OMD				
ρ	0.14***	0.24**	0.32***	0.37***
	(0.04)	(0.08)	(0.06)	(0.07)
β_k	0.08	0.09	-0.25	-0.30
	(0.12)	(0.19)	(0.18)	(0.23)
β_m	0.37***	0.32***	0.22***	0.31***
	(0.04)	(0.05)	(0.06)	(0.04)
β_l	0.02	0.04	0.11	0.15
	(0.05)	(0.05)	(0.10)	(0.08)
Returns to scale	0.47	0.45	0.09	0.16
Chi2-stat (RTS=1)	14.64	7.32	17.66	15.22
P-val(RTS=1)	0.00	0.01	0.00	0.00
J-stat OID-OMD	10.66	4.77	13.37	17.12
P-val (OID-OMD)	0.01	0.19	0.00	0.00
Panel C: Baseline log TFP estimates				
TFP mean	8.73	3.86	8.70	4.48
TFP SD	2.61	1.59	2.30	0.96

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: Panel A presents estimates of the reduced-form specification using different set of lags as gmm instruments. Columns(1) and (2) present estimates in the case of Farm-sector households. Columns (3) and (4) present estimates in the case of households from the off-farm sector. Each column presents coefficients estimated through "difference-GMM" (Arellano and Bond, 1991). Estimation is conducted based on a two-step approach and the resulting variance co-variance matrix has been corrected to account for potential small sample bias. Panel B reports structural estimates after imposing common factor restrictions to the reduced-form estimates. The estimates were obtained through Optimal Minimum Distance using the variance-co-variance matrix from the reduced-form coefficients as a weighting matrix. Standard errors are clustered at the household level to account for serial correlation of unknown form.

C.2 Number of instruments

One important trade-off in the estimation of dynamic panel models concerns the number of instruments. Including further lags as instruments involves using more information to more precisely estimate the parameters. However, adding too many of them would lead either to weak instruments or over-fitting. Because I allow for differences in elasticities by sectors, estimations are conducted over samples of roughly 300 household and 15 periods, in such a setting using all the available lags as instruments would lead to the case of instruments exceeding the number of households. Thus, a careful examination of the potential efficiency gains from adding further lags as instruments is important.

Table C2 reports reduced-form estimates and structural estimates on panels A and B. Columns (1) and (5) report the baseline estimates for farm and non-farm sectors, respectively. These estimates include the 3rd and 4th lags in levels of the endogenous variables. Columns (2) and (6) include lags 3 to 5 in the set of instruments. Relative to the baseline model, precision increases for each of the reduced-form and structural coefficients. Note that simply adding the fifth lag of each variable to the set of instruments expanded the total number of instruments by 40. Columns (3) and (7) include the sixth lag of each variable into the set of instruments and columns (4) and (8) include all the available lags as instruments. Note that relative to the model with lags 3 to 5, the point estimates are very similar. Moreover, there doesn't seem to be a gain in precision, but the number of instruments approaches the number of households in the sub-sample when using lags 3 to 6 as instruments, and exceeds the number of households when using all the available instruments. As there is no extra gain in precision, I choose the models in columns (2) and (4) as the main models for the empirical analysis in this paper.

Table C2: Estimates of factor elasticities using different set of lags as instruments

Panel A: Reduced form estimates

	Farm				Non-farm			
	(1)	**(2)**	(3)	(4)	(5)	**(6)**	(7)	(8)
Lag length	3 to 4	**3 to 5**	3 to 6	3 onward	3 to 4	**3 to 5**	3 to 6	3 onward
y_{t-1}	0.17	**0.18***	0.23**	0.35***	0.23**	**0.40*****	0.50***	0.54***
	(0.114)	**(0.096)**	(0.091)	(0.088)	(0.108)	**(0.080)**	(0.070)	(0.055)
k_t	0.26	**0.28**	0.27	0.24	-0.18	**0.26**	0.11	0.10
	(0.280)	**(0.226)**	(0.204)	(0.194)	(0.329)	**(0.263)**	(0.213)	(0.180)
k_{t-1}	-0.24	**-0.19**	-0.13	-0.15	0.14	**-0.10**	0.01	0.04
	(0.181)	**(0.164)**	(0.169)	(0.139)	(0.230)	**(0.176)**	(0.175)	(0.128)
m_t	0.37***	**0.35*****	0.35***	0.36***	0.18***	**0.20*****	0.19***	0.20***
	(0.074)	**(0.062)**	(0.057)	(0.038)	(0.060)	**(0.061)**	(0.054)	(0.048)
m_{t-1}	-0.14**	**-0.18*****	-0.18***	-0.20***	-0.21***	**-0.25*****	-0.25***	-0.22***
	(0.063)	**(0.056)**	(0.048)	(0.038)	(0.043)	**(0.039)**	(0.041)	(0.040)
l_t	0.08	**0.16****	0.13**	0.13***	0.34***	**0.31*****	0.23**	0.17*
	(0.070)	**(0.065)**	(0.056)	(0.039)	(0.131)	**(0.101)**	(0.100)	(0.091)
l_{t-1}	-0.05	**-0.02**	0.02	-0.03	-0.10	**-0.21**	-0.21	-0.21**
	(0.070)	**(0.060)**	(0.053)	(0.041)	(0.152)	**(0.131)**	(0.128)	(0.102)
Observations	3,283	**3,283**	3,283	3,283	2,279	**2,279**	2,279	2,279
AR(1) p	0.000	**0.000**	0.000	0.000	0.000	**0.000**	0.000	0.000
AR(2) p	0.614	**0.499**	0.330	0.072	0.431	**0.210**	0.098	0.062
Hansen stat	103.5	**144.8**	175.1	248.8	98.47	**142.1**	177.3	207.1
DF	82	**122**	158	302	82	**122**	158	302
P-val(Hansen)	0.0542	**0.0776**	0.167	0.989	0.104	**0.103**	0.139	1

Panel B: Common factor estimates - OMD

	(1)	**(2)**	(3)	(4)	(5)	**(6)**	(7)	(8)
ρ	0.24**	**0.25*****	0.27***	0.42***	0.37***	**0.57*****	0.59***	0.53***
	(0.08)	**(0.06)**	(0.05)	(0.04)	(0.07)	**(0.05)**	(0.05)	(0.04)
β_k	0.09	**0.20**	0.12	0.15	-0.30	**0.25**	0.13	-0.04
	(0.19)	**(0.16)**	(0.14)	(0.13)	(0.23)	**(0.19)**	(0.16)	(0.12)
β_m	0.32***	**0.30*****	0.33***	0.34***	0.31***	**0.37*****	0.28***	0.23***
	(0.05)	**(0.04)**	(0.04)	(0.03)	(0.04)	**(0.03)**	(0.04)	(0.03)
β_l	0.04	**0.15*****	0.14***	0.12***	0.15	**0.26*****	0.22**	0.17**
	(0.05)	**(0.04)**	(0.04)	(0.03)	(0.08)	**(0.07)**	(0.07)	(0.06)
Returns to scale	0.45	**0.66**	0.59	0.61	0.16	**0.88**	0.63	0.35
Chi2-stat (RTS=1)	7.32	**4.32**	8.09	7.87	15.22	**0.32**	3.91	19.11
P-val(RTS=1)	0.01	**0.04**	0.00	0.01	0.00	**0.57**	0.05	0.00
J-stat OID-OMD	4.77	**6.30**	6.07	5.80	17.12	**20.56**	17.83	16.36
P-val (OID-OMD)	0.19	**0.10**	0.11	0.12	0.00	**0.00**	0.00	0.00

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: The table presents estimates of the reduced-form equation and the structural model based on difference-GMM varying the length of lags included as instruments. Columns (1) to (4) report results for households in the farm sector. Columns (5) to (8) report results for the households in the off-farm sector. Coefficients are estimated using a two-step approach and the resulting variance co-variance matrix has been corrected to account for potential small sample bias. Panel B reports structural estimates after imposing common factor restrictions to the reduced-form estimates. The estimates were obtained through Optimal Minimum Distance using the variance-co-variance matrix from the reduced-form coefficients as a weighting matrix. Standard errors are clustered at the household level to account for serial correlation of unknown form.

C.3 Robustness to including further moment conditions: Difference vs. System GMM

An important concern in the context of the estimation of production functions is related to the source of variation that is used to estimate the output elasticity with respect to capital. By differentiating out the fixed effects from the reduced-form equation, the "difference-GMM" approach (Arellano and Bond, 1991) only exploits within subject variation to identify the factor elasticities. While in principle, this is enough to obtain consistent estimates, this approach may lead to downward biases of the elasticity of capital if most of the variation in capital is likely to be explained by cross-sectional variation rather than within subject variation. Such scenario is likely in the case of Thai households: Samphantharak and Townsend (2010) document that investments are rather lumpy in the case of the households in this sample. They find that only 11% of household-month observations in the initial 88 waves of the survey exhibit positive investments.

One limitation of the difference-gmm approach is that it does not make use of the full set of moment conditions to identify the reduced form parameters. Blundell and Bond (1998) propose an alternative estimator for the reduced-form equation that uses both within subject and cross-subject variation. It uses lagged variables in levels as instruments for first-differences, and suitable first differences as instruments for lagged levels. From a practical perspective, using these extra moment conditions increases precision and also exploits cross section variation to identify the parameters. However, exploiting additional moment conditions related to the equation in levels implies imposing additional identification assumptions that could be questionable in practice. Thus it is important to examine the extent the sensibility of the estimates to both approaches.

Table C3 presents coefficients estimated through difference-GMM and system GMM for comparison. Note that the coefficient related to capital is fairly stable suggesting that the Arellano Bond difference-GMM approach is likely to capture relevant variation in capital. However, the estimates are more precisely estimated in the case of the system-GMM estimates as they include more information to identify each parameter. In terms of the structural parameters, the main difference relies in the persistence of TFP: the system-GMM estimates yield higher persistent parameters ρ than the difference-GMM estimates. The former are more likely to satisfy the common factor restrictions as the J-Stat corresponding to the test of the validity of the parameter restrictions is not rejected (p-val>0.64 in both cases).

As the object of interest are the factor elasticities which are fairly stable across methods, I use the estimates from the system-GMM approach as they are estimated with higher precision. Because this choice comes at the cost of stronger identification assumptions than the ones required by the difference-GMM approach, mainly that first differences are not correlated with the initial levels of output and inputs, I report robustness results for the difference-GMM estimates for all the results in the paper.[41]

[41] While system-gmm includes cross-sectional variation and exploits a richer set of instruments to provide more precise estimates, it imposes assumptions regarding the relation of output and factor trends with the initial conditions. In particular, as the equation in levels uses first-differences as instruments for the endogenous variables, identification requires that factor and output growth are orthogonal to the initial levels of output. In words, businesses which start with a larger size should not systematically exhibit higher growth rates.

Table C3: Sensitivity of difference-gmm estimates to including moment restrictions associated to the equation in levels (System GMM)

Panel A: Reduced-form estimates

	Farm		**Non-farm**	
	Diff-GMM	**System-GMM**	**Diff-GMM**	**System-GMM**
	(1)	(2)	(3)	(4)
y_{t-1}	0.18*	0.67***	0.40***	0.75***
	(0.096)	(0.097)	(0.080)	(0.048)
k_t	0.28	0.32*	0.26	0.17
	(0.226)	(0.177)	(0.263)	(0.175)
k_{t-1}	-0.19	-0.28	-0.10	-0.08
	(0.164)	(0.175)	(0.176)	(0.171)
m_t	0.35***	0.39***	0.20***	0.33***
	(0.062)	(0.068)	(0.061)	(0.055)
m_{t-1}	-0.18***	-0.27***	-0.25***	-0.23***
	(0.056)	(0.049)	(0.039)	(0.045)
l_t	0.16**	0.12*	0.31***	0.31***
	(0.065)	(0.071)	(0.101)	(0.118)
l_{t-1}	-0.02	-0.06	-0.21	-0.30***
	(0.060)	(0.067)	(0.131)	(0.112)
Observations	3,283	3,584	2,279	2,586
AR(1) p	0.00	0.00	0.00	0.00
AR(2) p	0.499	0.00661	0.210	0.0277
Hansen stat	144.8	185.2	142.1	196.3
DF	122	170	122	170
P-val(Hansen)	0.0776	0.201	0.103	0.0813
Panel B: Common factor estimates - OMD				
ρ	0.25***	0.66***	0.57***	0.72***
	(0.06)	(0.06)	(0.05)	(0.03)
β_k	0.20	0.15**	0.25	0.26**
	(0.16)	(0.06)	(0.19)	(0.09)
β_m	0.30***	0.40***	0.37***	0.33***
	(0.04)	(0.04)	(0.03)	(0.04)
β_l	0.15***	0.14**	0.26***	0.28***
	(0.04)	(0.04)	(0.07)	(0.08)
Returns to scale	0.66	0.69	0.88	0.87
Chi2-stat (RTS=1)	4.32	17.26	0.32	1.31
P-val(RTS=1)	0.04	0.00	0.57	0.25
J-stat OID-OMD	6.30	1.69	20.56	1.70
P-val (OID-OMD)	0.10	0.64	0.00	0.64
Panel C: Correlates of TFPs difference and system GMM)				
OLS coefficient	1.010***		1.005***	
SE	(0.017)		(0.003)	

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: Panel A presents reduced-form coefficients estimated with two alternative methods: difference-GMM (Arellano and Bond, 1991) and system-GMM (Blundell and Bond, 1998). Both sets of coefficients are estimated using a two-step approach and the resulting variance co-variance matrix has been corrected to account for potential small sample bias. Panel B reports structural estimates after imposing common factor restrictions to the reduced-form estimates. The estimates were obtained through Optimal Minimum Distance using the variance-co-variance matrix from the reduced-form coefficients as a weighting matrix. Standard errors are clustered at the household level to account for serial correlation of unknown form. Panel C reports OLS coefficients from a regression of TFP estimated with the Diff-GMM approach on TFP estimated with the system-GMM approach.

C.3.1 Robustness to attrition

I use a balanced panel of 509 households that report information in each of the 172 household survey waves. These households, represent 72% of the total sample. Table C4 reports a comparison between the estimates from the preferred specification (System-GMM with lags 3 to 5 as instruments) using a fully balanced panel and an unbalanced panel including all the observations available in the survey.

It shows that the only coefficient that seems sensitive (in magnitudes) is the one corresponding to capital. By incorporating all the observations, the elasticity corresponding to capital is smaller and it leads to a rejection of the common factor restrictions in the case of off-farm households. In order to examine how much would the productivity estimates differ with respect to those based on the elasticities estimated from the balanced sample, Panel C reports OLS coefficients of a regression of household TFP from the benchmark specification on TFP measures using the unbalanced panel and village fixed effects. The results show that both measures are strongly correlated; a one-percent increase in one measure is related to a one-percent increase in the other, suggesting that results might be invariant to either specification.

Table C4: Sensitivity of estimates to attrition

Panel A: Reduced form estimates

	Farm		Non- Farm	
	(1)	(2)	(3)	(4)
Panel	Balanced	Unbalanced	Balanced	Unbalanced
y_{t-1}	0.67***	0.58***	0.75***	0.77***
	(0.097)	(0.079)	(0.048)	(0.046)
k_t	0.32*	0.26	0.17	0.02
	(0.177)	(0.179)	(0.175)	(0.163)
k_{t-1}	-0.28	-0.24	-0.08	0.06
	(0.175)	(0.184)	(0.171)	(0.147)
m_t	0.39***	0.38***	0.33***	0.30***
	(0.068)	(0.042)	(0.055)	(0.056)
m_{t-1}	-0.27***	-0.26***	-0.23***	-0.18***
	(0.049)	(0.046)	(0.045)	(0.039)
l_t	0.12*	0.13**	0.31***	0.22**
	(0.071)	(0.065)	(0.118)	(0.104)
l_{t-1}	-0.06	-0.05	-0.30***	-0.30***
	(0.067)	(0.057)	(0.112)	(0.098)
Observations	3,584	4,253	2,586	3,307
AR(1) p	3.48e-10	0	0	0
AR(2) p	0.00661	0.00586	0.0277	0.00400
Hansen stat	185.2	204	196.3	203.6
DF	170	170	170	170
P-val(Hansen)	0.201	0.0383	0.0813	0.0402
Panel B: Common factor estimates - OMD				
ρ	0.66***	0.58***	0.72***	0.76***
	(0.06)	(0.05)	(0.03)	(0.04)
β_k	0.15**	0.06	0.26**	0.08
	(0.06)	(0.05)	(0.09)	(0.12)
β_m	0.40***	0.37***	0.33***	0.28***
	(0.04)	(0.03)	(0.04)	(0.04)
β_l	0.14**	0.14***	0.28***	0.25**
	(0.04)	(0.04)	(0.08)	(0.08)
Returns to scale	0.69	0.58	0.87	0.61
Chi2-stat (RTS=1)	17.26	49.04	1.31	10.51
P-val(RTS=1)	0.00	0.00	0.25	0.00
J-stat OID-OMD	1.69	3.30	1.70	8.33
P-val (OID-OMD)	0.64	0.35	0.64	0.04
Panel C: Correlates of TFPs (balanced - unbalanced)				
OLS coefficient	0.982***		0.949***	
SE	(0.008)		(0.012)	

$***p < 0.01, **p < 0.05, *p < 0.1$

Note: The table presents estimates of the reduced-form equation and the structural model estimated using system-GMM exploiting a balanced panel and an unbalanced sample. Columns (1) to (2) report results for households in the farm sector. Columns (3) to (4) report results for the households in the off-farm sector. Both sets of estimates are computed using a two-step approach and the resulting variance co-variance matrix has been corrected to account for potential small sample bias. Panel B reports structural estimates after imposing common factor restrictions to the reduced-form estimates. The estimates were obtained through Optimal Minimum Distance using the variance-co-variance matrix from the reduced-form coefficients as a weighting matrix. Panel C reports OLS coefficients from a regression of TFP recovered using the balanced-sample model on village fixed effect and TFP recovered using the unbalanced-panel model. Standard errors are clustered at the household level to account for serial correlation of unknown form.

D Variable definition

D.1 Data used to compute TFP

- Gross revenues: Total revenues from cultivation, livestock sales and production of livestock produce, fishing and shrimping, wage labor provision and off-farm family business. The data is obtained from household income statements. It includes the value of the production sold to other households for cash or in-kind payments, used as inputs for other activities,[42] and the value of self-consumption (valued at sales prices). From an accounting perspective, revenues are registered when the sale is made rather than when the household obtains the resources.I aggregate the revenues by summing all the revenues obtained during an economic year: that is from April to March of the following calendar year.

- Stock of fixed capital: The value of land, livestock, as well as tools, machinery and other fixed assets used for cultivation, fishing or shrimping (including a pond) and off-farm businesses. It excludes liquid assets, thus it does not include working capital. The information comes from the Balance Sheet statement of each household measured as of April of each year, the first month of the Thai economic year. Each asset is valued at its acquisition cost. A fixed depreciation rate of 10% annual is linearly applied to each fixed asset other than land. The depreciation rate of animals is computed based on their age and life expectancy.

- Use of non-labor flexible inputs: Market value of input usage. Usage is registered as an operation cost in the Income Statement and is registered once the final product is sold or consumed by the household. In the case of cultivation it includes the value of fertilizer, seeds and pesticides as well as the costs of irrigation. In the case of livestock it includes the cost of feed and operations but excludes depreciation or capital loss. In the case of retail businesses it includes the purchase price of each item. Finally, in the case of wage labor provision it includes transportation costs. Note that this measure only includes input usage which is not necessarily the same as input purchase

[42] It is registered as if the household makes a sale in the market and re-purchases it in the same period.

as households my store inputs. Thus, for the large majority of items it is possible to think of costs as reductions of inventories associated to operations.

- Labor: total number of work hours across all household activities over an economic year. For each household activity, I counted the number of work hours corresponding to: *i)* wage workers, *ii)* workers under a non-wage agreement (includes labor sharing), and *iii)* household members. To account for the use of labor coming from labor provision by the household to other households or businesses I count the number of hours per month in which each household member worked outside the household and then aggregate across all household members.