IDB WORKING PAPER SERIES Nº IDB-WP-989

Targeting Credit through Community Members

Diego A. Vera-Cossio

Inter-American Development Bank
Department of Research and Chief Economist

December 2019

Targeting Credit through Community Members

Diego A. Vera-Cossio

Inter-American Development Bank

December 2019

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Vera-Cossio, Diego A.
Targeting credit through community members / Diego A. Vera-Cossio.
p. cm. — (IDB Working Paper Series ; 989)
Includes bibliographic references.
1. Credit-Thailand-Econometric models. 2. Microfinance-Thailand-Econometric models. 3. Village communities-Economic aspects-Thailand. I. Inter-American Development Bank. Department of Research and Chief Economist. II. Title. III. Series.
IDB-WP-989

http://www.iadb.org

Abstract*

Limited borrower information may create targeting distortions in credit markets. Community-based lending programs may reduce these distortions by exploiting information transmitted in local networks, but connections may create asymmetries in power. This paper analyzes how local leaders balance issues of neediness, productivity (TFP), risk, and favoritism to allocate subsidized loans to Thai villagers. Local leaders provided credit to richer, less-productive and elite-connected villagers. These connection-based distortions threatened the program's sustainability. Moreover, eliminating these distortions would increase village-level output by 1.5%. Finally, informal markets partially attenuated the targeting distortions by redirecting credit to unconnected households, albeit at high interest rates.

Keywords: Decentralization, Entrepreneurship, Targeting, Microfinance

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

---

*Research Department, Inter-American Development Bank.

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, Juan Pablo Chauvin, and attendees at the UCSD Applied Economics seminar series, as well as PACDEV 2018, ABCDE 2018, seminars at ITAM and the World Bank for their insightful comments. This paper reflects the author's own opinions and does not reflect the IDB's official positions. 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 (Casey, 2018; Mansuri and Rao, 2004). Their popularity is based on the premise that local leaders have accurate information about local needs. In the context of financial inclusion, delegating the allocation of loans to community members could lead to the delivery of capital to those who would benefit the most: poor but high-productivity households. However, the success of these schemes depends on whether community members can effectively keep local leaders accountable (Reinikka and Svensson, 2004; Björkman and Svensson, 2010). Although community-based approaches have been shown to be effective in implementing projects with objective and verifiable outcomes—e.g., providing antipoverty aid to the poor (Galasso and Ravallion, 2005), or building local infrastructure (Casey et al., 2012)— less is known about their effectiveness when the implementation of projects entails balancing multiple subjective dimensions, such as the case of allocating credit, which may limit the scope for community monitoring.

One important class of lending programs is that of government infusions of resources into villages for the establishment of community-managed credit funds.[1] Given that verifying borrower attributes can be costly (Townsend, 1979), and that the returns to credit are heterogeneous (Meager, 2019; Banerjee et al., 2019, 2018), the role of local socioeconomic networks is crucial in these schemes: local leaders may use them to obtain information about risk, neediness, and returns (Iyer et al., 2016; Alderman, 2002; Hussam et al., 2017). In contrast, as these schemes are prone to favoritism (Bardhan and Mookherjee, 2005), the lack of connections may hamper the ability of community members to monitor and challenge the allocation of credit. This paper empirically analyzes how community members solve these tensions in order to allocate credit to poor but high-productivity households, the role of local economic networks in reducing or creating allocative distortions, and the ability of informal credit markets to attenuate targeting errors.

This paper empirically assesses these issues in the context of one of the largest lending programs in developing countries, the Million Baht Village Fund (MBVF). Between 2001 and 2002, the Thai government donated resources to over 90% of rural villages for the creation of village credit funds that, on average, expanded the village supply of credit by 25%. The funds were fully managed by elected Village Fund committees (VFCs) made up of villagers, who decided to whom credit

[1] Four of the largest lending programs in developing countries decentralize the allocation and management of publicly provided loans to community members: the Million Baht Village Fund Program in Thailand (Kaboski and Townsend, 2012), the Village Banking Program in China (Cai et al., 2017), the Integrated Rural Development Program in India (Bardhan and Mookherjee, 2006), and the Rural Financial Institutions Programme, also in India.

would be extended and under what conditions. The program's stated objective was to deliver individual-liability loans to promote income generation.[2] Thus, Village Fund committees faced the problem of balancing issues of repayment, neediness, and productivity. Importantly, the rollout of the program overlaps with the Townsend Thai Project Monthly Survey (Townsend, 2014), which collects three years of pre-program information regarding household enterprises, loans, and cross-household transactions, as well as several post-program waves.

The analysis in this paper follows three steps. First, I exploit 14 years of panel data to structurally estimate gross-revenue production functions following the approach of Blundell and Bond (2000). I then use the estimated factor elasticities to recover *pre-program* estimates of total factor productivity (TFP) associated with all household businesses.[3] I combine these estimates with baseline repayment and per-capita consumption data to test whether baseline repayment, poverty, and TFP predict program borrowing. Second, I combine detailed data on baseline within-village transactions and social interactions to document the relationship between socioeconomic connections and program selection, the mechanisms behind such a relation, and the efficiency gains from eliminating connection-based distortions. Third, I use quasi-experimental variation in the rollout of the program to test for within-village reallocation of credit through informal credit markets.

This paper first provides a descriptive characterization of the allocation of program credit. A simple model of constrained entrepreneurs suggests that, without allocative distortions, committee members would allocate more credit to poorer and high-productivity households. However, the data do not support that prediction. While estimated pre-program TFP is correlated with household education and risk-adjusted returns over assets (Samphantharak and Townsend, 2018), it does not predict higher chances of obtaining credit, and predicts lower total program borrowing. In contrast, richer and less-vulnerable households—with higher baseline per-capita consumption and lower consumption volatility—obtained more program credit. Moreover, program credit was delivered to households who already had access to formal credit, even to those with preexisting history of delinquent payments.

Although program credit was not allocated based on productivity, poverty, or repayment, I find evidence that resources were disproportionately allocated to households with connections to

[2]See (Government of Thailand, 2004) for a further description of program objectives, and Kaboski and Townsend (2012) for an analysis of the program's effects.

[3]All sample households have at least one enterprise, which makes the estimation of TFP household by household possible. Identification of the production function relies on timing restrictions that provide a set of suitable instruments for input usage. Capital is measured as the stock of total fixed assets registered in household balance sheets compiled by Samphantharak and Townsend (2010).

members of the Village Council—i.e., the village government. Council members and households with direct connections to council members were 30 and 16 percentage points more likely to obtain program resources than unconnected households. The results are robust to controlling for health and production shocks that could trigger demand for credit and suggest that there were important connection-based allocative distortions.

Second, I analyze two nonexclusive mechanisms through which connections could relate to program participation. One mechanism is that elite-connected households were better located in the village socioeconomic networks, which lowered the costs of transmitting information to program committee members. After controlling for the total number of links in the village network, the correlation between program participation and elite connections falls sharply. This result suggests that VFCs were indeed able to extract information through local networks. However, even after controlling for network centrality, per-capita consumption, consumption volatility, TFP, repayment history, and exposure to health and production shocks, village council members are still 20 percentage points more likely to obtain program credit, which raises the suspicion of favoritism.

There are two causes that would explain why elite-connected households obtain more program credit: better enforcement and favoritism. However, better enforcement should be profitable for the program while favoritism should be costly for the program. Following Shaban (1987) and Khwaja and Mian (2005), I empirically test for favoritism by following a double difference approach. For the same borrower, I compute differences in ex post returns to the lender between program loans and loans from private community-based organizations.[4] I then compare these differences in returns between connected and unconnected borrowers.

I find that the *ex post* internal rate of return (IRR) on program loans to connected households was 2.7 percentage points lower than that on loans from member-funded community lending groups (on average 7%), relative to similar comparisons in the case of unconnected households. These losses suggest favoritism. Moreover, there were no differences in repayment, but elite-connected households benefited from larger and cheaper loans for a similar level of risk, even exceeding borrowing caps imposed by the central government. These patterns could be the consequence of incomplete contracts (Holmstrom and Milgrom, 1991; Hart et al., 1997): despite government incentives to induce high repayment, size and pricing decisions were left to the committee's discretion.

An alternative interpretation is that the financial losses of lending to elite-connected borrowers

[4] I use a sample of 6,700 loans made to 335 households, who borrowed both from the program and other member-funded village credit groups. These groups 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.

represent the price of ensuring the long-term stability of the Village Funds or the cost of fostering local institutions. However, 10 years after the program was initially rolled out, Village Funds in villages that originally allocated a larger share of credit to the local elite grew less—some even decreased—than those in villages with more-egalitarian allocations. Despite these distortions, VFCs could have still outperformed other policy-relevant targeting criteria. However, under the assumption that nonborrowers indeed demanded credit, a repayment-score targeting criterion would have targeted less-risky and more-productive households, without favoring elite-connected households.[5] One explanation is that scoring models, while imperfect, provide objective targeting rules that are less vulnerable to elite influence.

All in all, the results suggest that asymmetries in connections to the local elite can generate targeting distortions. Thus, reducing these distortions may lead to efficiency gains. I find that simply redistributing the excess program credit obtained by elite-connected borrowers (11% of Village Fund portfolio) to nonborrowers would increase village-level output by 1.5%. Behind this increase, there are substantial returns to credit for non-borrowers: for each THB of credit reallocated to nonborrowers, village-level output would increase by THB 3. These gains are evenly explained by information-transmission frictions and favoritism, and they suggest that the connection-based distortions are not only related to program profitability and sustainability but also translate to village-level costs in output.

Finally, while the program might not have directly reached unconnected households, the program's rollout indirectly delivered loans to unconnected households through informal credit markets. Exploiting cross-village variation in the monthly rollout of the program, difference-in-differences estimates reveal that borrowing from informal lenders increased by 30% in the case of unconnected households. These loans were mostly obtained from relatives at an average annual interest rate of 14%, which is twice as high as that of program loans. This result suggests that lower program borrowing in the case of unconnected households was not mainly driven by lack of demand, as they did borrow, albeit at higher rates. It also suggests that targeting frictions may have generated arbitrage opportunities in local credit markets. Indeed, connected borrowers increased the probability of lending due to the program rollout. However, informal markets only mildly attenuated targeting distortions: back-of-the-envelope calculations suggest that these effects only account for 10% of the

[5] Exploiting pre-program data from 3,800 loans, and household financial and demographic characteristics, I use the least absolute shrinkage and selection operator (LASSO) to estimate a repayment probability model, and identify the set of clients who, based on predicted repayment, would have been eligible for a loan, while holding program coverage constant.

program-borrowing gap between connected and unconnected households.

This paper contributes to the literature on community-based approaches to targeting productive resources in two ways. The first contribution is methodological. Bardhan and Mookherjee (2006); Menkhoff and Rungruxsirivorn (2011) and Coleman (2006) analyze the ability of local leaders to deliver loans to the poorest households, while Basurto et al. (2017) analyze whether village chiefs deliver fertilizer subsidies to high-return farmers. The former two studies lack pre-program information regarding productivity, while the latter relies on postprogram, self-reported measures of average returns. This paper contributes to these studies by using three years of pre-program measures of entrepreneur productivity (TFP) that were unaffected by the program itself. This key feature allows this paper to provide novel evidence on the efficiency losses due to connection-based frictions in terms of program sustainability but also in terms of village-level output. Second, although Coase (1960) predicts that in the absence of transaction costs, secondary private arrangements could overcome allocative distortions, evidence on the role of secondary markets in the context of community-based programs is scarce (Giné et al., 2019). This paper complements previous studies by showing that informal credit markets can (partially) attenuate connection-based distortions.

This paper also builds on the literature analyzing the means through which local elites obtain rents form public resources(Anderson et al., 2015; Goldstein and Udry, 2008; Acemoglu et al., 2014) by documenting that connections to the local elite can generate distortions in the allocation of subsidized loans. This finding is consistent with evidence of favoritism based on socioeconomic links to elites in financial markets in different contexts (Haselmann et al., 2017; Schoenherr, 2018; Khwaja and Mian, 2005; Agarwal et al., 2016). However, the results contrast with evidence on the modest incidence of elite capture in cash-transfer programs (Alatas et al., 2019, 2012). One explanation is that, unlike the case of targeting antipoverty resources to the needy, targeting credit entails balancing dimensions that are costly to verify (Townsend, 1979), which may limit the ability of unconnected households to push for more pro-poor, less-risky or more-efficient allocations (Reinikka and Svensson, 2004).

Lastly, while there were costly allocative distortions based on connections, the MBVF program did increase consumption (Kaboski and Townsend, 2012), and it mildly reduced capital-market failures, relative to a no-program scenario (Shenoy, 2017b). However, the program only increased the profits of high-TFP entrepreneurs who were able to borrow (Banerjee et al., 2018), and a simple cash-transfer program would have been more cost-effective (Kaboski and Townsend, 2011). This

paper provides an explanation for such modest results: connection-based allocative distortions prevented the delivery of more credit to the most productive. More broadly, it suggests that targeting frictions may explain the nontransformative effects of other microcredit programs (Banerjee et al., 2015).

2 Context

This paper studies the context of the Million Baht Village Fund (MBVF) program in Thailand. Starting in 2001, the Thai government donated THB 1 million (USD 22,500 at 1999 values) to each participating village.[6] The funds were used as seed capital for the creation of village credit funds that would provide loans to community members. Any villager was allowed to apply for a loan,[7] and borrowers were expected to repay with interest. Once a loan was repaid, both the principal and revenues from interest were reinvested in the Village Funds and were allocated to other local borrowers. 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 of village credit funds, and by 2004 its gross lending portfolio exceeded USD 3 billion (Haughton et al., 2014).

The program represented a large unexpected increase in the supply of credit at the village level. It was announced following a change in government in January 2001 and was initially rolled out between June 2001 and April 2002. 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.

A unique feature of the program is its community-based management. Each village elected a Village Fund committee (VFC) made up of nine to 12 community members. Committee members were elected in community meetings for a two-year term and received a small amount of compensation for their service. Most of them continued in the position for several years, however (Haughton et al., 2014). The VFC was responsible for allocating loans and monitoring repayment, yet no specific training was provided.[8] Committee members met once or twice a year to review loan applications and authorize disbursements into borrowers' bank accounts in the state-owned Bank for

[6] Around 95% of all Thai villages participated in the program, including all the 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).

[7] In order to apply, households were required to purchase a share of the fund at a very low cost that was mainly symbolic.

[8] 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 (five years of schooling).

Agriculture and Agricultural Cooperatives (BAAC). While VFC decisions were subject to a set of restrictions regarding loan size and term,[9] VFC members had full discretion to approve or deny applications and to set interest rates.

The program offered individual liability loans, which did not require a collateral, but did require 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 preexisting sources of credit in the villages, program loans exhibited the lowest interest rate in the market: the second-lowest interest rate was that of bank loans (11% per annum, see Appendix Table AIX). In terms of maturity, the program offered loan terms that were similar to those of quasi-formal lenders:[11] shorter than those of banks, but longer than those of 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 at the village level. If the repayment rate was high, a village would be rewarded with further infusions of resources. In contrast, if the default rate was high, government transfers and funding for other programs to the village would be suspended. However, no direct incentives for or sanctions of VFC members were built into the program.

2.1 The Program and the Local Political Elite

The program was implemented in villages with well-established local political elites such as the village council (village head and advisers). Village council members are generally elected by villagers, they report to district authorities, and usually serve in office until retirement.[12] The village council represents the main link between community members and higher-level authorities: 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 five years of schooling), they are richer and

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

[10] Average loan size is THB 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] This was the case during the study period. However, a reform in 2011 set terms at five years but allowed village heads to run for reelection.

hold almost twice as much land as their fellow villagers.

Even though the VFC was de jure an independent entity, it is possible that the village council members had enough de facto authority to influence VFC decisions. For instance, when elections could not take place, VFC members were appointed by the village head.[13] Moreover, the local elite 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, while 13% of households in the sample were first-degree relatives of village council members.

3 Theoretical Framework

The program's stated objectives were to expand access to institutional credit, and promote career development and income generation (Government of Thailand, 2004), which suggests that poverty, productivity, and repayment are central to the understanding of program participation.[14] However, there were no clear program guidelines regarding how these criteria would be balanced.

In this section, I first sketch a simple framework that characterizes the optimal allocation of loans by the VFC, assuming that VFC members balance villagers' utilities as well as program revenues.

The VFC's problem: committee members decide the amount of credit (b_i) that each of their N_v fellow villagers obtains from the program at a given interest rate r. They do so in order to maximize a weighted sum of the utilities of all the community members as well as the revenues related to program credit, subject to a resource-availability constraint determined by the village endowment of MBVF funds B_v:

$$\max_{\{b_1,\ldots,b_{N_v}\}} \sum_{i=1}^{i=N_v} \psi_i \left[(1+r)b_i + V(b_i)\right] \tag{1}$$

s.t.

$$\sum_{i=1}^{i=N_v} b_i \leq B_v \tag{2}$$

[13] Haughton et al. (2014) document that 15% of village fund committee members were appointed directly by either the village head or the village council

[14] For instance, access to institutional credit was low among poorer households, the government claimed publicly that resources were allocated to productive activities (Pasuk and Baker, 2004), and program sustainability relies on repayment.

Here, V_i denotes household i's utility, which is increasing and concave in b_i. For the sake of simplicity, I assume that Village Fund loans are always repaid, but that achieving repayment is rather costly for VFCs and hence they will value the returns on each loan differently across households. This simplifying assumption is consistent with evidence of high repayment rates in the sample villages (See Appendix Table AIX) and with the idea that peer monitoring may be effective at enforcing repayment (Bryan et al., 2015). Thus, each VFC weights the returns on each loan and the associated household utility by a weight ψ_i such that $\sum_i \psi_i = 1$ for each village.

Political favoritism, social norms, and preferences or non-pecuniary costs may determine the weights associated with each village member (ψ_i), which are exogenous with respect to the allocation problem. Concretely, I assume that ψ_i is an increasing function of household wealth w_i, as wealthier households may have more means to challenge or influence VFCs' decisions. Likewise, ψ_i is an increasing function of whether a household is connected to the local elite (d_i), as better-connected entrepreneurs may use their connections to obtain more-favorable allocations, or because connections may lead to lower screening and/or enforcement costs for committee members. Finally, ψ_i is a decreasing function of the ex ante probability of repayment q_i, as committee members have to exert more effort in order to achieve the repayment of risky loans. Thus, $\psi_i = \psi(w_i, d_i, q_i)$.

Finally, the sustainability constraint (2) dictates that the total value of loans does not exceed the Village Fund endowment B_v. The first-order conditions imply:

$$\psi_i \left[\frac{\partial V_i}{\partial b_i} + (1+r)\right] = \psi_j \left[\frac{\partial V_j}{\partial b_j} + (1+r)\right] \tag{3}$$

Expressed in words, VFC members allocate resources such that the weighted marginal utilities of their fellow villagers—including the revenues from program loans—are equalized. Equation (3) highlights the complexity of the VFC's problem as it involves returns, risk, and the potential influence of richer or well-connected households. It can be shown that, in the case of entrepreneurs choosing capital in order to maximize profits, subject to borrowing limits given by the amount of program credit b_i, the optimal allocation of program credit can be written as:[15]

[15] Consider the very simple case of an entrepreneur choosing capital k to maximize profits, and hence utility from consumption, subject to a credit constraint:

$$\max_k U = c_i = A_i k_i^\alpha - rk - b_i$$

$$s.t. k_i = w_i + b_i$$

If the budget constraint is binding, then $k_i^* = w_i + b_i$, the value function is $V(b_i) = A_i[w_i + b_i]^\alpha - (1+r)[b_i] - rw_i$,

$$b_i^* = \frac{(\psi_i A_i)^{\frac{1}{1-\alpha}}}{\sum_i^{N_v} (\psi_i A_i)^{\frac{1}{1-\alpha}}} \left[B_v + \sum_i^{N_v} w_i \right] - w_i \tag{4}$$

Here, A_i denotes the entrepreneurs' total factor productivity (TFP), and $\alpha \in [0, 1]$ denotes the capital elasticity corresponding to a production function $y = A_i k_i^{\alpha}$.

To understand the importance of the trade-offs faced by committee members, consider first the case in which household weights are constant across villagers or, equivalently, a context in which neither richer nor elite-connected households can influence the allocation of resources and there are no differences in ex ante repayment probabilities. In such a setting, program credit is allocated according to $b_i^* = \frac{(A_i)^{\frac{1}{1-\alpha}}}{\sum_i^{N_v} (A_i)^{\frac{1}{1-\alpha}}} \left[B_v + \sum_i^{N_v} w_i \right] - w_i$. Thus, committee members allocate more resources to the relatively more productive households in the village. Moreover, the village fund committees would also consider neediness and allocate more resources to poorer households. Thus, in such a simplified context, committee members would achieve an efficient and progressive allocation.

In practice, $\psi_i = \psi(w_i, d_i, q_i)$ captures the allocative distortions preventing marginal returns from equalizing across borrowers as in (Hsieh and Klenow, 2009). For instance, richer households could exert more pressure on committee members, leading to a regressive allocation. Committee members may favor elite-connected households over the most productive or prioritize less-risky households over more-productive but possible riskier entrepreneurs. Empirically, finding that resources are systematically delivered to more-productive and needy households would imply that differences in household weights generate smaller distortions. In contrast, observing program credit flowing towards richer or lower-TFP households would indicate that either ex ante risk or the ability of households to influence committee decisions drives the allocation. Which of these forces dominates is an empirical question.

4 Data and Measurement

The context of the MBVF program coincides with the availability of high-frequency, detailed data from the Townsend Thai Project Monthly Survey (Townsend, 2014). Starting in September 1998, the survey followed 710 households on a monthly basis for over 14 years and includes three years of

and the marginal utility of an extra unit of program credit is: $\frac{\partial V}{\partial b} = \alpha A_i [w_i + b_i]^{\alpha - 1} - (1 + r)$. Plugging $\frac{\partial V}{\partial b}$ into equation (3) and solving for b_i yields the expression on equation (4).

pre-program information, which is essential to provide a full characterization of potential borrowers at baseline. While the survey covers only 16 villages,[16] 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.

The dataset provides high-frequency information regarding transactions with other households in the village, the portfolio of loans held by each household, purchases, sales, and use of inputs as well as the destination of final output. Additionally, it is possible to link the survey to households' financial statements (cash flows, income, and balance sheets) that were constructed by Samphantharak and Townsend (2010). Table I reports summary statistics. Over 80% of households own land, and over one-third of household revenues correspond to agricultural production. However, the average household obtains revenues from four different economic activities, such as raising and selling livestock, fishing and shrimping, or wage labor provision. Off-farm business ownership is not rare, either: 15% of sampled households obtained income from these businesses. Such a context and the richness of the data allow the study of households as corporate firms.

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. The fact that cash represented over 30% of household assets 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%), although it should be noted that this low delinquency coincides with high shares of loan term extensions (36%).

4.1 Measuring Baseline Neediness, Productivity, Repayment, and Connections

Using three years of pre-program panel data, I characterize the set of potential borrowers according to four important dimensions: neediness, productivity, repayment, and connections with local leaders.

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

Table I: Summary Statistics: Household Characteristics

Variable	N	Mean	S.D.	Percentile 10th	Percentile 90th
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	4.09	1.77	2	6
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 (THB)	688	2366	8229	-67	5992
Income volatility (coefficient of variation)	699	2.19	2.17	0.60	3.65
Per-capita consumption (THB)	688	1563	1420	638	3027
Consumption volatility (coefficient of variation)	694	0.86	1.06	0.32	1.79
Income-consumption co-movements	710	0.11	0.24	-0.12	0.42
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
Elite connected (Any link with Village Council members)	710	0.65	0.48	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's implementation. The latter information is only available for a subset of households who reported borrowing from any source during the baseline periods. Per-capita measures are adjusted by household composition and age of household members using household equivalent scales. Income and consumption volatility are computed as the coefficient of variation of monthly income or consumption during the pre-period.

Neediness. The analysis focuses on two dimensions of neediness: levels and variance. To assess pre-program neediness I compute measures of average monthly per-adult equivalent consumption during the three years preceding the program's rollout.[17] Although on average some households

[17]To adjust for family composition I follow Deaton (1997) and compute the number of adult equivalents (AE) as

may exhibit higher levels of consumption, different type of shocks may affect households who are less able to smooth out such shocks. To proxy the ability of households to smooth consumption, I use the pre-program monthly data to compute the coefficient of variation of household consumption, household by household.

I complement the aforementioned two measures with proxies of transitory neediness by using the pre-period to compute exposure to shocks that could have triggered the 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. In the case of health symptoms, the Townsend Thai Project Monthly Survey asks each household member to report all symptoms from a provided list that they experienced during the previous survey wave.[18] In the case of shocks to business operations, the survey records self-reported information on the number of times a family business experienced problems with production, output sales, input purchases, and the reception of payments.

Repayment. To measure repayment history, I exploit self-reported information regarding loans households had taken out from banks, credit groups, and informal personal lenders. In each survey wave, enumerators recorded information regarding loan characteristics for all new loans and followed up with information regarding repayment throughout the life of each loan. I use this information to identify borrowers who either failed to make loan payments or extended the term of loan.[19] One potential limitation of self-reported measures of repayment is that, since default might be socially undesirable, interviewees may underreport delinquent payments. However, as this paper uses mainly pre-period data, the potential bias should be similar among those households who later on borrow from the program and those who do not.

Connections to the Local Elite. First, I use detailed baseline information regarding several economic interactions to elicit baseline undirected, unvalued socioeconomic village networks,[20] which I complement with data about kinship relations. Second, I combine the network data with pre-program information regarding participation in the village councils to quantify elite connectedness. A household is defined as connected with the local elite if any of its members are part of the

$AE = N_{Adults} + 0.3 * N_{Under15}$.

[18] Such data have been shown to be predictive of reductions in business revenues (Kinnan et al., 2019).

[19] I use delinquency rates, rather than default rates, since default is uncommon in the sample: recovery rates are on average over 97% (see Table AIX).

[20] The transactions can be categorized into 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 and advising. Following Banerjee et al. (2013), I consider all possible transactions, as different interactions may transmit relevant information. See Online Appendix Section B.5 for a detailed explanation of the construction of the variables.

village council, are first-degree kin of a council member, or have engaged in at least one transaction, of any type, with a council member during the baseline period.[21]

4.2 Productivity

I focus on estimates of baseline revenue total factor productivity (TFP) for household enterprises as a proxy for marginal returns to credit. Estimating TFP at the household level is possible as all sample households have either a farm or an off-farm business. By capturing variation in output unexplained by input use, TFP captures the ability of a household to generate revenue, while holding input usage constant. Intuitively, higher TFP may increase demand for inputs, which may lead to more binding credit constraints for businesses without access to credit.

Consider a log-production function in which output $y_{i,t}$, corresponding to household i during period t, is a function of labor $l_{i,t}$, nonlabor variable inputs $m_{i,t}$—i.e., intermediates, fixed capital $k_{i,t}$, household productivity $a_{i,t}$, and shocks to production $\epsilon_{i,t}$. Household productivity is known to the household but is unobserved by the researcher. It captures managerial ability, household-specific business opportunities, and economic conditions that may affect the choice of inputs. Unforeseen shocks to production ($\epsilon_{i,t}$) are neither observed by the researcher nor considered by the household in relation to input choice (e.g., production loss due to theft, spoilage, or unexpected natural disasters).

$$y_{i,t} = \beta_l l_{i,t} + \beta_k k_{i,t} + \beta_m m_{i,t} + a_{i,t} + \epsilon_{i,t} \tag{5}$$

In order to estimate (5), I construct an annual panel by aggregating the balances of monthly income statements and labor and time use data over each Thai economic year.[22] This approach prevents seasonality from driving the results and captures household behavior over the full production cycle. I then use these data to measure annual revenues and total hours of labor (hired and household labor). Next, I combine this dataset with balance sheets from household operations measured at the beginning of each economic year (April), which I then use to measure capital.

I focus on total output and input usage across all household economic activities which include agriculture, livestock farming and production of animal produce, fishing and shrimping, off-farm

[21]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).

[22]Thai economic years begin in April and end in March. The beginning of the year coincides with the beginning of the rainy season, and an economic year captures two rice-production cycles. All monetary variables are deflated with respect to 1999 prices.

family businesses, and wage work outside the household. I do so as households may simultaneously optimize resources across all economic activities: all sampled households have at least two sources of revenue and, on average, derive income from four different sources (see Table I).[23] However, I also report estimates that exclude revenue, cost and time use related to the provision of labor outside the household for robustness. As factor elasticities may vary across economic activities, I estimate (5) separately for households which are mainly involved in the farm sector (agriculture, livestock, fishing and shrimping) and households who mostly obtain revenues from the off-farm sector (businesses or wage labor).[24]

I proxy total output with gross revenue from all household activities in a given year. I measure nonlabor input usage as the cost of all the variable inputs used for production, which include fertilizer, seeds, feed, merchandise, fuel, transportation, and tools required for nonfarm family businesses. Consistent with time-to-build models, 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, nonfarm business assets, machinery, and other household assets. Finally, labor is measured as the 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.

4.2.1 Identification of the Gross-Revenue Function

While exogenous variation in capital, labor, and inputs is not available in the Thai context, it is possible to exploit the panel structure of the data to attenuate endogeneity. Intuitively, when credit constraints or other market frictions limit the ability of households to respond to unforeseen productivity shocks, past choices of inputs may be informative about future input choice and may be orthogonal to unforeseen productivity shocks (Shenoy, 2017a).[25]

Following Blundell and Bond (2000), I assume that household productivity $a_{i,t}$ is a function of a time-invariant component (α_i) and a time-variant component following a first-order autoregressive process ($\omega_{i,t} = \rho\omega_{i,t-1} + \zeta_{i,t}$), such that $a_{i,t} = \alpha_i + \omega_{i,t} = \alpha_i + \rho\omega_{i,t-1} + \zeta_{i,t}$. While this assumption

[23] As production functions are product specific, aggregating across sources of revenues comes at the cost of interpretation of the factor shares. This may not be a first-order concern, however, as the factor shares themselves are not the focus of this paper.

[24] I coded a household as being part of the farm sector if the baseline share of farm revenues was higher than 0.5 and coded a household as being part of the off-farm sector if the share of farm revenues was below 0.5.

[25] As Shenoy (2017a) suggests, the possibility of credit constraints may invalidate identification assumptions typically invoked by traditional structural approaches such as Olley and Pakes (1996), Levinsohn and Petrin (2003), or Ackerberg et al. (2015).

imposes linearity, the implied structure is quite flexible, as it allows production and input choice to respond to both time-invariant abilities and transitory but persistent business opportunities. Appendix Section B.2 shows that, under that assumption, equation (5) can be written as:

$$\begin{aligned}\Delta y_{i,t} &= \gamma_{l1}\Delta l_{i,t} + \gamma_{l2}\Delta l_{i,t-1} + \gamma_{k1}\Delta k_{i,t} + \gamma_{k2}\Delta k_{i,t-1} \\ &\quad + \gamma_{m1}\Delta m_{i,t} + \gamma_{m2}\Delta m_{i,t-1} + \gamma_y \Delta y_{i,t-1} + \tilde{\epsilon}_{i,t} \end{aligned} \tag{6}$$

where $\Delta x_{i,t} = x_{i,t} - x_{i,t-1}$, $\gamma_{j1} = \beta_j$, $\gamma_{j2} = -\rho\beta_j$ (for $j = \{l, k, m\}$), $\gamma_y = -\rho$, and $\tilde{\epsilon}_{i,t} = \zeta_{i,t} - \zeta_{i,t-1} + \epsilon_{i,t} - (1+\rho)\epsilon_{i,t-1} - \epsilon_{i,t-2}$. This transformation is important for identification in two ways. First, it provides a set of restrictions that maps the reduced-form parameters to the "structural" parameters in (5). Second, by eliminating α_i and $\omega_{i,t-1}$, identification of equation (6) only requires assumptions regarding the relation between input choice and the current unforeseen shocks to productivity and production, as opposed to previous productivity shocks. In particular, the parameters from (6) are identified by the following moment condition:

$$\mathbf{E}[\zeta_{i,t} - \zeta_{i,t-1} + \epsilon_{i,t} - (1+\rho)\epsilon_{i,t-1} - \epsilon_{i,t-2} | I_{i,t-3}] = 0 \tag{7}$$

Here, $I_{i,t-3}$ denotes the information set available to household i at the end of period $t-3$. Equation (7) suggests that lagged versions of output and inputs at $t-3$ and backward are valid instruments. In this case, the behavioral assumption is that while households may choose inputs based on their productivity forecasts ($\rho\omega_{i,t-1}$), time-invariant characteristics α_i, and contemporary and past unexpected shocks (ζ, ϵ), households do not choose current inputs in anticipation of future, unexpected shocks. While this assumption is rather mild, there could be other unobserved confounding variables such as output and input prices. To minimize that risk, estimations include village and year fixed effects, as well as interactions of baseline share of agricultural revenues with rainfall and external rice prices to control for variation in prices across time, space, and sectors. Turning to relevance, it is important that lagged versions of output and inputs are predictive of changes in output and inputs. Appendix Table AX shows that lagged levels do indeed have strong predictive power.

4.2.2 Estimation

Estimating baseline productivity involves three steps. First, I estimate equation (6) using suitable instruments through the generalized method of moments (GMM). This process yields seven 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 implement this process separately for two sectors: *i*) households whose revenues from farm activities (cultivation, livestock farming and fishing) account for more than 50% of their baseline revenues, and for *ii*) households whose revenues mostly come from off-farm activities. Third, I combine the estimated factor elasticities with three years of pre-program data regarding input usage and revenues to back out estimates of baseline household productivity for all potential borrowers: $\hat{a}_{i,t} = y_{i,t} - \hat{\beta}_l l_{i,t} - \hat{\beta}_k k_{i,t} - \hat{\beta_m} m_{i,t}$. Finally, I average TFP over the pre-program years and use these estimates to study selection into program credit.

Without further assumptions, estimating the reduced-form parameters from equation (6) requires at least four periods, but there are only three available pre-program periods. Instead, I use 14 years of panel data including pre- and postprogram years. At the cost of assuming that factor elasticities are time invariant, this approach provides enough variation to implement dynamic panel estimation techniques. Concretely, I estimate the reduced-form model using Blundell and Bond (1998)'s system-GMM estimator.[26] The system-GMM approach incorporates both within- and cross-household variation to recover reduced-form parameters as opposed to Arellano and Bond (1991)'s difference-GMM approach that only exploits within-household variation. This approach is appealing in contexts in which capital is generally fixed over time; however, it requires assuming that first differences in output and inputs are orthogonal to initial levels of output.[27]

Appendix Table AXI reports the reduced-form coefficients as well as factor elasticities by sector using the difference- and system-GMM approaches. Panel A shows that, regardless of the method, the reduced-form specifications are likely to pass the Hansen test for overidentifying restrictions highlighting the validity of the instruments. Panel B presents OMD estimates of $\hat{\beta}_k, \hat{\beta}_m, \hat{\beta}_l$, and $\hat{\rho}$. Columns (1) and (2) report estimates obtained by estimating the reduced-form equation through Arellano and Bond (1991)'s difference-GMM estimator, and columns (3) and (4) report estimates based on the system-GMM approach. While the factor elasticities are similar across methods, they

[26] In particular, I use lags 3 to 5, which balances issues of precision with the risk of overfitting due to too many instruments. An econometric discussion of the specification choice is detailed in Online Appendix Section B.1.

[27] This assumption seems appropriate in contexts in which firms operate around their steady state.

are more precisely estimated in the case of the system-GMM estimator. Importantly, the structural restrictions imposed to the reduced-form estimates are likely to hold in the case of the system-GMM estimates. Thus, the system-GMM estimates are the preferred specification in this paper.

4.2.3 Alternative Specifications

Throughout this paper, I report results that are robust to three different approaches to estimating productivity:

Robustness to only using pre-program data. Estimating equation (6) is quite data-demanding. However, it is possible to estimate a restricted version of equation (6) using only the three periods preceding the program rollout. Columns (1) and (2) from Appendix Table AXII report GMM estimates of factor elasticities assuming away the presence of a time-invariant component of TFP, but allowing TFP shocks to be persistent over time (Panel A)—i.e., $a_{i,t} = \omega_{i,t}$ as opposed to $a_{i,t} = \alpha_i + \omega_{i,t}$.[28] Panel B shows that the TFP estimates using only pre-program data are strongly correlated with the benchmark estimates.

Robustness to measurement error in fixed capital. Fixed capital is likely to be measured with error (Kim et al., 2016; Collard-Wexler and De Loecker, 2016). Failure to account for this problem may lead to underestimating the elasticity of fixed capital and overestimating productivity for capital-intensive entrepreneurs. To tackle this issue, instead of using lagged capital in levels to instrument for current changes in capital, I follow Collard-Wexler and De Loecker (2016) and instrument current changes in capital with suitable lags of investment—i.e., the cash flows associated with capital expenses. Columns (3) and (4) in Appendix Table AXII show that correcting for measurement error yields larger capital elasticities in both sectors. However, these estimates are imprecise, as investment is lumpy in the Thai context.[29] Reassuringly, despite the differences in the estimated elasticities, panel C shows that baseline TFP measures are strongly correlated with the benchmark estimates.

Robustness to excluding operations related to off-household labor. The empirical approach in this paper has not imposed restrictions on how households allocate inputs and time across their sources of income. In fact, this paper considers households as complex productive units involved in the production of goods, retail activities, and the provision of inputs and services to other businesses. However, there is the concern that including wage-labor activities may introduce

[28] Estimation details are presented in Appendix section B.1.

[29] For example, Samphantharak and Townsend (2010) find that only 11% out of 55,000 household-month observations during the first 84 months of the Townsend Thai Project Monthly Survey recorded positive investments.

bias into the TFP estimates. Columns (5) and (6) from panel A in Appendix Table AXII report factor elasticities that were estimated excluding wage-labor related data. Concretely, I used measures of revenues and costs that excluded labor earnings and the costs of labor provision (mainly transportation). I also excluded the number of hours worked outside the households from the measures of labor. Reassuringly, factor elasticities are quite similar to the benchmark estimates.

4.2.4 Validating the TFP Estimates

Correlates with education, rainfall, and financial returns. To assess whether the baseline TFP estimates capture meaningful variation, as opposed to simply capturing noise, Table II shows correlates of baseline TFP with demographic characteristics and potential productivity shifters, controlling for village fixed effects and baseline wealth. As expected, in the case of farm-oriented households, rainfall interacted with the share of agricultural revenues is strongly and positively correlated with TFP. Likewise, experiencing shocks to agricultural businesses predicts declines in productivity in farm-oriented households. Moreover, TFP is negatively correlated with household age, which is consistent with a sector with physically intensive tasks. In contrast, for nonfarm households, TFP is correlated with the probability that the household head has completed primary school, which is consistent with the idea that better-educated households may have comparative advantages in off-farm businesses. Finally, the TFP estimates are correlated with measures of risk-adjusted return over assets (RoA) computed by Samphantharak and Townsend (2018). Thus, increases in TFP are correlated with idiosyncratic changes in average financial returns.[30] These results are robust across methods.

Persistence. The idea of this paper is to analyze whether local committees directed resources to higher-productivity entrepreneurs. Using pre-program data has the advantage of attenuating issues of reverse causality, but this advantage comes at the cost of assuming that past variation in TFP is a good predictor of TFP associated with the periods in which committees make decisions. In other words, the estimates of TFP are useful only to the extent that they are persistent. Appendix Figure AV shows that the TFP estimates are quite persistent. More formally, panel B of Table AXI shows that the TFP persistence parameters (ρ) are significant and substantial (between 0.6 to 0.75 in the case of our preferred measure). This result suggests that the baseline TFP estimates are informative about household TFP around the rollout of the program.

[30] More precisely, I use data from 520 households for which Samphantharak and Townsend (2018) recovered a risk-adjusted measure of the returns over assets (RoA), net of village level fluctuations.

Table II: Baseline Correlates of TFP Estimates and Entrepreneur's Characteristics

	System-GMM		Measurement Error		Excluding Labor Earnings		Only pre-program data	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Farm	Non Farm	Farm	Non Farm	Farm	Non Farm	Farm	Non Farm
Age of household's head	-0.003	0.005	-0.004	0.006	0.002	-0.002	-0.001	-0.005
	(0.003)	(0.005)	(0.004)	(0.006)	(0.004)	(0.010)	(0.003)	(0.004)
Household's head completed primary school	0.103	0.270**	0.121	0.300*	0.148	0.197	0.122	0.173
	(0.092)	(0.131)	(0.108)	(0.157)	(0.114)	(0.271)	(0.095)	(0.121)
Head of household gender (male)	0.094	0.078	0.150	0.100	0.048	0.065	-0.003	-0.103
	(0.092)	(0.094)	(0.102)	(0.115)	(0.099)	(0.206)	(0.092)	(0.100)
Number of adults	0.045	0.064	0.044	0.036	0.091***	-0.066	-0.030	-0.080*
	(0.030)	(0.052)	(0.035)	(0.063)	(0.034)	(0.085)	(0.028)	(0.042)
Number of elder	0.000	-0.027	0.013	-0.049	-0.036	-0.299*	-0.068	-0.102
	(0.048)	(0.081)	(0.053)	(0.095)	(0.055)	(0.176)	(0.052)	(0.087)
Number children under 5	0.021	0.004	0.020	0.001	-0.043	0.265	0.027	0.015
	(0.052)	(0.074)	(0.052)	(0.088)	(0.062)	(0.214)	(0.063)	(0.086)
Share of females in the household	-0.164	-0.123	-0.211	-0.076	-0.014	0.493	-0.073	0.095
	(0.174)	(0.288)	(0.210)	(0.359)	(0.166)	(0.449)	(0.146)	(0.243)
Average age in household	-0.007	-0.004	-0.010**	-0.005	-0.007*	0.000	-0.002	0.009
	(0.004)	(0.005)	(0.005)	(0.006)	(0.004)	(0.010)	(0.004)	(0.005)
Average education level in household	0.013	0.012	0.013	-0.013	-0.002	0.015	0.001	0.022
	(0.020)	(0.030)	(0.021)	(0.038)	(0.024)	(0.058)	(0.022)	(0.028)
Count of health symptoms	0.005**	-0.000	0.005*	0.002	0.003	0.001	0.004	0.004
	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)	(0.011)	(0.003)	(0.003)
Count of shocks to non farm business	-0.012	-0.013	-0.008	-0.036*	-0.029**	0.008	-0.027*	-0.019
	(0.010)	(0.015)	(0.011)	(0.018)	(0.014)	(0.025)	(0.014)	(0.017)
Count of shocks to livestock business	0.005	0.023	0.015	0.022	-0.015	0.016	-0.023*	-0.054***
	(0.014)	(0.029)	(0.016)	(0.033)	(0.014)	(0.035)	(0.013)	(0.020)
Count of shocks to agriculture	-0.037*	-0.011	-0.027	-0.010	-0.046*	-0.046	-0.058**	-0.020
	(0.021)	(0.029)	(0.022)	(0.034)	(0.027)	(0.057)	(0.027)	(0.033)
Idiosyncratic Return over Assets	0.009*	0.022***	0.010*	0.024***	0.010**	0.027***	0.009	0.028***
	(0.005)	(0.005)	(0.006)	(0.007)	(0.005)	(0.009)	(0.006)	(0.005)
Share of agricultural revenues	3.808***	-1.393	2.499	-2.343	2.086	3.411	4.379***	0.138
	(1.308)	(1.723)	(1.572)	(2.145)	(2.191)	(3.188)	(1.326)	(1.414)
Share of agricultural revenues X rainfall	7.172***	-1.077	4.548	-2.584	3.737	5.333	8.242***	0.715
	(2.482)	(3.108)	(2.980)	(3.849)	(4.039)	(6.040)	(2.556)	(2.662)
Observations	292	228	292	228	292	195	292	228
R-Squared	0.448	0.543	0.493	0.571	0.420	0.258	0.389	0.416
Adjusted R-squared	0.375	0.463	0.426	0.495	0.343	0.100	0.308	0.313

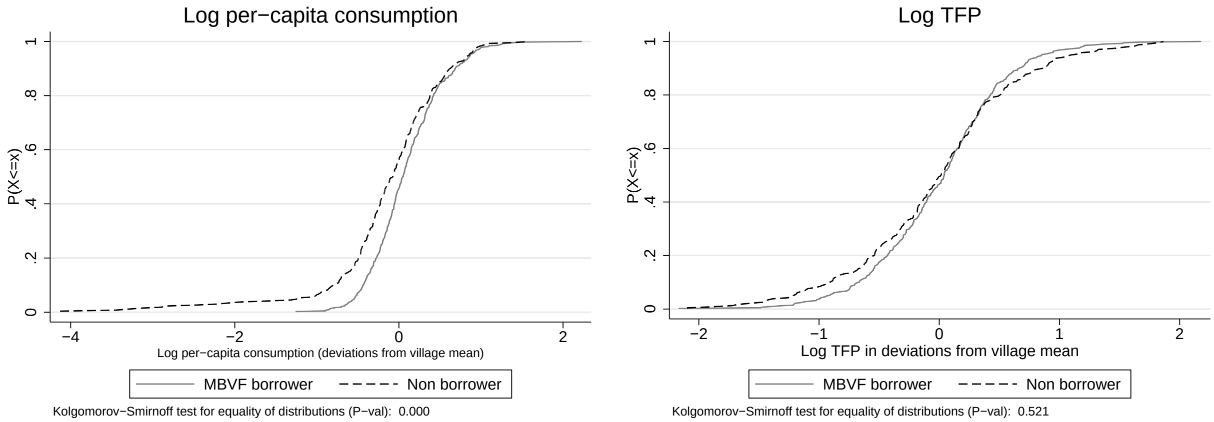
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: The table presents baseline correlates of household TFP and baseline characteristics estimated through OLS by sector and different estimation method: System GMM (Columns 1 and 2), TFP corrected for potential measurement error in capital (Columns 3 and 4), TFP excluding income and costs associated with the provision of labor to other firms (Columns 5 and 6), and TFP estimates using only pre-program data assuming no time-invariant component in the TFP process (Columns 7 and 8). 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 household level.

5 Who Obtains Program Credit?

This section provides a descriptive analysis assessing the main predictors of program participation. While the program currently operates in several villages, the analysis is based on the two years following the rollout of the program, as baseline characteristics are more likely to be informative of the context around that period.[31]

[31] I choose two years in order to capture households who may not have needed credit during the first year but may have requested it during the second year. Also, there were some modifications to the program in 2004, three years after its initial rollout. 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.



(a) Per-capita consumption (b) Measured TFP

Figure I: Cumulative Distribution Function of Baseline Log Per-Capita Consumption and TFP

Note: The figure plots the cumulative distribution function (CDF) of log per-capita consumption and TFP, measured at baseline, for households with access to credit from the program (59%) and households who did not 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 and is adjusted for family size. Both measures are standardized with respect to the village mean in order to perform within village comparisons, and they are winsorized with respect to the top and bottom 1%.

The theoretical framework suggests that, in the absence of allocative distortions, program credit should have been delivered to the neediest and most-productive households. I begin by analyzing whether preperiod per-capita consumption—adjusted by household composition—is predictive of program participation. Figure I plots the distribution of per-capita consumption for program borrowers and nonborrowers, normalized with respect to the village mean. It shows that consumption is higher for program borrowers at each point in the distribution, suggesting that selection into the program was not consistent with neediness as a targeting criterion. Column (1) in panel A from Table III shows that, controlling for village fixed effects, a 1% increase in baseline per-capita consumption predicts an increase of 16 percentage points in the probability of borrowing from the program. A similar strong negative correlation is found when analyzing total program credit (see panel B). Appendix Table AXIII shows that this pattern is also robust to using pre-program wealth—assets net of liabilities—as a proxy for household neediness. The results suggest that neither resource-constrained entrepreneurs nor poorer consumers were targeted by committee members.

Table III: Correlates of Pre-Program Characteristics and Program Participation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Correlates of probability of borrowing from MBVF and baseline characteristics											
Per-capita consumption (logs)	0.159***						0.119***	0.155***	0.122**	0.222***	0.174***
	(0.025)						(0.032)	(0.039)	(0.057)	(0.077)	(0.038)
Consumption volatility (log Coeff. of Variation)		-0.101***					-0.078***	-0.072**	-0.056	-0.074**	-0.070**
		(0.027)					(0.028)	(0.030)	(0.037)	(0.037)	(0.031)
TFP (logs)			0.030				0.006	-0.004	0.001	-0.000	-0.012
			(0.019)				(0.019)	(0.020)	(0.022)	(0.028)	(0.020)
Access to instutitional credit (dummy)				0.342***			0.225***	0.199***	0.166**	0.168*	0.190***
				(0.041)			(0.049)	(0.051)	(0.068)	(0.088)	(0.053)
Ever missed a payment (dummy)					0.149***		0.031	0.025	0.027	0.030	0.031
					(0.047)		(0.048)	(0.049)	(0.050)	(0.052)	(0.049)
Connected with Village Council						0.163***	0.095**	0.096**	0.097*	0.257**	0.093**
						(0.043)	(0.045)	(0.045)	(0.052)	(0.100)	(0.047)
Observations	692	694	648	710	710	710	646	642	524	538	588
Adjusted R-squared	0.110	0.092	0.074	0.161	0.082	0.091	0.151	0.153	0.085	0.073	0.130
Within-village R-Squared	0.040	0.019	0.002	0.097	0.011	0.021	0.087	0.090	0.040	N.A.	N.A.
Panel B: Correlates of average MBVF borrowing and baseline characteristics											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Per-capita consumption (logs)	4,439.332***						4,138.054***	4,937.005***	5,323.785***	8,867.354***	5,253.873***
	(980.661)						(1,092.034)	(1,365.637)	(1,726.663)	(2,120.040)	(1,428.160)
Consumption volatility (log Coeff. of Variation)		-621.590					-1,420.943**	-1,248.557	-931.609	-2,332.215**	-1,392.176*
		(590.965)					(710.385)	(761.920)	(924.491)	(908.304)	(832.954)
TFP (logs)			-368.908				-678.944	-961.049**	-953.424*	-1,221.247**	-1,180.481**
			(450.481)				(429.938)	(446.983)	(512.563)	(593.707)	(467.356)
Access to instutitional credit (dummy)				6,123.178***			4,612.165***	3,479.754***	2,417.497**	3,835.432**	3,432.255***
				(847.640)			(874.775)	(903.393)	(1,062.226)	(1,828.759)	(983.883)
Ever missed a payment (dummy)					2,874.155***		1,402.494	1,373.808	1,367.388	2,190.303*	1,586.122
					(1,055.912)		(1,065.238)	(1,054.195)	(1,074.317)	(1,220.678)	(1,066.377)
Connected with Village Council						2,793.364***	1,890.379**	1,982.503**	2,698.368***	4,198.767**	2,167.977**
						(911.568)	(885.425)	(862.671)	(1,002.992)	(2,091.180)	(930.916)
Observations	650	652	619	652	652	652	617	614	511	531	562
Adjusted R-squared	0.233	0.202	0.200	0.255	0.210	0.213	0.274	0.302	0.289	0.228	0.284
Within-village R-Squared	0.040	0.000	0.000	0.066	0.010	0.014	0.094	0.130	0.115		
F-Stat (1st stage- log TFP)										191.7	4220
F-Stat (1st stage- Access)										135	
F-Stat (1st stage- log Cons.)										18.03	
F-Stat (1st stage-Connected)										24.90	
Excludes HH with no credit history	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Controls (shocks + demographics)	No	No	No	No	No	No	YES	YES	YES	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. Standard errors are clustered at the household level. All regressions control for village fixed effects. Columns (8) to (11) include demographic characteristics (household head's gender, age and education, average household age and education, number of adults, children (younger than 15) and elderly in the household) and dummies indicating whether the household experienced health, or production shocks during the preperiod. Column (10) reports IV estimates instrumenting access to institutional credit, per-capita consumption, TFP and connections with the local elite, measured the year before the program was implemented with their lags–measured two years before the program. Column (11) reports IV estimates of log TFP using three alternative TFP measures as instruments (TFP accounting for potential measurement error in capital, TFP measured excluding income and costs from wage labor, TFP estimated only with three years of pre-program data but not allowing for a fixed effect component in the TFP process) Panel B reports correlates between the average gross stock of program credit (over the two years following the program rollout) and baseline characteristics following similar specifications.

One limitation of analyzing levels of consumption is that it captures the permanent component of neediness as opposed to the vulnerability to adverse shocks (Alatas et al., 2012). Column (2) in panel A in Table III shows that households with larger monthly consumption volatility—measured as the coefficient of variation of monthly per-capita consumption—are also less likely to obtain program credit,[32] suggesting that credit was not delivered to more-vulnerable households. Panel A of Appendix Table AXIII shows that program borrowers were not more exposed to health shocks—measured as the self-reported number of health symptoms—relative to non-borrowers. This result is consistent with Kinnan et al. (2019) who show that the main source of insurance against such shocks are gifts from other households rather than loans. One explanation is that since VFCs only meet a couple times a year, their ability to promptly respond to transitory shocks is limited.[33]

It is possible that VFCs did not allocate credit to the needy but did allocate credit to high-productivity households. However, while the estimates of pre-program TFP are correlated with household education and risk-adjusted returns over assets (see Appendix Table II), column (3) in panel A from Table III shows that baseline productivity is not a good predictor of program participation. Panel B of Appendix Table AXIII shows that the results are similar using TFP estimates from three different methods.[34] While, on average, baseline productivity is uncorrelated with obtaining program credit, VFCs were indeed able to screen out lower-productivity households. Figure I compares the cumulative distribution functions of TFP between program borrowers and nonborrowers—centered at the village-sector mean. It shows that lower-TFP households were screened out, but VFCs were unable to lend to the most-productive entrepreneurs. The role of productivity is less ambiguous in the case of total program borrowing. Panel B from Table III shows that higher TFP predicts lower amounts of program credit. Thus, the potential efficiency losses seem to be concentrated in the intensive rather than the extensive margin of program credit.

Since neediness and productivity could be correlated with other borrower characteristics, column (7) in Table III shows that the correlates between program participation and baseline per-capita

[32]More formally, household consumption volatility is computed as $log(\frac{\sum_i^{T_{pre}}(c_{it}-\frac{\sum_i^{T_{pre}}c_{it}}{T_{pre}})^2}{\sum_i^{T_{pre}}c_{it}})$, where T_{pre} denotes the number of pre-period months, which ranges between 34 and 42 months depending on the rollout of the program at the village level.

[33]However, panel A from Appendix Table AXIII shows that program borrowers were more exposed to negative shocks to livestock farming during baseline periods. Nevertheless, it is unlikely that these shocks account for most of the differences in program participation, as average revenues from livestock farming represent only 8% of total household revenues (see Table I).

[34]The three methods are the model accounting for potential measurement error in capital, the model excluding operations related to wage labor provision, and the restricted model using only pre-program data for estimation that assumes away the existence of time-invariant components of productivity.

consumption, consumption volatility, and TFP are robust to controlling for pre-program credit usage, the probability of missing arrears, and connectedness with the local government. Column (8) shows that the results persist after controlling for household demographic characteristics and preperiod exposure to health and production shocks.[35] Finally, column (9) shows that the results persist even within a subsample of households with preexisting credit history (including loans from informal lenders). The same patterns are present in the case of total program borrowing in panel B.

Given that the neediest and most-productive households ended up not being program borrowers, I examine the importance of possible sources of allocative distortions. Columns (7) to (9) in panel A from Table III show that households with high pre-program access to institutional credit were around 20 percentage points more likely to obtain program resources than households without pre-program access to institutional credit. Moreover, 80% of program borrowers had pre-program experience with formal or quasiformal lenders. While reaching unbanked households did not seem to be a priority for the VFCs, it is possible that they used experience with formal credit as a proxy for creditworthiness. If that was the case, then program borrowers should exhibit better credit history. Column (9) shows that among households with baseline credit history (76% of sample households), households who missed loan payments at baseline were not penalized by committee members. If anything, panel C from Appendix Table AXIII shows that program borrowers' repayment history was rather poor. Among households with credit history, program borrowers were 12 percentage points more likely to have requested term extensions at baseline.

The results suggest that neither repayment history, nor poverty, nor productive efficiency was a relevant targeting criterion. One explanation is that committee members may have weighted households differently based on their connections with the local elite.[36] I find strong evidence that being connected to the members of the village council at baseline predicts higher chances of obtaining program credit. Column (6) in panel A from Table III shows that households who are either members of the village council or have a direct link in the socioeconomic network with council members are 16 percentage points more likely to obtain credit after controlling for village fixed

[35]I control for age, education, and gender of the household head, as well as average household age and education. In addition, I control for the number of adults, children (under 15 years old) and elderly household members (65 or older).

[36]Since only village council members in the sample can be identified, as opposed to all village council members, there is a potential downward bias in measuring connections to elites. Thus, the results based on comparisons between connected and unconnected households represent the lower bounds of the true differences. However, this bias should not be strong, as village council members represent only 10% of the households and at least one committee member is observed in each village in the sample.

effects. Panel D in Appendix Table AXIII shows that this correlation is robust across alternative measures of elite connectedness.[37] Columns (7) to (9) show that even after controlling for risk, neediness and productivity, being connected to the local leaders is a key predictor of program participation. In addition, panel B shows that elite-connected households obtained around THB 2,000 extra program credit even after controlling for other borrower attributes. This difference is substantial, as it represents 20% of average program borrowing.

Overall, this descriptive exercise suggests the allocation achieved by VFCs differs substantially from the optimal allocation rule in the absence of allocative distortions. These distortions seem to be related with power; richer and well-connected households with more means to influence or challenge VFCs decisions obtained more resources.

5.1 Robustness

While using pre-program data minimizes the risk of reverse causality, pre-program measures of borrower characteristics capture meaningful variation only if they are strong predictors of borrowers' attributes around the program rollout. To verify that the results are driven by the persistent component of borrowers attributes, I use observations corresponding to the year preceding the program implementation to compute average per-capita consumption, TFP, credit usage, and connections with local leaders.[38] I then use the observations corresponding to two years preceding the program rollout to compute lagged versions of these attributes, and I use these lagged versions to instrument for borrower characteristics during the year preceding the program implementation. Column (10) shows that the results are quite robust to this specification. Moreover, these estimates suggest a stronger negative correlation between program participation and per-capita consumption, and a stronger positive correlation between program participation and elite connectedness—being connected to the local elite predicts a 20 percentage-point increase in the probability of borrowing from the program. Importantly, the bottom panel of column (10) shows large F-stats corresponding to the first stage for each attribute. Panel B shows that the same patterns replicate in the case of total program borrowing. In all cases, attributes measured two years before the program are strong predictors of the attributes measured the year before the program.

[37] By using the extensive margin of transactions to define connections, it is possible that a household is identified as connected on the basis of one isolated interaction. Since the relative salience of each interaction cannot be identified or valued, I provide robustness checks using an elite-connectedness index based on the first principal component associated with the different types of transactions.

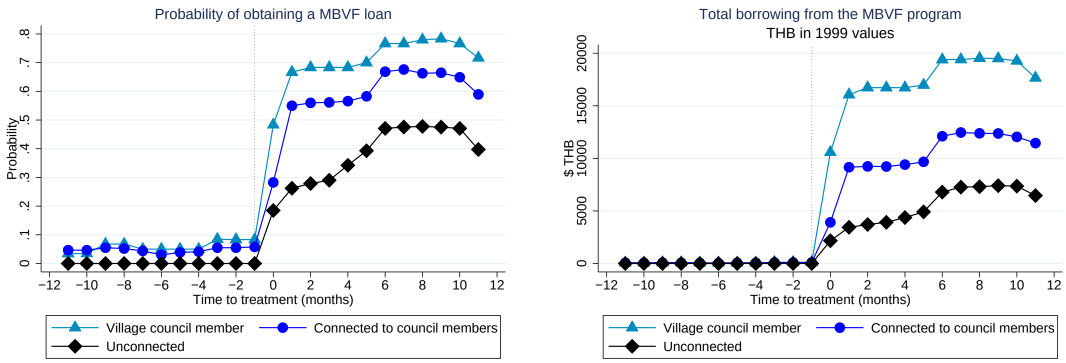
[38] Given that I use 34 months of pre-program data to recover consumption volatility and credit history, splitting the periods to compute lagged versions of these variables would lead to rather noisy measures.

Across specifications, baseline TFP does not seem to be a good predictor of the probability of program borrowing. One explanation is attenuation bias due to classical measurement error. To test the importance of this source of bias, column (11) uses three alternative estimates of baseline TFP to instrument for the preferred TFP measure. It shows no evidence of positive correlation between preperiod TFP and program borrowing. The point estimate did increase in absolute values, but it yields an imprecise negative correlation. In turn, these negative point estimates are consistent with the negative and significant correlation between pre-program TFP and total program borrowing (see panel B).

Finally, TFP is only a variable of interest as long as, before the program, all households were credit constrained, which is what is assumed in the theoretical framework in Section 3. For instance, if some high-TFP households were not credit constrained, then it would not be surprising that TFP fails to predict program participation. If that was the case, one would like to test whether program credit was allocated to those households with larger marginal revenue products of fixed capital or intermediate inputs. Appendix Figure AVI shows that marginal revenue products of fixed capital are similar in the case of program borrowers and non-borrowers, but that the marginal revenue product of intermediate inputs—i.e., feed, fertilizer, fuel, or merchandise—is lower in the case of program borrowers. On average, Panel B in Appendix Table AXIII shows that a 1% increase in baseline marginal product of intermediate inputs predicts a 5 percentage-point decline in the probability of obtaining a program loan. However, this negative correlation reduces when other borrower attributes are added as controls (see Appendix Table AXIV).

6 The Role of Connections to the Local Elite

The results from the previous section show that connections to local leaders are highly predictive of program participation. To further illustrate this result, Figure II shows how access to program credit and loan size over time vary with the type of relationship 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 obtain program credit sooner than unconnected households.



(a) Probability of borrowing

(b) Total credit

Figure II: 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.

There is also a clear gradient with respect to the type of relationship with village council members (elite members). A year after the program's rollout, elite members and households with connections to elite members were 30 and 20 percentage points more likely, respectively, to hold program credit than unconnected households. Moreover, elite members obtained double the amount of program resources obtained by nonelite households with socioeconomic connections to elite members, and almost three times the amount obtained by unconnected households. Such differences correspond to a large share of households in the sample; while village council members only represent 10% of the sample, non-elite households with direct connections to elite members represent 40% of the households in the sample. Thus, it is important to understand which mechanisms are consistent with the connection-based gap in program participation: connections may ease the transmission of information, but they may also lead to favoritism. The next subsections aim to test the salience of both mechanisms.

6.1 Do Connections to Local Elites Transmit Information?

If VFC members interpreted connections to the village council as a signal of creditworthiness or profitability, elite-connected households should be, on average, better potential borrowers. If that were true, the observed committee's allocation could be a result of statistical discrimination.

However, Appendix Table AXV shows that elite-connected households are, if anything, riskier than unconnected households. For instance, among households with baseline credit history, elite-connected households are more likely to have had delinquent payments and expanded the term of their loans at baseline. Such patterns are stronger for Village Council members, but still present for households with direct economic links to elite households.

Table IV: Program Borrowing and Connections with Village Council

	Borrowed from the program (dummy)									
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Social Connections										
Connectedness with Village Council	0.163***	0.096**	0.067	0.063	0.078*					
	(0.043)	(0.045)	(0.046)	(0.048)	(0.045)					
Village Council member						0.330***	0.190***	0.195***	0.150**	0.167**
						(0.060)	(0.068)	(0.066)	(0.072)	(0.067)
Directly transacted with council member						0.166***	0.086*	0.076	0.057	0.074
						(0.043)	(0.046)	(0.047)	(0.048)	(0.046)
First-degree relative to council member						-0.004	0.043	0.004	0.042	0.037
						(0.054)	(0.054)	(0.052)	(0.054)	(0.051)
Degree (count of links)			0.014***	0.006**	0.007***			0.013***	0.005**	0.006***
			(0.002)	(0.002)	(0.002)			(0.002)	(0.002)	(0.002)
Other baseline characteristics										
Access to institutional credit		0.200***		0.188***	0.203***		0.194***		0.183***	0.195***
		(0.051)		(0.052)	(0.045)		(0.052)		(0.052)	(0.045)
Ever missed a payment		0.026		0.022			0.018		0.016	
		(0.049)		(0.049)			(0.049)		(0.049)	
log Per capita consumption		0.153***		0.134***	0.119***		0.149***		0.131***	0.118***
		(0.039)		(0.040)	(0.025)		(0.039)		(0.039)	(0.025)
log Cons. Volatility		-0.072**		-0.069**	-0.074***		-0.074**		-0.071**	-0.077***
		(0.030)		(0.030)	(0.027)		(0.030)		(0.030)	(0.027)
log TFP		-0.003		-0.001			-0.003		-0.001	
		(0.020)		(0.020)			(0.020)		(0.020)	
Operation and health shocks										
Cultivation		0.000		0.001	0.017		0.001		0.001	
		(0.015)		(0.015)	(0.012)		(0.015)		(0.015)	
Livestock		0.016		0.015			0.017		0.016	0.018
		(0.012)		(0.012)			(0.012)		(0.012)	(0.012)
Non-agricultural business		-0.008		-0.009	-0.004**		-0.007		-0.008	
		(0.008)		(0.008)	(0.002)		(0.008)		(0.008)	
Health symptoms		-0.004*		-0.004**			-0.003*		-0.004**	-0.004**
		(0.002)		(0.002)			(0.002)		(0.002)	(0.002)
Observations	710	642	710	642	691	710	642	710	642	691
Control for demographics	NO	YES	NO	YES	YES	NO	YES	NO	YES	YES
Adjusted R-squared	0.09	0.15	0.14	0.16	0.21	0.10	0.15	0.14	0.16	0.21
Within-village adjusted R2	0.02	0.09	0.07	0.09	0.14	0.03	0.09	0.07	0.09	0.15

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

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 (5)) and on indicators of membership to the council and connectedness through transaction and kinship networks (Columns (6) to (10)). All regressions include village fixed effects. Columns (3) to (5), (8) to (10) control for the number of links in the baseline transaction network. Columns (4) and (9) control for 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 (5) and (10) only include variables selected by the LASSO. 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. Standard errors are clustered at the household level, and they are reported in parentheses.

One alternative explanation is that, while elite-connected households may not be better borrowers, their location in the network may allow them to better transmit information to VFC members.

Thus, the correlation between program participation and elite connectedness should vanish after controlling for the location of each household in the village network. One key advantage of the data in this study is that, on top of identifying links to elite members (vertical relations), it is possible to compute horizontal relations, that is, the number of fellow villagers to which each household is connected (degree centrality).

Column 1 from Table IV reports within-village correlations of program participation and being connected to the village council. Column 3 shows that adding degree centrality as a control reduces the difference in program participation based on connections to the elite from 16 to 6 percentage points, which is no longer significant. Moreover, simply including the number of links in the network improves the explanatory power of the model almost as much as adding several controls. Column 5 reports results from a predictive model using the least absolute shrinkage and selection operator (LASSO) to select relevant predictors of program borrowing.[39] Importantly, the LASSO model selected both the indicator of being connected to the local elite and network centrality. This model-selection process led to a model with larger explanatory power relative to those in Columns (1) to (4), as it explains 20% and 14% of the overall and within-village variation in program participation.[40] In this model, being connected to the local leaders is still significant at 10%.

The result suggests that connections may partially ease the transmission of information. In principle, such finding is encouraging, as community-based targeting is supposed to exploit information transmitted through 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. Finally, columns (6) to (10) from Table IV show that there are striking differences by type of connections to the elite. Even after controlling for network location and relevant borrower characteristics, elite members are still 17 percentage points more likely to obtain program credit, raising suspicions regarding favoritism.

6.2 Was There Favoritism?

As VFC members decided the size, term, and interest rate of each loan on a case-by-case basis, it is possible that elite-connected households were able to influence loan characteristics. Decisions regarding interest rates, loan size and term were decentralized to VFC members, and it is possible that connection with the local elite may have influenced those decisions.

[39] The LASSO penalty parameter is chosen through 10-fold cross-validation.

[40] The explanatory power of this model is similar to that of the predictive model in Crépon et al. (2015) used to characterize households that select into microfinance.

Thus, committee members could not only provide more-favorable loans to elite-connected households, but also use connections to achieve better enforcement. Empirically, both motives would predict the provision of larger and cheaper loans to connected households. However, unlike the enforcement motive, which is profitable for the lender, favoritism should be costly for the program. This rationale provides the theoretical foundation to 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 while controlling for unobserved borrower and lender 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. 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. I recover the ex post IRR by numerically computing the rate at which the net present value of all loan cash flows equals the principal.[41] This calculation is performed for each loan that was either fully repaid or defaulted on after some payments and was declared as not active. As the IRR is not defined for loans with no payments (0.1% of sample loans), the calculations exclude loans for which no payment was ever made. Finally, the resulting IRR is multiplied by 12 in order to obtain annual rates.

To test for favoritism while accounting for unobserved borrower characteristics, I exploit the following insight: while connections to the elite may be salient in the case of the program, elite connectedness should be less salient in the case of privately funded sources of credit. Although, on average, loans from the program may exhibit different ex post IRRs than loans from private lenders, larger IRR differences in the case of loans to elite-connected borrowers relative to loans to unconnected borrowers should be indicative of favoritism. Similar insights have been used to distinguish different monitoring models in agriculture (Shaban, 1987), and to test for favoritism towards firms with connections to the central government in Pakistan (Khwaja and Mian, 2005).

The Thai context offers an ideal setting in which 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

[41] That is, the rate (IRR) that solves the following equation: $Principal = \sum_{h=1}^{h=H} \frac{Payment_h}{(1+IRR)^h}$. Here h denotes time from disbursement and H denotes loan duration.

comparison group for program loans.[42] While comparison lenders and the Village Funds are both managed by community members, their source of funding is different: the Village Funds are fully funded by the central government, while the local credit groups are self-funded by 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.[43]

I bring this idea to the data by exploiting a subsample of 6,741 loans that 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. Following Khwaja and Mian (2005), I estimate the following specification:

$$IRR_{kijt} = \alpha_i + \theta_j + \beta \text{Connected}_i \times \text{MBVF}_j + \delta_{vt} + \epsilon_{kijt} \quad (8)$$

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.[44] 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 to the elite 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. Standard errors are clustered at

[42] 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).

[43] See Table AIX for comparative summary statistics for different sources of credit.

[44] Lenders include the 16 Village Funds as well as other village-specific credit and savings groups.

the lender level.

Table V: Differences in Loan Outcomes by Connections with Village Council Member and Lender

	Means				**Difference (MBVF-CG)**		**Double difference**			
	Connected		**Unconnected**		**Connected**	**Unconnected**	**All**			
	MBVF	**CG**	**MBVF**	**CG**	**(1)-(2)**	**(3)-(4)**		**(5)-(6)**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Returns to the lender										
Ex post IRR (annual)	0.061	0.079	0.068	0.057	-0.018	0.011	-0.027**	-0.028**	-0.021**	-0.027**
					(0.011)	(0.013)	(0.012)	(0.012)	(0.009)	(0.012)
Panel B: Loan outcomes										
Any delinquent payment	0.008	0.016	0.002	0.000	-0.004	0.003	-0.007	-0.009	-0.005	-0.007
					(0.009)	(0.002)	(0.008)	(0.009)	(0.007)	(0.009)
Delinquent payments as a share of due payments	0.006	0.010	0.001	0.000	-0.002	0.001	-0.004	-0.005	-0.003	-0.004
					(0.006)	(0.001)	(0.005)	(0.005)	(0.004)	(0.006)
Any loan extension	0.481	0.398	0.379	0.340	0.009	0.031	-0.023	-0.026	-0.064**	-0.021
					(0.052)	(0.061)	(0.034)	(0.019)	(0.024)	(0.032)
Panel C: Loan characteristics										
Initial interest rate (annual)	0.053	0.078	0.058	0.067	-0.019*	-0.007	-0.011	-0.015	-0.010	-0.011
					(0.011)	(0.009)	(0.009)	(0.009)	(0.008)	(0.009)
Term (months)	11	12	11	12	-0.189	-1.134	0.907	0.833	0.892	0.878
					(0.706)	(1.013)	(0.677)	(0.627)	(0.549)	(0.685)
Loan size (THB-1999 prices)	15,111	4,117	11,690	3,673	10,778.672***	9,070.316***	1,587.253*	1,722.826*	1,201.648	n.a
					(1,061)	(969)	(899)	(908.550)	(838)	
Loan exceeds maximum amount (> THB 20,000)	0.060	0.016	0.016	0.007	0.028*	0.007	0.022*	0.021*	0.020*	n.a
					(0.016)	(0.006)	(0.012)	(0.012)	(0.012)	
Borrower fixed effect					YES	YES	YES	YES	YES	YES
Lender fixed effect					NO	NO	YES	YES	YES	YES
Village -year fixed effects					YES	YES	YES	YES	YES	YES
Demographic controls					NO	NO	NO	YES	NO	NO
Weights for number of loans					NO	NO	NO	NO	YES	NO
Weights for loan size					NO	NO	NO	NO	NO	YES
Number of borrowers					260	75	335	323	335	335
Observations (loans)					5,274	1,404	6,741	6,050	6,741	6,741

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

Note: The sample corresponds to loans obtained after the rollout of the program that were fully repaid, reached maturity or were declared as defaulted on. It includes only loans belonging to households who 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 (10) report double difference estimates following several specifications. Column (8) includes demographic characteristics, measured the month before each loan was taken out, as controls. These characteristics include age and education of the household head, average age and years of schooling of all household members, and the number of adults, children and elderly people in the household. Standard errors are clustered at the lender. Lenders include the 16 village funds in the sample as well as each of the local community-based lenders in each village (production credit groups, woman's groups, and other similar lenders.)

Column (7) of Table V reports estimates of β corresponding to the specification in equation (8). 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 robust to controlling for time-variant demographic characteristics (see column (8)),[45] and it is neither driven by loans from borrowers who rarely borrow nor by smaller loans: Column (9) shows that the results are robust to weighting observations by the number of loans corresponding to each borrower in the sample, and Column (10) shows that the results are robust to weighting each observation by loan size.

The differences in returns seem to be driven by more-favorable loans to elite-connected borrowers. Panel B shows that there are no significant differences in repayment behavior,[46] suggesting that elite-connected households at least complied with their payments schedules. Panel C shows that elite-connected households seem to have obtained lower initial interest rates in the case of program loans (see Column 5). However, the double differences are not precisely estimated (Columns 7 to 10). In addition, elite-connected households do obtain substantially larger amounts, even exceeding the maximum amount allowed by program regulations.[47] This last result is consistent with the results in Section 5 showing that lower TFP predicts larger amounts of program credit. Put together, the results suggest that, for a similar level of risk, program committee members delivered more-favorable loans to elite-connected borrowers. As a result, there were forgone returns to the lender of the order of 2.7 percentage points per THB lent to elite-connected borrowers.

The previous results suggest favoritism, as better loan conditions to elite-connected households coincide with lower ex post returns on their program loans. The set of incentives faced by VFCs could explain these results. For instance, the central government offered village-level incentives and punishments for good and bad Village Fund management. However, there were no explicit sanctions of VFCs, and the costs of mismanagement were paid by the village and not internalized by the VFCs. In principle, community monitoring should align with VFC incentives, but, as allocating credit involves balancing attributes that are not easy to verify, the ability of poorer and unconnected community members to challenge the VFCs' decisions was limited. In addition, committee members did not seem to be compromising repayment, which could trigger government sanctions, but seem to be even willing to exceed lending caps in order to provide larger and cheaper

[45] Specifically, I control for household head's age, gender, and education, as well as average age and years of schooling in the household. I also include the number of adults, elderly household members and children in the household. All demographic characteristics were measured the year preceding the disbursement of the loan.

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

[47] Without special approval, program loans should not exceed THB 20,000.

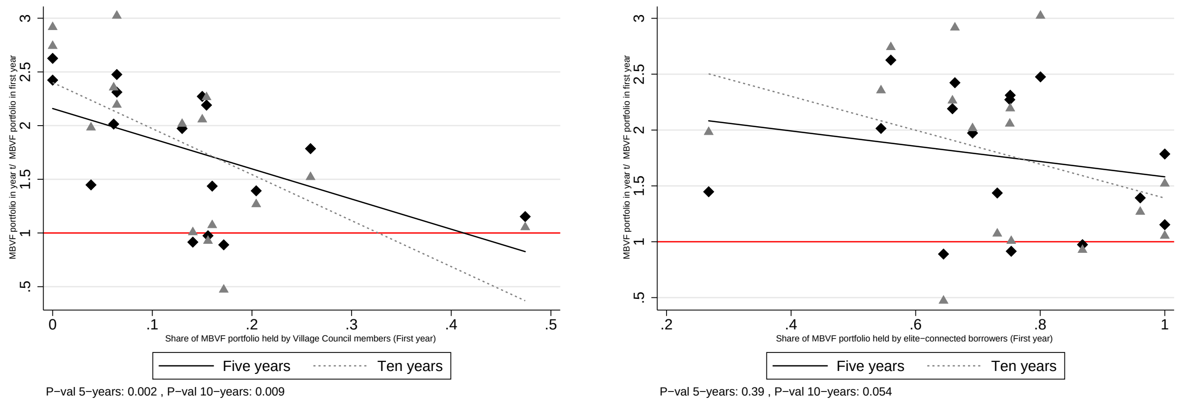
loans to elite-connected households. This behavior is consistent with models of incomplete contracts (Holmstrom and Milgrom, 1991; Hart et al., 1997): despite government incentives to induce high repayment, the decisions regarding loan attributes were left to the discretion of the VFCs.

It is also possible that the committee's choices simply reflect social norms or community preferences (Alatas et al., 2012). If that was the case, then the estimates would capture the financial cost of community preferences rather than the cost of favoritism. Given that loans from the comparison group (member-funded local credit groups) are also likely to be exposed to such preferences, concerns regarding social norms or community preferences seem unlikely to drive the results. In any case, the results suggest that decentralizing the allocation of credit does come at the cost of program profitability.

Finally, it is worth noting that these estimates are only valid for the subset of loans taken out by households who borrowed from both sources, which might be different than those who only borrowed from the program. Thus, these results suggest the existence of favoritism but are unable to explain the potential consequences for the full portfolio of program loans.

6.3 Connections and Program Capacity Building

VFCs could have allocated loans to the elite in order to boost the long-term success of the program or build institutional capacity at the local level. I focus on Village Fund growth over time as a measure for capacity building for two reasons. First, VFCs were supposed to reinvest revenues from interest on the Village Funds. Second, the central government committed to adding more funds to Village Funds with good performance. Thus, if VFCs lent to the elite in order to maximize Village Fund growth, then the Village Funds should have grown more in places in which the elite initially received more credit. That is not the case in the data. Figure III plots the relative growth of the Village Funds five and 10 years after their rollout, with respect to their initial size, as a function of the share of program resources allocated to village council members (panel a) and to elite-connected borrowers (panel b) during the first program year. Villages in which a large share of program credit was delivered to the members of the local government show substantially lower growth rates—in some cases negative, relative to villages in which program credit was not concentrated in the local elite. This correlation is stronger 10 years after the introduction of the funds.



(a) Elite borrowers

(b) Elite and elite-connected borrowers

Figure III: Village Fund Growth and Loans to Elite-Connected Borrowers

Note: The vertical axis measures total village fund portfolio 5 and 10 years after the initial rollout as a share of initial village fund portfolio. In panel a, the horizontal axis measures the share of the initial Village Fund portfolio allocated to members of the village council, and, in panel b, the share of the Village Fund portfolio allocated to elite-connected borrowers. P-values correspond to regressions of Village Fund growth on shares, and they are computed based on standard errors clustered at the village level.

7 Village-level Gains from Reducing Targeting Frictions

This section quantifies the potential returns of reallocating program credit. It considers two counterfactual exercises: The first involves reallocating excess program credit from elite-connected borrowers to nonborrowers. The second involves reallocating program credit from borrowers who would be ineligible based on repayment risk to low-risk households who did not obtain program credit.

In both cases there are two measures of interest: village-level output gains from reallocation, relative to the observed program allocation—$log(Y_v^c/Y_v)$, where Y_v^C and Y_v denote output under the counterfactual and actual regimes, and returns to reallocating program credit, which is measured as the THB changes in output $(Y_v^C - Y_v)$ per THB of reallocated program credit (B_v^C). Actual and counterfactual output are computed by imputing actual and counterfactual program credit into the production functions estimated in Section 4.2.2. For this, I assume that borrowers allocate a fraction κ of program credit to increase fixed capital, and the rest to working capital.[48]

Gains from redistributing excess credit to elite-connected borrowers. I use the coefficients associated with being connected to the elite from column (8) in panel B in Table III, as a proxy for excess program credit. This coefficient captures the amount of program credit that

[48]For each household, I compute output under each regime as: $Y_i = TFP_i(K_i + \kappa b_i)^{\beta_k}(M_i + (1-\kappa)b_i)^{\beta_m}(L)^{\beta_l}$, where b_i denotes program borrowing under each regime. I also use baseline averages of fixed capital (K), materials (M)—i.e., working capital, and labor (L).

elite-connected households obtained that is unexplained by neediness, TFP, credit history and other demographic characteristics. I then compute counterfactual allocations following three steps. First, I subtract the excess credit from the observed program borrowing amount of each elite-connected borrower. Second, I add up the total excess credit in each village (B_v^C), which on average accounts for 11% of program credit. Finally, within each village, I split B_v^C equally across nonborrowers.

Table VI: Gains from Reallocation of MBVF Loans

Panel A -From elite-connected borrowers to nonborrowers			
	$\kappa = 1$	$\kappa = 0.61$	$\kappa = 0$
Gains from eliminating information-transmission frictions and favoritism			
% Output gains	0.03%	1.50%	2.45%
THB Output change per THB of reallocated program credit	0.11	3.74	6.72
Share of reallocated MBVF portfolio		0.11	
Gains from eliminating only favoritism			
% Output gains	0.06%	0.79%	1.33%
THB Output change per THB of reallocated program credit	0.19	5.23	10.18
Share of reallocated MBVF portfolio		0.03	
Panel B - From elite-connected borrowers to unconnected households			
	$\kappa = 1$	$\kappa = 0.61$	$\kappa = 0$
Gains from eliminating information transmission frictions and favoritism			
% Output gains	0.00%	0.78%	1.28%
THB Output change per THB of reallocated program credit	0.06	2.91	5.12
Share of reallocated MBVF portfolio		0.11	
Gains from eliminating only favoritism			
% Output gains	0.00%	0.34%	0.57%
THB Output change per THB of reallocated program credit	0.10	4.31	7.84
Share of reallocated MBVF portfolio		0.03	
Panel C - From overincluded to overexcluded households (scoring model)			
	$\kappa = 1$	$\kappa = 0.61$	$\kappa = 0$
% Output gains	1.15%	1.33%	0.79%
THB Output change per THB of reallocated program credit	0.66	0.54	0.21
Share of reallocated MBVF portfolio		0.35	

Note: The table reports the average village-level output changes due to reallocation, average changes in village-level output per THB of reallocated program credit, and the average reallocated program credit as a share of village-level MBVF portfolio. Each panel reports results under three assumptions about the allocation of loans between fixed capital and working capital—κ. Panel A reports results from reallocating excess program credit, from elite-connected borrowers to nonborrowers. Panel B reports results from reallocating excess program credit, from elite-connected borrowers to unconnected households. Panel C reports results from reallocating credit from ineligible borrowers under the repayment-scoring criterion, to eligible nonborrowers. The results are computed using truncated measures of village-level output (top and bottom 5%).

Table VI shows the gains from reallocation for three alternative values of κ: assuming that all program credit is invested in fixed-capital ($\kappa = 1$), assuming that program credit is split between

fixed capital and working capital based on pre-program cost shares ($\kappa = 0.61$),[49] and assuming that all program credit is used as working capital ($\kappa = 0$). The first two rows of Panel A show that reallocating resources to nonborrowers would increase village-level output by 0.3% to 2.4% depending on the value of κ.[50] In particular, since nonborrowers had exhibited higher marginal returns to intermediate inputs (see Appendix Figure AVI), output gains are higher for lower values of κ. These gains are achieved by reallocating, on average, only 11% of the Village Fund porfolio. Thus, the returns of reallocating excess program credit from elite-connected borrowers to non-borrowers seem substantial: each $THB of program credit reallocated to nonborrowers increases village-level output by $ THB 0.11 to 6.72. Panel B shows a similar pattern in the case of reallocating excess program credit from connected borrowers to unconnected households (either borrowers or non-borrowers). Overall, the results imply important returns to eliminating connection-based targeting frictions.

Note that the gains arise from eliminating the elite-connection advantage in program credit due to both information-transmission frictions and favoritism. To isolate the elite-advantage due to favoritism, I replicate the reallocation exercise by using a network-centrality adjusted measure of excess credit due to connections. I do so by including network centrality as a control in the predictive model, and then using the coefficient associated with being connected to the elite to compute excess program borrowing.[51] Panel A and B show that the gains from eliminating favoritism account for half of the village-level output gains due to eliminating the total elite-connection advantage. These gains are achieved by reallocating only 3% of the Village Fund portfolio, suggesting high returns to reallocation: village-level output would increase by THB 0.19 to 10.8 per each THB of program credit obtained through favoritism. This is consistent with the idea that favoritism may create a large degree of misallocation (Hsieh and Klenow, 2009).

Gains from redistributing credit based on a credit-scoring benchmark criterion. Despite the presence connections-based allocative distortions, VFCs could outperform other policy-relevant ways of allocating credit. I analyze whether program borrowers would have been eligible to borrow under a repayment-based targeting criterion, and the potential gains from reallocating

[49] $\kappa = \frac{rK}{rK+M}$, where r is the cost of capital set to 0.05 following the Kaboski and Townsend (2011)'s estimates of borrowing and lending rates in Thai villages. K is the average stock of fixed capital at baseline, and M is the average baseline value of working capital—i.e, materials, intermediate inputs.

[50] This increase is sizable, as the Thai economy grew at an average rate of 5.12% during the five years following the recovery from the financial crisis of 1998. Source: World Bank national accounts data.

[51] This process suggests that Elite Borrowers get on average THB 3800 of extra credit, while non-elite households with connections to the elite get only THB 30 of extra credit after accounting for the number of links they have in pre-program networks.

resources. The repayment-based criterion is a policy-relevant benchmark, as microfinance institutions (MFIs) in developing countries often rely on scoring models to screen applicants (Schreiner, 2000), including Thailand's state-owned Bank for Agriculture and Agricultural Cooperatives (Limsombunc et al., 2005).[52]

I use self-reported information corresponding to over 3,800 pre-program loans from different types of lenders (formal and informal), combining it with several baseline household financial and demographic characteristics. I then estimate a LASSO model of the probability of having a delinquent payment, for a given loan, as a function of loan, lender, and borrower characteristics, as well as village and year fixed effects.[53] I purposely exclude consumption, consumption volatility, TFP, and connections data from the LASSO estimations in order to prevent these variables from being mechanically related to the repayment scores. I then use the LASSO coefficients to estimate delinquency risk for all potential borrowers and create within-village rankings based on the risk estimates. For each village, the households with the $k-th$ lowest positions (lower risk) are classified as the benchmark target group. Finally, k is picked such that the number of borrowers targeted by the hypothetical criterion coincides with the number of program borrowers in each village. Thus, it is possible to think of this exercise as a hypothetical change in the targeting criterion, while holding the number of program beneficiaries constant.

Table VII shows that 40% of potential borrowers who would have been targeted by the repayment criterion also obtained credit from the program. However, panel B shows that 34% of program borrowers would have been 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 previous sections: repayment risk was not a relevant dimension considered by committee members.

[52]In addition, it is unclear whether scoring models that rely on hard information and provide objective targeting rules outperform community-based approaches that rely on soft information and are more discretionary; most of the existing analysis is focused on comparing soft/hard information in for-profit banks(Paravisini and Schoar, 2013; Iyer et al., 2016; Bryan et al., 2015).

[53]To obtain a parsimonious model and minimize the risk of overfitting, I use the LASSO with a penalty parameter chosen through 10-fold cross-validation. Consistent with other scoring models in Thailand (Limsombunc et al., 2005), household debt-to-assets ratio predicts a 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 AXVI 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.

Table VII: Targeting Errors with Respect to the Hypothetical Repayment-Score Criterion

Panel A: Distribution of households by access to program credit and hypothetical eligibility criterion

	N	%
Group A: hhs who borrowed and would be eligible by the repayment-score criterion	278	39.15
Group B: hhs who did not borrow and would be eligible by the repayment-score criterion	139	19.58
Group C: hhs who borrowed but would be ineligible by the repayment-score criterion	147	20.70
Group D: hhs who did not borrow and would be ineligible	146	20.56

Panel B: Targeting errors with respect to hypothetical eligibility criterion

Inclusion error: % of ineligible hh who borrowed (C / (A+ C))	34.6%
Exclusion error: % of eligible hh who did not borrow (B/ (B+D))	48.8%

Panel C: Correlates of probability of hypothetical eligibility and baseline characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Per-capita consumption (logs)	0.095***						0.086**	0.130***	0.126**
	(0.031)						(0.041)	(0.045)	(0.055)
Pval (difference with MBVF)	0.08						0.47	0.67	0.96
Consumption volatility (log Coeff. of Variation)		-0.016					0.003	-0.029	0.001
		(0.028)					(0.032)	(0.032)	(0.038)
Pval (difference with MBVF)		0.08					0.05	0.30	0.25
TFP (logs)			0.048**				0.040*	0.030	0.054**
			(0.020)				(0.021)	(0.020)	(0.022)
Pval (difference with MBVF)			0.52				0.23	0.22	0.09
Access to instutitional credit (dummy)				0.110***			0.064	0.023	0.018
				(0.043)			(0.049)	(0.049)	(0.065)
Pval (difference with MBVF)				0.00			0.01	0.01	0.07
Ever missed a payment (dummy)					-0.105**		-0.147***	-0.116**	-0.117**
					(0.052)		(0.055)	(0.051)	(0.051)
Pval (difference with MBVF)					0.00		0.01	0.03	0.03
Connected with Village Council						0.006	-0.010	0.015	0.045
						(0.042)	(0.044)	(0.041)	(0.047)
Pval (difference with MBVF)						0.01	0.09	0.17	0.43
Observations	692	694	648	710	710	710	646	642	524
Adjusted R-squared	0.080	0.068	0.075	0.074	0.071	0.065	0.088	0.220	0.194
Within-village R-Squared	0.01	0.00	0.01	0.01	0.00	0.00	0.02	0.16	0.15
Excludes HH with no credit history	NO	NO	NO	NO	NO	NO	NO	NO	YES
Controls (shocks + demographics)	NO	NO	NO	NO	NO	NO	NO	YES	YES

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 nonborrowers who would have been eligible under the hypothetical repayment-based criterion (over-exclusion error). Panel C reports coefficients of a regression of the probability of being eligible by the repayment-based criterion on baseline characteristics. Columns (8)-(9) include demographic characteristics (household head's gender, age and education, average household age and education, number of adults, children younger than 15 and elderly in the household) and dummies indicating whether the household has experienced illness or issues with livestock, agricultural and non-agricultural production. Column (9) excludes households with no pre-program credit history. Standard errors are clustered at the household level. All regressions control for village fixed effects. P-values testing the null that the coefficients are different to those from a regression of program borrowing on the same baseline characteristics (see Table III) are estimated through seemingly unrelated regressions (SUR).

Panel C of Table VII reports correlates of repayment-based eligibility with baseline characteristics. Similar to the case of the allocation achieved by VFCs, the repayment-based criterion targets richer households with pre-program experience with institutional lenders. In contrast, it targets high-TFP households, while it does not give an advantage to elite-connected households. This suggest that elite-connected households, on average, were not less-risky borrowers. All results persist after controlling for demographic characteristics and exposure to pre-program shocks (see columns

(8) to (10)). One explanation is that scoring models provide an objective rule for the allocation of resources that may be less prone to the influence of connections. In contrast, while the VFCs could have used soft information to allocate resources more efficiently, it is not clear that the set of incentives of the VFCs are well aligned with program objectives.

Table VI shows that reallocating program credit from overincluded to overexcluded households would increase village-level output by 0.8% to 1.3%. Note, however, that these gains are achieved by redistributing, on average, 35% of program credit. Since the repayment-based criterion targets wealthier (and probably unconstrained) households, the returns from reallocation in this case seem smaller than those of reducing the elite-connection advantage.

8 Redistribution and Informal Credit Markets

The results from the previous sections suggest that connections to the local elite may create targeting frictions that result in unconnected households obtaining substantially fewer program resources. If these frictions prevented creditworthy, poor households from obtaining program loans, then other well-informed lenders in the village should be willing to lend to unconnected households. This section aims to test this hypothesis.

I analyze whether the credit supply shock generated by the program indirectly increased borrowing from nonprogram lenders 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 after the rollout of the program, aggregate borrowing increased by 24% in the sample villages. Second, the presence of active informal lenders in the study villages provides a potential mechanism for redistribution in the short run,[54] as opposed to institutional lenders, which may react slowly as they have to follow formal application processes.

I exploit monthly variation in the 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 use pre-program measures of elite connectedness to test for heterogeneity in borrowing from local informal lenders.

Identification of effects of the rollout of the program is achieved under the assumption that,

[54] For instance, using the first 88 waves of the Townsend Thai Project Monthly Survey, (Kinnan and Townsend, 2012) documented that among households without access to formal credit, being connected to a household with access contributes to consumption smoothing.

conditional on household time-invariant characteristics, the rollout of the program was not related to unobserved time-varying shocks that determined household decisions to obtain credit. This assumption seems plausible, as the timing of the program was mostly generated by differences in the timing of the establishment of VFCs, which is arguably orthogonal to the village economic environment. In order to examine the presence of pre-program trends and the dynamic program effects, I estimate the following flexible difference-in-differences model:

$$Y_{ivt} = \alpha_i + \delta_t + \sum_{j=-5, j\neq -1}^{j=4} \beta_j \mathbb{I}[\tau_{vt} = j] + \epsilon_{ivt} \tag{9}$$

Here, Y denotes total gross borrowing from local informal lenders by household i, in village v, at quarter t. I collapse time variation by quarters for a parsimonious graphical presentation of the results, and I focus on the six quarters preceding and following the introduction of the program. τ_{vt} denotes time to treatment for each village in a given quarter. Household fixed effects are denoted by α_i, and δ_t denotes a set of month and year fixed effects. The coefficients of interest are $\{\beta_j\}_{j=-5}^{4}$, which capture the difference between borrowing in period $\tau_{vt} = j$ relative to the quarter preceding the release of the funds ($\tau_{vt} = -1$), compared to the difference in borrowing in villages where funds were not released by that month.

Appendix Figure AVIII plots treatment effects of the rollout of the program on borrowing from informal credit, as well as confidence intervals based on 500 wild-bootstrap replications at the village level to account for a small number of clusters (Cameron and Miller, 2015). 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. Though imprecisely estimated, there is a clear jump in informal borrowing for unconnected households. Figure IV analyzes the effects of the rollout of the program on borrowing from relatives and nonrelatives. There are neither substantial nor significant pre-program trends, but there is a clear, significant surge in borrowing from relatives in the case of unconnected households after the release of program funds. Note that this pattern is the opposite of that observed in the context of program credit (see Appendix Figure AVII): while program credit was directly delivered to connected households, unconnected households indirectly obtained resources through relatives.

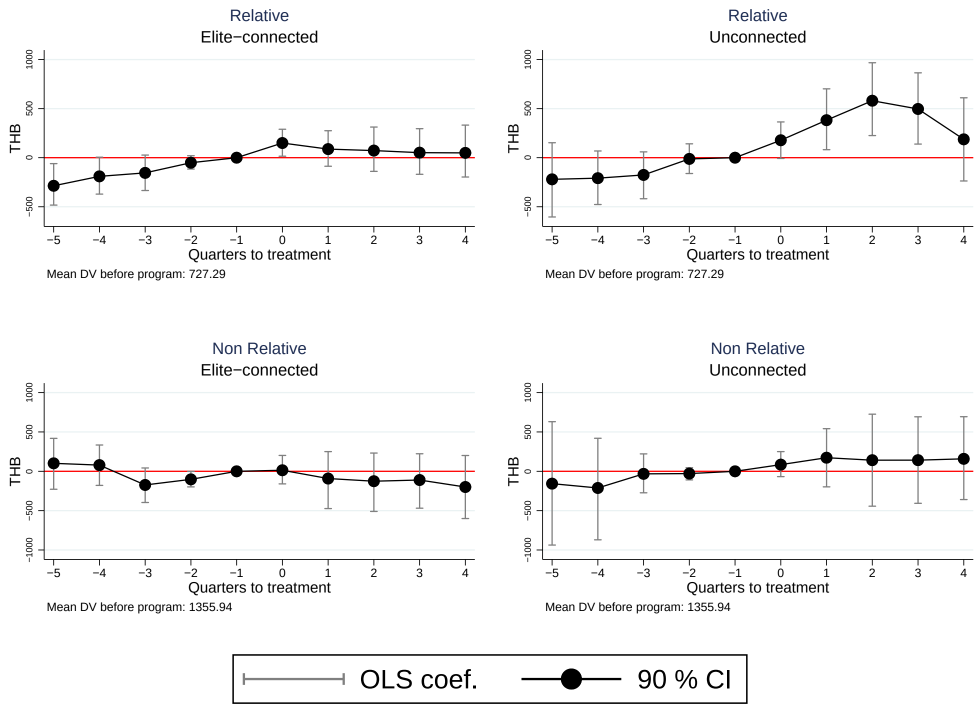


Figure IV: Short-Term Effects on Borrowing from Relatives and Non-Relatives

Note: The figure depicts OLS point estimates from a flexible difference-in-difference model following equation (9). 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 and winsorizing the top 1% of the depending variable. 90% Confidence intervals are based on 500 bootstrapped samples following the procedure suggested by Cameron and Miller (2015).

The previous set of results suggests that the allocative distortions in program credit were attenuated through redistribution. Coase (1960) predicts that in the absence of transaction costs, secondary private arrangements should overcome allocative distortions. The results in this section provide nuanced support for this prediction. First, while unconnected households ended up obtaining loans from relatives, the program-borrowing gap between connected and unconnected households was only partially offset. Table VIII shows that, on average, borrowing from local informal lenders increased by THB 470 in the case of unconnected households.[55] This increase only represents 10% of the connection-based gap in program borrowing (THB 4,400, see Appendix Table

[55] We compute difference-in-difference estimates using the following specification to approximate the average treatment effect corresponding to the post-rollout periods:

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

AXVII). Second, redistribution is costly, as around the rollout of the program the interest rates of loans taken out from relatives more than doubled those of program loans (14% and 6% annual, respectively). One implication is that differences in program participation based on connections are not mainly driven by lack of demand: unconnected households did indeed borrow, even at higher prices.

The targeting frictions may have generated arbitrage opportunities for secondary lending. Some households who obtained program loans may have used program resources to make loans to those who did not. Indeed, columns (7) to (9) in panel B of Table VIII suggest that the rollout of the program increased the probability of lending to other households in the case of elite-connected lenders. The latter result suggests that it is unlikely that general equilibrium effects, rather than redistribution, fully account for the connection-based patterns. An increase in overall economic activity should have created demand for liquidity for both connected and unconnected households, equally reducing their incentives for lending.

This set of results highlights the importance of informal markets for redistribution in the context of community-based targeted programs. Although informal credit markets may provide credit at high interest rates, they are also important for attenuating targeting distortions. An alternative explanation is that VFC members purposely targeted households with direct connections to council members, expecting that connected households would share credit with their unconnected relatives. While possible, that mechanism is hard to reconcile with the evidence of favoritism in the previous section. Overall, the results suggest that one important dimension to be considered in the design of targeting schemes is the role of redistribution. This is particularly important in programs that infuse large amounts of resources into local economies, as they are likely to have spillovers (Angelucci and De Giorgi, 2009; Kinnan and Townsend, 2012).

Table VIII: Short-Term Effects of the Program on Borrowing from Informal Lenders

Panel A: Effects on total borrowing and lending in informal markets

	Borrowing						**Lending**		
	Relatives			**Non-relatives**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	All	Connected	Unconnected	All	Connected	Unconnected	All	Connected	Unconnected
$Post_{vt}$	224.667**	144.465*	424.360*	135.302	178.916	85.621	497.649	568.285	544.470
	(97.820)	(99.447)	(206.351)	(92.115)	(136.506)	(95.439)	(488.744)	(751.044)	(483.713)
Bootstrap p-value	[0.024]	[0.091]	[0.064]	[0.420]	[0.316]	[0.719]	[0.176]	[0.400]	[0.156]
Observations	23,013	15,030	7,983	23,019	14,966	8,053	23,783	15,522	8,261
R-squared	0.681	0.740	0.555	0.684	0.685	0.671	0.834	0.870	0.647
P-val (Connected-Unconnected)		[0.22]			[0.461]			[0.976]	
Baseline DV mean	592	623.8	532.1	939.5	1166	517	4888	6023	2764
# of households	671	439	232	669	438	231	685	444	241

Panel B: Effects on the probability of borrowing and lending in informal markets

	Borrowing						**Lending**		
	Relatives			**Non-relatives**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	All	Connected	Unconnected	All	Connected	Unconnected	All	Connected	Unconnected
$Post_{vt}$	0.010	-0.002	0.033*	-0.000	-0.001	0.003	0.019**	0.021*	0.015
	(0.009)	(0.010)	(0.018)	(0.009)	(0.013)	(0.011)	(0.009)	(0.013)	(0.014)
Bootstrap p-value	[0.372]	[0.916]	[0.080]	[0.972]	[0.988]	[0.924]	[0.020]	[0.044]	[0.236]
Observations	23,228	15,143	8,085	23,228	15,143	8,085	25,560	16,488	9,072
R-squared	0.640	0.680	0.559	0.647	0.652	0.604	0.791	0.784	0.805
P-val (Connected-Unconnected)		[0.068]			[0.808]			[0.676]	
Baseline DV mean	0.0707	0.0733	0.0658	0.105	0.130	0.0577	0.225	0.239	0.200
# of households	671	439	232	671	439	232	710	458	252

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

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 moneylenders 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. Dependent variables are winsorized with respect to the top 1%. Panel A shows results for total gross borrowing and lending (winsorizing the top 1% of observations), and Panel B reports results for probability of holding a loan, and the probability of lending to other households. 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.

Robustness. The program was initially rolled out in different calendar months across villages and reached different villages in different phases of the agricultural cycle. To test if the results were driven by the phase of the agricultural cycle as opposed to the program rollout, I constructed a placebo sample by allocating a placebo treatment to each villages exactly 24 months from the actual rollout date.[56] For example, if the program funds were released in December of 2001 in village v, the corresponding placebo shock would have taken place in December of 1999, and captured potential dynamics associated with the agricultural cycle. I then estimated the placebo effects using equation (9). Appendix Figure AIX reports flexible difference-in-differences estimates corresponding to the placebo shock using pre-program data. Reassuringly, the placebo exercise does not yield patterns that are similar to the ones observed after the effective rollout.

9 Concluding Remarks

Three of the largest lending programs in developing countries decentralize the allocation and management of loans to community members (see Bardhan and Mookherjee (2006) in the case of India, Kaboski and Townsend (2012) in the case of Thailand, and Cai et al. (2017) in the case of China). Given that it is well-documented that community members have important information regarding neediness (Alderman, 2002; Alatas et al., 2012) and returns (Hussam et al., 2017), it is natural to ask whether they indeed use such information to balance issues of risk, poverty, and productivity in lending schemes that are also prone to favoritism. This paper argues that decentralizing the allocation of credit to locally elected committees can lead to allocations that do not prioritize risk, neediness, or productivity. Instead, more-powerful households, both in terms of wealth and connections, ended up obtaining more program credit, and unconnected households ended up obtaining informal loans at higher rates. In turn, there are forgone financial returns from allocating resources to elite-connected households with subsequent consequences for program sustainability. Moreover, a traditional repayment-based targeting criterion would have eliminated the influence of connections and improved productive efficiency.

The evidence in this paper has three core policy implications. First, it suggests that there are costs in terms of productive efficiency, risk, and returns associated with decentralized approaches. In a sense, although these approaches are easier and cheaper to implement than traditional banking models, it seems that the cost is transmitted from the policymaker to the community. Thus, the

[56] I computed time to treatment in the placebo sample as $\tilde{\tau}_{v,t} = \tau_{v,t} + 24$, where τ denotes time to treatment with respect to the actual shock and $\tilde{\tau}$ denotes time to the placebo shock.

choice between centralized and decentralized approaches to expanding access to credit involves balancing the costs faced by the community against the costs of relying on financial institutions to expand credit.

Second, if policymakers are willing to face such costs with the aim of building local capacity, then it is important to design incentives to hold local leaders accountable. While community monitoring can increase accountability (Björkman and Svensson, 2009), monitoring requires that community members have the means to verify whether the decisions of local leaders are in their best interest. When resources are allocated based on attributes that are hard to verify, challenging the actions of community members is costly and community monitoring might not be enough (Reinikka and Svensson, 2004). In the Thai case, the central government offered village-level incentives and punishments for good and bad Village Fund management. However, there were no explicit sanctions on members of the committees. As the poorer and unconnected may not have the means to challenge the program committees, committee members may not have internalized the cost of their actions. Unsurprisingly, Village Funds grew less—some even contracted—when there was a larger share of loans allocated to the local elite.

Third, the results also suggest that policymakers should consider the possibility of redistribution and the costs associated with it. One justification for subsidized credit programs is the idea that cheaper credit would crowd out borrowing from usurers. The results from this paper highlight a more nuanced view of the role of informal markets: they can attenuate allocative distortions, but that process is costly.

Finally, it should be noted that despite the evidence of allocative distortions in program lending, the results in this study do not imply that the MBVF program reduced village welfare. The MBVF program did increase consumption (Kaboski and Townsend, 2012), and it mildly reduced capital-market failures, relative to a no-program scenario (Shenoy, 2017b). However, the program only increased the profits of high-TFP entrepreneurs who were able to borrow (Banerjee et al., 2018), and a simple cash-transfer program would have been more cost-effective (Kaboski and Townsend, 2011). This paper provides an explanation for such modest results: connection-based allocative distortions undermined the success of the program. More broadly, it suggests that targeting frictions may explain the nontransformative effects from other microcredit programs (Banerjee et al., 2015).

Diego A. Vera-Cossio. Research Department (RES). Inter-American Development

Bank. 1300 New York Ave. NW, Washington, D.C. 20577, USA.

References

Acemoglu, D., T. Reed, and J. A. Robinson (2014). Chiefs: Economic development and elite control of civil society in sierra leone. *Journal of Political Economy 122*(2), 319–368.

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

Agarwal, S., B. Morais, C. Ruiz Ortega, and J. Zhang (2016, February). The political economy of bank lending : evidence from an emerging market. Policy Research Working Paper Series 7577, The World Bank.

Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, R. Purnamasari, and M. Wai-Poi (2019, May). Does elite capture matter? local elites and targeted welfare programs in indonesia. *AEA Papers and Proceedings 109*, 334–39.

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.

Alderman, H. (2002). Do local officials know something we don?t? decentralization of targeted transfers in albania. *Journal of Public Economics 83*(3), 375 – 404.

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.

Banerjee, A., E. Breza, E. Duflo, and C. Kinnan (2019, October). Working Paper 26346, National Bureau of Economic Research.

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

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

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.

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 (2006). 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.

Björkman, M. and J. Svensson (2009, 05). Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda*. *The Quarterly Journal of Economics 124*(2), 735–769.

Björkman, M. and J. Svensson (2010). When is community-based monitoring effective? evidence from a randomized experiment in primary health in uganda. *Journal of the European Economic Association 8*(2?3), 571–581.

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.

Bryan, G., D. Karlan, and J. Zinman (2015, August). Referrals: Peer screening and enforcement in a consumer credit field experiment. *American Economic Journal: Microeconomics 7*(3), 174–204.

Cai, S., A. Park, and S. Wang (2017). Microfinance can raise incomes: Evidence from a Randomized Control Trail in China. Working paper.

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.

Casey, K., R. Glennerster, and E. Miguel (2012, 11). Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan*. *The Quarterly Journal of Economics 127*(4), 1755–1812.

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

Coase, R. H. (1960). The problem of social cost. *The Journal of Law & Economics 3*, 1–44.

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

Collard-Wexler, A. and J. De Loecker (2016, July). Production function estimation with measurement error in inputs. Working Paper 22437, National Bureau of Economic Research.

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.

Deaton, A. (1997). *The Analysis of Household Surveys : A Microeconometric Approach to Development Policy.* The World Bank.

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.

Giné, X., B. Barboza Ribeiro, and I. Valley (2019, 9). Targeting Inputs: Experimental Evidence from Tanzania. Working Paper Series 9013, The World Bank.

Goldstein, M. and C. Udry (2008). The profits of power: Land rights and agricultural investment in ghana. *Journal of Political Economy 116*(6), 981–1022.

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

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.

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.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing tfp in china and india*. *The Quarterly Journal of Economics 124*(4), 1403–1448.

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.

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.

Kim, K. i., A. Petrin, and S. Song (2016). Estimating production functions with control functions when capital is measured with error. *Journal of Econometrics 190*(2), 267 – 279. Endogeneity Problems in Econometrics.

Kinnan, c., K. Samphantharak, R. Townsend, and D. Vera-Cossio (2019). Insurance and propagation in village networks. Technical report.

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

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.

Meager, R. (2019, January). Understanding the average impact of microcredit expansions: A bayesian hierarchical analysis of seven randomized experiments. *American Economic Journal: Applied Economics 11*(1), 57–91.

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.

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

Paravisini, D. and A. Schoar (2013, August). The incentive effect of scores: Randomized evidence from credit committees. Working Paper 19303, National Bureau of Economic Research.

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

Reinikka, R. and J. Svensson (2004, 05). Local Capture: Evidence from a Central Government Transfer Program in Uganda*. *The Quarterly Journal of Economics 119*(2), 679–705.

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

Samphantharak, K. and R. M. Townsend (2018, February). Risk and return in village economies. *American Economic Journal: Microeconomics 10*(1), 1–40.

Schoenherr, D. (2018, January). Political connections and allocative distortions. *Journal of Finance Forthcoming*.

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.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory 21*(2), 265 – 293.

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

A APPENDIX: Supplementary Results

A.1 Supplementary Figures

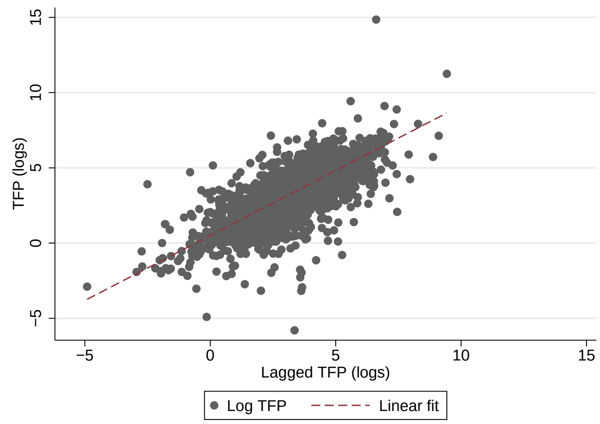
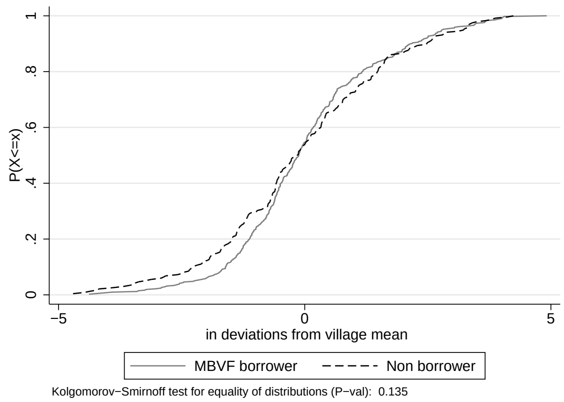
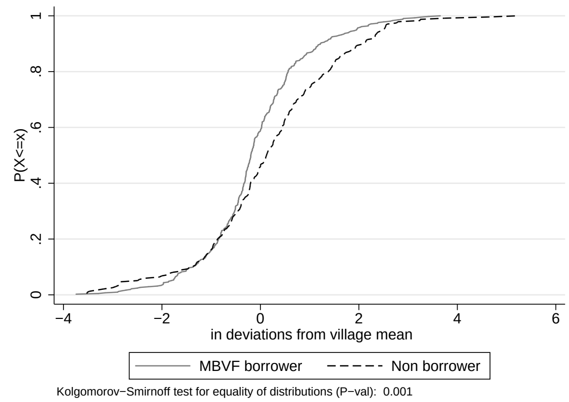


Figure AV: TFP persistence

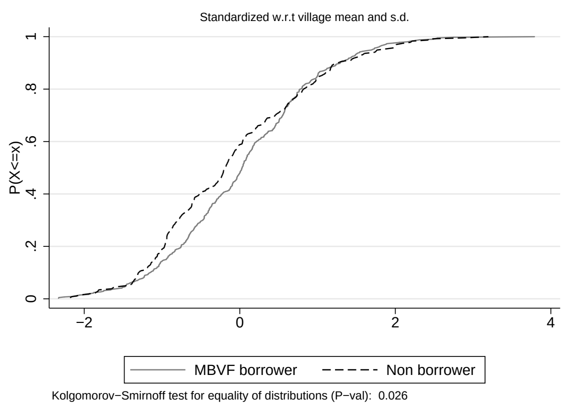
Note: The figure depicts estimated log TFP as a function of its first lag. TFP are estimated using the System-GMM method. TFP measures are winsorized with respect to the top and bottom 1%.



(a) Capital (k)



(b) Intermediates (m)



(c) Labor (l)

Figure AVI: Distribution of Baseline Marginal Revenue Products by Program Participation

Note: The figures depict CDFs of baseline log marginal revenue products for program borrowers and nonborrowers. All measures are centered with respect to the village-sector mean. Marginal products are computed based on the factor elasticities reported in Columns (3) and (4) of Panel B in Table AXI and our preferred estimates of TFP. All measures are truncated with respect to the top and bottom 1% values. Program participation is measured as an indicator of whether a household borrowed from the program within the first two years of program implementation.

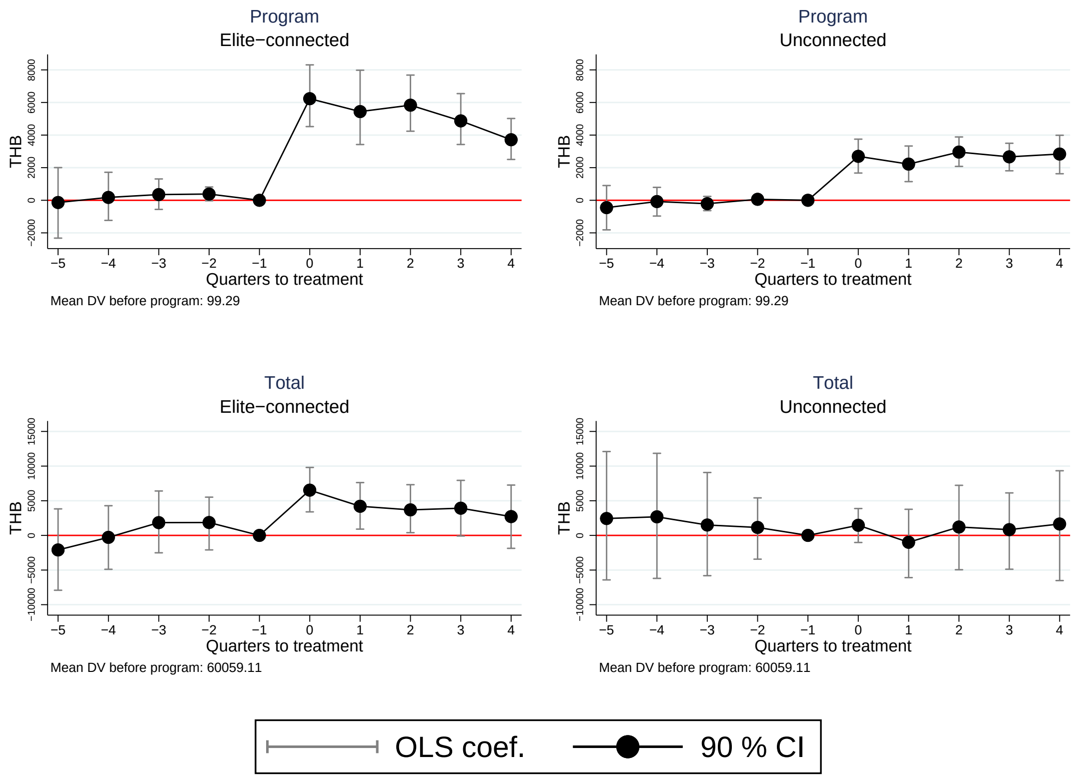


Figure AVII: Short-Term Effects on Program and Total Credit by Elite-Connectedness

Note: The figure depicts OLS point estimates from a flexible difference-in-difference model following equation (9). 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, and winsorizing the dependent variables with respect to the top 1%. 90% Confidence intervals are based the 5th and 95th percentiles of the bootstrap distribution of point estimates obtained by re-sampling villages following the wild-bootstrap procedure suggested by Cameron and Miller (2015) to account for a small number of clusters.

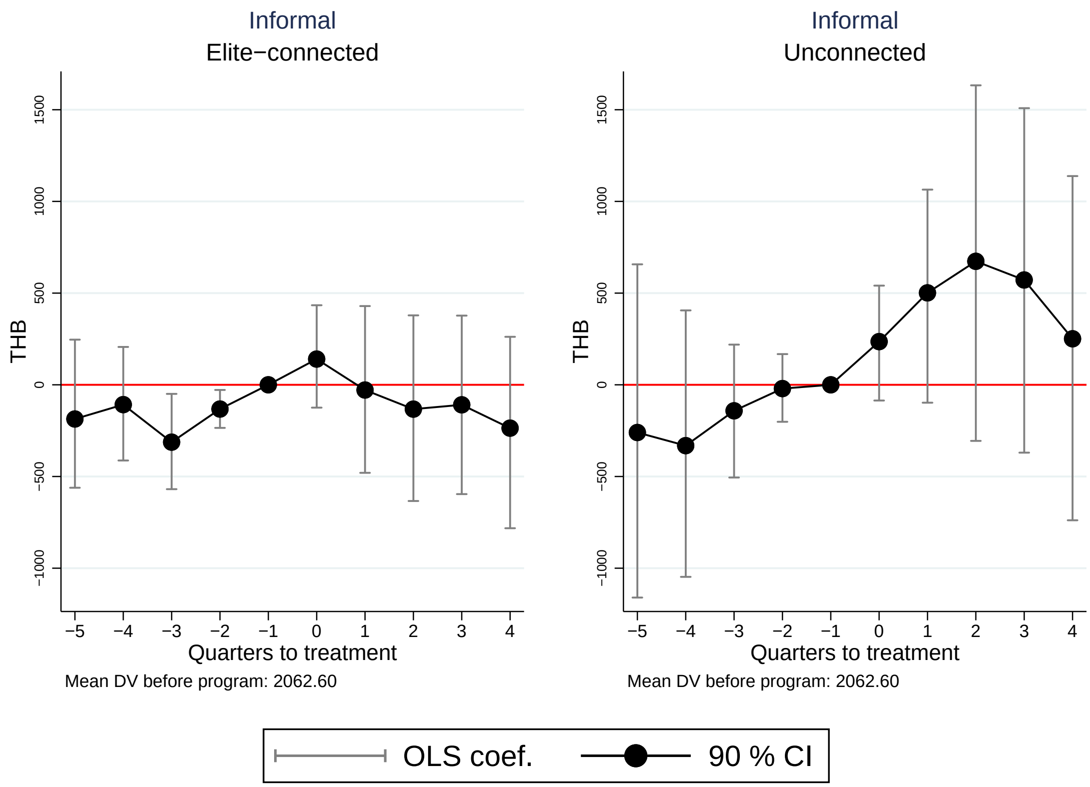


Figure AVIII: Short-Term Effects on Credit from Informal Lenders by Elite-Connectedness

Note: The figure depicts OLS point estimates from a flexible difference-in-difference model following equation (9). 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, and winsorizing the dependent variables with respect to the top 1%. 90% Confidence intervals are based the 5th and 95th percentiles of the bootstrap distribution of point estimates obtained by re-sampling villages following the wild-bootstrap procedure suggested by Cameron and Miller (2015) to account for a small number of clusters. Informal lenders include personal lenders and relatives in the villages.

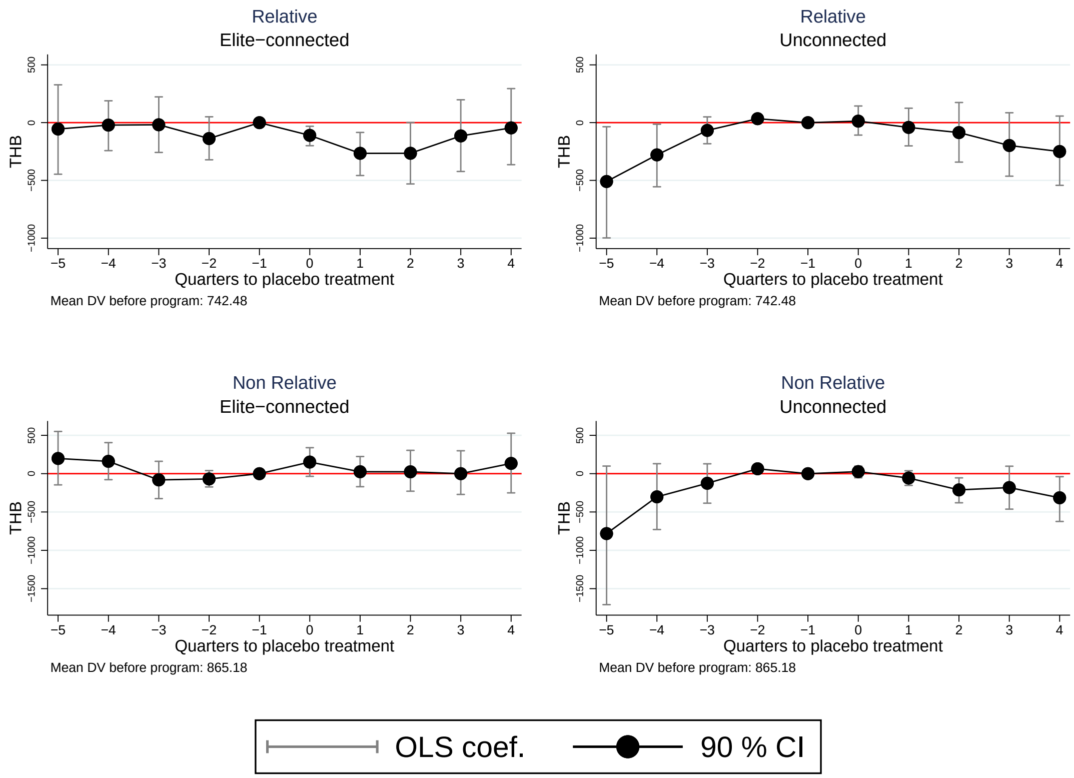


Figure AIX: Changes in Informal Borrowing in the Placebo Sample

Note: The figure depicts OLS point estimates from a flexible difference-in-difference model following equation (9). 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 placebo shock (2 years before the actual rollout of the program). Estimations were performed using all the available observations for the 18 months before and after the placebo shock in each village. 90% Confidence intervals are based the 5th and 95th percentiles of the bootstrap distribution of point estimates obtained by re-sampling villages following the wild-bootstrap procedure suggested by Cameron and Miller (2015) to account for a small number of clusters. Informal lenders include personal lenders and relatives in the villages.

A.2 Supplementary Tables

Table AIX: Summary Statistics: Loan Characteristics in the Village Financial System

Panel A: Distribution of loan by type of lender	Formal		Quasi-formal		Personal lender		Relative		Village Fund	
	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 (THB)	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%

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, women'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 AX: Correlates between Current First-Differences of Inputs and Output 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 AXI: Estimates of factor elasticities

Panel A: Reduced form estimates				
	Diff GMM		System GMM	
	(1)	(2)	(3)	(4)
	Farm	Non Farm	Farm	Non Farm
y_{t-1}	0.18*	0.40***	0.67***	0.75***
	(0.096)	(0.080)	(0.097)	(0.048)
k_t	0.28	0.26	0.32*	0.17
	(0.226)	(0.263)	(0.177)	(0.175)
k_{t-1}	-0.19	-0.10	-0.28	-0.08
	(0.164)	(0.176)	(0.175)	(0.171)
m_t	0.35***	0.20***	0.39***	0.33***
	(0.062)	(0.061)	(0.068)	(0.055)
m_{t-1}	-0.18***	-0.25***	-0.27***	-0.23***
	(0.056)	(0.039)	(0.049)	(0.045)
l_t	0.16**	0.31***	0.12*	0.31***
	(0.065)	(0.101)	(0.071)	(0.118)
l_{t-1}	-0.02	-0.21	-0.06	-0.30***
	(0.060)	(0.131)	(0.067)	(0.112)
Hansen stat	144.8	142.1	185.2	196.3
DF	122	122	170	170
P-val(Hansen)	0.0776	0.103	0.201	0.0813
Panel B: Estimates of factor elasticities				
	Farm	Non Farm	Farm	Non Farm
ρ	0.25***	0.57***	0.66***	0.72***
	(0.06)	(0.05)	(0.06)	(0.03)
β_k	0.20	0.25	0.15***	0.26***
	(0.16)	(0.19)	(0.06)	(0.09)
β_m	0.30***	0.37***	0.40***	0.33***
	(0.04)	(0.03)	(0.04)	(0.04)
β_l	0.15***	0.26***	0.14***	0.28***
	(0.04)	(0.07)	(0.04)	(0.08)
Obs	3,283	2,279	3584	2586
Returns to scale	0.66	0.88	0.69	0.87
Chi2 (RTS=1)	4.32	0.32	17.26	1.31
P-Val (RTS=1)	0.04	0.57	0.00	0.25
J-stat OID-OMD	6.30	20.56	1.69	1.70
P-val (OID-OMD)	0.10	0.00	0.64	0.64
Panel C: Summary statistics and correlations across methods				
	Farm	Non Farm	Farm	Non Farm
Mean	4.72	2.66	4.58	2.72
SD	0.83	1.16	0.79	1.15
Correlates with System-GMM	1.010***	1.005***	N.A.	N.A.
	(0.017)	(0.003)		

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

Note: Panel A reports reduced-form estimates for Farm and Non-farm sectors. Columns (1) and (2) report results from the Arellano and Bond (1991)'s diff-GMM estimator using lagged levels of inputs and output as instruments of contemporaneous first differences. Columns (3) and (4) report coefficients estimated through Blundell and Bond (1998)'s system-GMM approach using lagged levels of inputs and output as instruments for contemporaneous first differences and suitable lags of first-differences as instruments for contemporaneous levels of input and output. All regressions use lags 3 to 5 of output and input as instruments (see Appendix Section B.1 for details). 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. 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 and productivity persistence parameters (ρ), which were estimated through Optimal Minimum Distance (OMD) using the reduced-form variance-covariance matrix as weighting matrix. Panel C reports summary statistics of baseline TFP and correlation of TFP across methods, which corresponds to the regression coefficient from a regression of log TFP (diff-GMM method) on log TFP (system GMM method) and village fixed effects.

Table AXII: Factor Elasticities Using Alternative Methods

Panel A: Estimates of factor elasticities

	Pre-program only		**Measurement Error**		**Excluding labor provision**	
	Farm	Non Farm	Farm	Non Farm	Farm	Non Farm
	(1)	(2)	(3)	(4)	(5)	(6)
β_k	0.09***	0.24**	0.30	0.39*	0.06	0.24***
	(0.03)	(0.11)	(0.24)	(0.22)	(0.05)	(0.05)
β_m	0.47***	0.16	0.33***	0.38***	0.45***	0.32***
	(0.10)	(0.15)	(0.04)	(0.04)	(0.04)	(0.03)
β_l	0.38**	0.86***	0.08*	0.28***	0.14***	0.36***
	(0.16)	(0.33)	(0.05)	(0.07)	(0.05)	(0.05)
Obs	237	140	3283	2279	3505	1921
Returns to scale	0.95	1.26	0.72	1.05	0.65	0.91
Chi2 (RTS=1)	0.16	0.90	1.49	0.04	33.87	2.20
P-Val (RTS=1)	0.69	0.34	0.22	0.84	0.00	0.14
J-stat OID-OMD	N.A.	N.A.	7.35	18.14	3.61	4.96
P-val (OID-OMD)	N.A.	N.A.	0.06	0.00	0.31	0.17

Panel B: Summary statistics and correlations across methods

	Farm	Non Farm	Farm	Non Farm	Farm	Non Farm
Mean	2.80	0.08	3.66	0.66	5.19	2.52
SD	0.83	1.24	0.85	1.23	0.96	1.42
Correlates with System-GMM	0.834***	0.835***	0.872***	0.914***	0.473***	0.269***
	(0.016)	(0.017)	(0.016)	(0.007)	(0.069)	(0.038)

Standard errors in parentheses

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

Note: Panel A reports OMD estimates for Farm and Non-farm sectors using three different methods. Columns (1) and (2) report results using only the three years of pre-program data for the estimation of a restricted model that does not allow for fixed effects. Columns (3) and (4) report coefficients estimated by instrumented first-differences in capital with suitable lags of investment. Columns (5) and (6) report system-GMM estimates that exclude operations from off-household labor activities. Estimates in columns (1) and (2) use first and second lags of capital and input and second lags of labor and intermediate inputs as instruments. Estimates from columns (3) to (6) use lags 3 to 5 of output and inputs as instruments (see Appendix Section B.1 for details). 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. 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 summary statistics of baseline TFP and correlation of TFP with the benchmark specification for each alternative method.

Table AXIII: Baseline Characteristics and Program Participation

Panel A - Neediness

	OLS Coefficient	S.E. (Diff)	P-val	Hochberg Adj-Pval
log Per capita consumption	0.16***	(0.02)	0.000	0.00
log Per capita wealth	0.05***	(0.01)	0.000	0.00
Consumption volatility (log coef. of variation)	-0.10***	(0.03)	0.000	0.00
Health symptoms	0.00	(0.00)	0.482	0.96
Problems with cultivation operations	0.02	(0.01)	0.145	0.44
Problems with livestock operations	0.04***	(0.01)	0.000	0.00
Problems with business operations	0.00	(0.01)	0.696	0.70
Panel B - Productivity				
	OLS Coefficient	S.E. (Diff)	P-val	Hochberg Adj-Pval
log TFP (Blundell & Bond)	0.03	(0.02)	0.120	0.72
log TFP (Measurement Error in k)	0.02	(0.01)	0.211	0.84
log TFP (Excluding wage-labor revenues)	0.01	(0.01)	0.457	0.91
log TFP (pre-program estimation sample)	0.01	(0.01)	0.360	1.00
log TFP Marginal Revenue Product of Capital (MRPK)	-0.00	(0.01)	0.921	0.92
log TFP Marginal Revenue Product of Intermediates (MRPM)	-0.05***	(0.01)	0.000	0.00
log TFP Marginal Revenue Product of Labor (MRPL)	0.03	(0.02)	0.171	0.85
Panel C- Risk and credit history				
	OLS Coefficient	S.E. (Diff)	P-val	Hochberg Adj-Pval
	(1)	(2)	(3)	(4)
Ever borrowed from institutional lender	0.34***	(0.04)	0.000	0.00
Leverage rate	-0.09	(0.07)	0.180	0.72
Income volatility (log coef. of variation)	0.001	(0.02)	0.969	1.00
Share of loans with delinquent payments	-0.00	(0.14)	0.989	0.99
Missed a payment (dummy)	0.07	(0.05)	0.162	0.81
Share of loans with term extensions	0.09	(0.07)	0.188	0.56
Extended loan (dummy)	0.15***	(0.05)	0.004	0.03
Panel D - Connections with local leaders				
	OLS Coefficient	S.E. (Diff)	P-val	Hochberg Adj-Pval
Any connection to council members	0.16***	(0.04)	0.000	0.00
Inverse distance to council members	0.48***	(0.09)	0.000	0.00
# of links with council members	0.11***	(0.02)	0.000	0.00
Connectedness PCA index	0.21***	(0.04)	0.000	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. Marginal revenue products are computed using the main TFP measure (Blundell and Bond (2000)'s method). The correlations between program credit and baseline share of loans with delinquent payments, the indicator of ever missing a payment, share of extended loans and the indicator of ever extending the term of a loan are estimated over a sub-sample of households with self-reported information about their credit history.

Table AXIV: Correlates of Program Participation and Total Program Borrowing with Alternative Measures of Baseline TFP and Marginal Products

Panel A: Program participation

	TFP measures Measurement error	 Excluding labor operations	 Pre-program data only	Marginal products
	(1)	(2)	(3)	(4)
Per-capita consumption (logs)	0.154***	0.175***	0.155***	0.155***
	(0.039)	(0.039)	(0.039)	(0.039)
Consumption volatility (log Coeff. of Variation)	-0.072**	-0.070**	-0.072**	-0.073**
	(0.030)	(0.032)	(0.030)	(0.030)
TFP (logs)	-0.005	-0.009	-0.008	
	(0.013)	(0.012)	(0.014)	
Marginal revenue product of Capital (logs)				-0.002
				(0.012)
Marginal revenue product of Intermediates (logs)				-0.012
				(0.015)
Marginal revenue product of Labor (logs)				-0.021
				(0.020)
Access to institutional credit	0.201***	0.190***	0.203***	0.198***
	(0.051)	(0.054)	(0.052)	(0.053)
Ever missed a payment	0.025	0.031	0.023	0.021
	(0.049)	(0.050)	(0.049)	(0.049)
Connected with Village Council	0.096**	0.095**	0.097**	0.091*
	(0.045)	(0.048)	(0.045)	(0.046)
Observations	642	588	642	642
Adjusted R-squared	0.15	0.13	0.15	0.15
Within-village adjusted R-squared	0.09	0.08	0.09	0.09

Panel B: Average program borrowing

	TFP measures Measurement error	 Excluding labor operations	 Pre-program data only	Marginal products
	(1)	(2)	(3)	(4)
Per-capita consumption (logs)	4,795.035***	5,363.482***	5,054.347***	5,020.536***
	(1,356.630)	(1,481.872)	(1,361.152)	(1,428.112)
Consumption volatility (log Coeff. of Variation)	-1,221.224	-1,320.586	-1,276.212*	-1,229.168
	(759.247)	(861.556)	(761.709)	(774.250)
TFP (logs)	-684.099**	-685.731**	-887.045***	
	(303.577)	(290.579)	(332.400)	
Marginal revenue product of Capital (logs)				-144.298
				(231.621)
Marginal revenue product of Intermediates (logs)				-422.042
				(323.204)
Marginal revenue product of Labor (logs)				-805.567*
				(455.570)
Access to institutional credit	3,540.143***	3,282.973***	3,650.949***	3,137.194***
	(908.496)	(1,016.469)	(920.968)	(967.070)
Ever missed a payment	1,393.296	1,664.826	1,338.559	1,409.975
	(1,054.026)	(1,122.250)	(1,044.807)	(1,067.195)
Connected with Village Council	2,022.572**	2,301.551**	2,038.609**	1,809.904**
	(861.184)	(944.983)	(862.625)	(855.200)
Observations	614	562	614	614
Adjusted R-squared	0.303	0.286	0.306	0.299
Within-village adjusted R-squared	0.131	0.125	0.136	0.126

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

Note: The table presents OLS coefficients from regressions of program participation on baseline characteristics. Column (1) uses TFP calculations correcting for potential measurement error in capital as a regressor. Columns (2) and (3) use alternative TFP measures excluding wage-labor activities and using only pre-program data for estimation, respectively. Column (4) includes marginal products (in logs) of fixed capital, intermediate inputs, and labor instead of productivity. Marginal products are computed based on the factor elasticities reported in Columns (3) and (4) of Panel B in Table AXI and our preferred estimates of TFP. All regressions include village fixed effects. Standard errors are clustered at the household level.

Table AXV: Baseline Correlates between Connections with the Village Council and Creditworthiness, and Profitability

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

VARIABLES	(1) Delinquent loans (share)	(2) Missed a payment	(3) Term extensions (share)	(4) Loan term extension	(5) Per capita consumption	(6) Cons. Volatility	(7) TFP
Connected with Village Council	-0.009	0.064*	-0.030	0.081*	0.036	0.023	0.072
	(0.018)	(0.038)	(0.035)	(0.047)	(0.055)	(0.060)	(0.091)
Observations	544	544	544	544	692	694	648
R-squared	0.068	0.093	0.056	0.117	0.123	0.035	0.172

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

VARIABLES	(1) Delinquent loans (share)	(2) Missed a payment	(3) Term extensions (share)	(4) Loan term extension	(5) Per capita consumption	(6) Cons. Volatility	(7) TFP
Village council member	0.175**	-0.005	0.132**	0.001	-0.142***	0.232***	0.066
	(0.068)	(0.047)	(0.064)	(0.004)	(0.055)	(0.073)	(0.152)
Direct transactions with council member	0.056	-0.033	0.084*	0.004	0.024	0.021	0.064
	(0.040)	(0.036)	(0.049)	(0.003)	(0.037)	(0.059)	(0.097)
First-degree relative with council member	-0.049	-0.007	-0.059	-0.002	0.002	-0.063	0.051
	(0.054)	(0.036)	(0.053)	(0.003)	(0.055)	(0.070)	(0.117)
Base category mean: unconnected	0.06	0.13	0.33	0.61	7.18	-0.47	3.71
Observations	544	544	544	658	694	692	648
R-squared	0.100	0.054	0.117	0.061	0.132	0.131	0.170

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

Note: The table presents OLS coefficients from a regression of baseline characteristics on different measures of connectedness with the village council. Columns (1) to (4) report estimates on a subsample of households who had pre-existing credit history, either from formal or informal lenders. Columns (6) and (7) use all sample households in the estimations. All regressions include village fixed effects. Standard errors are clustered at the household level.

Table AXVI: 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 AXVII: Short-Term Effects on Program and Total Credit

Panel A: Effects on credit from the program

	Connected		Unconnected	
	(1)	(2)	(3)	(4)
VARIABLES	Any loan	Total borrowing	Any loan	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	
	(1)	(2)	(3)	(4)
VARIABLES	Any loan	Total borrowing	Any loan	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. 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.

ONLINE APPENDIX: FOR ONLINE PUBLICATION ONLY

B ONLINE APPENDIX

B.1 Online Appendix: Production function estimation

B.2 Identification of the Gross-Revenue function

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 household productivity $a_{i,t}$ has a time-invariant component α_i, and a time-varying component $\omega_{i,t}$ which follows a first-order autoregressive process:

$$a_{i,t} = \alpha_i + \omega_{i,t} \tag{10}$$

$$\omega_{i,t} = \rho\omega_{i,t-1} + \zeta_{i,t} \tag{11}$$

where $\zeta_{i,t}$ denotes unforeseen productivity shocks which are unobserved by the researcher but known to the household. Plugging in (10) and (11) into (5), and subtracting $\rho y_{i,t-1}$ from both sides of the equation yields:

$$\begin{aligned} y_{i,t} &= \gamma_{l1}l_{i,t} + \gamma_{l2}l_{i,t-1} + \gamma_{k1}k_{i,t} + \gamma_{k2}k_{i,t-1} + \gamma_{m1}m_{i,t} + \gamma_{m2}m_{i,t-1} + \gamma_y y_{i,t-1} \\ &\quad +\alpha_i + \zeta_{i,t} + \epsilon_{i,t} - \rho\epsilon_{i,t-1} \end{aligned} \tag{12}$$

where $\gamma_{j1} = \beta_j$, $\gamma_{j2} = -\rho\beta_j$ (for $j = \{l, k, m\}$), and $\gamma_y = -\rho$. A direct consequence of the equations (10) and (11) is that the error term in the reduced-form equation is no longer a function of past productivity shocks ($\omega_{i,t-1}$) and is only a function of the unforeseen part of productivity $\zeta_{i,t}$. However, the fixed effect is still present, which is particularly problematic in dynamic panel models as input choice is still correlated with α_i. Arellano and Bond (1991) solves this problem by taking first differences of (12) which yields the following reduced-form equation in first-differences:

$$
\begin{aligned}
\Delta y_{i,t} &= \gamma_{l1}\Delta l_{i,t} + \gamma_{l2}\Delta l_{i,t-1} + \gamma_{k1}\Delta k_{i,t} + \gamma_{k2}\Delta k_{i,t-1} \\
&\quad + \gamma_{m1}\Delta m_{i,t} + \gamma_{m2}\Delta m_{i,t-1} + \gamma_{y}\Delta y_{i,t-1} + \tilde{\epsilon}_{i,t} \qquad (13) \\
\tilde{\epsilon}_{i,t} &= \zeta_{i,t} - \zeta_{i,t-1} + \epsilon_{i,t} - (1+\rho)\epsilon_{i,t-1} - \epsilon_{i,t-2} \qquad (14)
\end{aligned}
$$

where $\Delta x_{i,t} = x_{i,t} - x_{i,t-1}$. Equations (10) and (11) are less restrictive than assuming time-invariant TFP or than assuming that there is no persistence in productivity shocks ($\rho = 0$). However, they impose linearity. While other methods allow productivity to be non-parametrically related to past realizations (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015), those methods do not accommodate a fixed effect and impose stronger assumptions regarding input choice to relax linearity.[57]

Timing. Identification of the reduced-form parameters requires imposing assumptions regarding how households adjust inputs in order to accommodate unforeseen productivity shocks. The error term in equation (14) suggests that output and input usage observed before period $t-2$ are valid candidates for instruments. 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 AX shows that indeed lagged levels have predictive power.

B.3 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 attriters, it is important to evaluate the sensitivity of the productivity estimates to attrition.

[57] Other approaches achieve identification assuming that households could perfectly adjust investment or demand for intermediate inputs to accommodate productivity shocks (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015). However, Shenoy (2017a) argues that when households or firms face credit constraints, this assumption is likely to be violated leading to a failure of such approaches.

B.3.1 Timing Restrictions and Lag Selection

Note that the structure of the error term from equation (13) 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 18 present reduced-form estimates estimated using Arellano and Bond (1991)'s approach by sector. 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 farm-sector households and are not rejected in the case of 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 18: 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. 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.

B.3.2 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 to weak instruments and/or over-fitting. Thus a careful assessment between the tradeoff between precision and potential bias is important. Table 19 reports reduced-form estimates and structural estimates on panels A and B. Each column varies the lag length of the instruments. 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 does not 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 it 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 19: Estimates of Factor Elasticities Using Different Sets 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								
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.

B.3.3 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. That 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-sectional variation to identify the parameters.

Table 20 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).

I use the estimates from the system-GMM approach as the main specification in this paper. While they are similar to the difference-GMM estimates, they are estimated with higher precision. This choice comes at the cost of stronger identification assumptions than the ones required by the difference-GMM approach. Mainly, identification requires that first differences are not correlated with the initial levels of output and inputs.[58]

[58] 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 other words, businesses which start with a larger size should not systematically exhibit higher growth rates.

Table 20: Sensitivity of Difference-GMM Estimates to Including Moment Restrictions Associated with 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.

B.3.4 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 21 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 the productivity estimates would 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 21: 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.

B.4 Alternative Estimates

B.4.1 Correction for Potential Measurement Error in Capital

I correct for measuring error in capital by relying on lagged variation in investment (both levels and first-differences). This approach aims at correcting measurement errors that are related to corrections for depreciation or imperfect recall regarding the initial levels of capital. The intuition is that investment captures flows of resources that increase the stock of capital, but it is unrelated to measurements error in depreciation or measuring errors in the initial level of capital. The drawback of using this variation is that investment in developing countries tends to be lumpy and can lead to imprecise estimates.

To implement this approach I estimate (13) using suitable lags of investment instead of lagged values of capital. The required moment conditions are still similar to those from (7). The only difference is that the vector $\{k_{t-3}, ..., k_0, \Delta k_{t-3}, ..., \Delta k_1\}$ is not included in the set of instruments ($\mathbf{I}_{t-3}$). Instead, I include the vector of suitable investment lags $\{i_{t-3}, ..., i_0, \Delta i_{t-3}, ..., \Delta i_1\}$. Once the reduced-form equation (13) is estimated, I implement OMD to back out the structural parameters.

B.4.2 Estimates of a Restricted Model Using Three Pre-Program Years Only

The main estimation procedure in this paper exploits the moment conditions suggested by (7). Such approach is suitable for estimating models that allow for persistence in the productivity shocks (ω) and allow for a time-invariant component of ω. The existence of a fixed-effect requires first-differencing in order to purge out the time-invariant component, and thus we lose observations regarding the first period. As this process yielded a moment condition that suggested that 3rd lags were suitable instruments, it was impossible to estimate the production function only using the three pre-program periods.

To test the sensitivity of my estimates to using all years for the structural estimation, I estimate a restricted model that rules out a fixed effect but still allows productivity to evolve following a first-order autoregressive process. With this distinction it is possible to simply $\rho - differentiate$ equation (5) which yields the following set of moment conditions:

$$\mathbf{E}[\zeta_{i,t} + \epsilon_{i,t} - \rho\epsilon_{i,t-1} | I_{t-2}] = 0$$

Thus, information regarding input usage that is known at the end of period $t-2$ can be used to

identify the structural elasticities. Note that this means that only three periods are needed for the estimation of the simplified model as second lags would be suitable instruments under the simplified specification.

The estimation process is detailed below:

1. First I subtract ρy_{t-1} from both sides of equation (5):

$$\begin{aligned}(y_{i,t} - \beta_k k_{i,t} - \beta_l l_{i,t} - \beta_m m_{i,t}) =& \rho(y_{i,t-1} - \beta_k k_{i,t-1} - \beta_l l_{i,t-1} - \beta_m m_{i,t-1}) \\ &+ \zeta_{i,t} + \epsilon_{i,t} - \rho\epsilon_{i,t-1}\end{aligned}$$

2. I obtain candidate estimates for factor elasticities by estimating (5) through OLS.

3. Using candidate values for β_k, β_m, and β_l, I compute $\tilde{\omega_{i,t}}$:

$$\begin{aligned}\tilde{\omega}_{i,t} &= (y_{i,t} - \beta_k k_{i,t} - \beta_l l_{i,t} - \beta_m m_{i,t}) \\ \tilde{\omega}_{i,t} &= \omega_{i,t} + \beta_0\end{aligned}$$

4. As I allow ω to follow a first-order autoregressive process, for a given values of β_k, β_m, and β_l I estimate:

$$\tilde{\omega}_{it}(\beta_l, \beta_k, \beta_m) = \rho\tilde{\omega}_{it-1}(\beta_l, \beta_k, \beta_m) + \delta_{vt} + \tilde{\zeta}_i$$

   where δ_{vt} include a full set of village-year fixed effects.

5. The resulting residuals $\hat{\zeta}_i(\beta_l, \beta_k, \beta_m)$ are used to construct the sample analog of:

$$\mathbf{E}\left[(\omega_{i,t} - \rho\omega_{i,t-1}) \otimes \begin{pmatrix} l_{i,t-2} \\ k_{i,t-1} \\ k_{i,t-2} \\ m_{i,t-2} \end{pmatrix}\right] = 0$$

6. $\hat{\beta}_l$, $\hat{\beta}_m$ and $\hat{\beta}_k$ are estimated using GMM.

7. Note that the first lag of capital is a suitable instrument as capital is measured at the beginning of each period. Also, to estimate this model we only need three periods. Thus, I conduct the whole process using pre-program data only.

B.5 Online Appendix: Variable definition

B.6 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 cash or in-kind sales, the value of home-produced goods used as inputs for other activities,[59] and 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 construct annual revenues by summing all the revenues obtained between April and 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. See Chapter 4 in Samphantharak and Townsend (2010) for more details.

- 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 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

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

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.

- Farm revenues as a share of total revenues at baseline: Total cumulative revenues associated to cultivation, livestock, fishing and shrimping over the baseline periods divided by the total cumulative revenues from all activities.

- Farm sector household: A household whose baseline share of farm revenues is larger than 0.5. Off-farm sector household: A household whose baseline share of farm revenues is lower than 0.5.

B.7 Pre-program characteristics

Repayment history

- Share of loans with missed payments: Number of pre-program loans for which the borrower failed to make a payment when it was due, divided by the total number of pre-program loans obtained by each household. Note that this information is only available for households who ever borrowed from either formal or informal lenders.

- Share of loans with term extensions: Number of pre-program loans with an extension in its repayment period divided by the total number of loans obtained by each household. This information is only available for households with pre-program credit history.

Shocks to family operations:

- Operations: Number of self-reported issues or inconveniences regarding household operations. Three type of issues are considered by enumerators: issues collecting payments from final costumers, issues regarding the provision of inputs for production and issues regarding the production process itself–i.e., loss of production due to pests or extreme weather conditions. Enumerators collect this information for cultivation, livestock and off-farm family businesses.

- Health shocks: Count of the number of times household members were reported to have felt different health symptoms during each economic year. In the case of targeting analysis, I use the total over the pre-program survey waves.

Connections with the Village Council

- Village council membership: An indicator of whether a household member is either: the village chief or one of the members of the Village Council during the pre-program periods.

- Connected to the village council: An indicator of whether a household is either a council member, has any direct transaction links with the council member or is a first-degree relative of any council member. Transactions and kinship links are computed based on pre-program interactions.

- Number of links with the council: number of links to different council members in the baseline transaction network.

- Degree: Number of links a household has in the network. That is count of other households in the village that a particular households is connected to.

Borrower characteristics

- Leverage ratio: Total liabilities divided by total assets, averaged during the pre-program periods. Total liabilities include loans as well as short-term debts (accounts payable). Total assets include fixed assets, cash in hand as well as loans provided to other households.

- Wealth: Total assets net of total liabilities, averaged over the pre-program periods.

- Asset turnover ratio: Gross revenues from all operations divided by the average stock of assets in an economic year, averaged over the pre-program periods.

- Previously borrowed from lender: Indicator of whether a household has previously borrowed from a particular lender at the time in which a new loan is obtained. For each loan, I count the number of loans obtained by the borrower from the same lender during the periods preceding its disbursement.

- Number of outstanding loans: Count of outstanding loans, from any source, at the moment of obtaining a new loan.

- Demographic characteristics: average household age, average household years of schooling, household head's age, gender and years of schooling, number or working age adults in the households.

B.8 Loan characteristics

- Loan term: I group loans in four categories based on terms: 1-6 months, 6-12 months, more than 12 months and loans without a settled repayment period, which is the omitted category for the repayment model.

- Interest rate: self-reported interest rate (annual). In cases in which that information is missing, I calculate ex-ante interest rates as $(1 - \frac{ExpectedPayments}{Principal})/(Term(years))$. For each loan, enumerators ask the amount that the borrower is expected to pay to the lender at the end of the loan term. In the case of the repayment model, I group loans in 5 categories based on the initial interest rate: loans with zero interest rates, loans with positive interest rates that are less than 5%, 5-10%, 10-20% and more than 20%.

- Other loan characteristics: Indicators of whether the loan required a cosigner, collateral or the loan was a joint-liability loan.

B.9 Loan outcomes

- Ex-post internal rate of return: It is calculated as the interest rate such that the net present value of all the cash flows related to the loan equals zero. This statistic is computed for each loan that was: a) fully repaid, b) defaulted on after some payments and declared as not active. It excludes loans for which no payment was ever made, as the IRR is not defined in that case (0.1% of sample loans)

- Loans with delinquent payments: Indicator of any payment was missed for a given loan during its repayment period. A loan is coded as missed if the borrower reported not making a payment when a payment was due.

B.10 Delinquency risk

- Probability of a delinquent payment: I use loan, borrower and lender characteristics to predict repayment fitting a LASSO regression model (least absolute shrinkage and selection operator). The penalty parameter is selected through 10-fold cross validation. I then use the fitted model and borrower characteristics to predict repayment probabilities for all households in the survey sample, that is households with and without credit history.

- Eligibility under the repayment-score criterion: For each village I obtain the number of sample households who borrowed from the program during the first two years of program implementation k. For each village, I code the k households with the lowest delinquency risk as eligible.