Access to Credit and Productivity:

Evidence from Thai Villages*

Abhijit Banerjee, Emily Breza, Robert Townsend and Diego Vera-Cossio

PRELIMINARY VERSION-FOR DISCUSSION ONLY

September 5, 2019

Abstract

Approaches to underdevelopment based on misallocation of resources have two premises. First, that there is huge heterogeneity in terms of underlying productivity among potential and actual entrepreneurs. Second, that the mechanisms that guide resource allocation do not necessarily result in the resources going to the most productive entrepreneurs. Using a long panel of small businesses, and quasi-experimental variation from the rollout of a large-scale credit expansion program in Thailand, we show evidence for both these premises. First, exploiting a long time series of pre-intervention information, we estimate productivity household by household. We then show that the effect of the program, which was a source of additional short-term credit in the village, varies dramatically by pre-program productivity. There is no discernible effect in terms of income or business profits among low pre-program productivity households, but the higher-productivity households show a large increase in profits (more than 1.5 THB increase in profits for 1 THB in loans). This effect doubles when we restrict to high-productivity households that had a non-agricultural business before the intervention. On the other hand, program credit is not allocated based on baseline productivity. However, market credit partly mitigates the disparity.

Keywords: Credit, misallocation, productivity.

JEL: O16, G21, D21.

*Abhijit Banerjee: Department of Economics, MIT. Emily Breza: Department of Economics, Harvard University. Robert Townsend: Department of Economics, MIT. **Corresponing author:** Diego Vera-Cossio: Research Department, Inter-American Development Bank (IADB).Email: diegove@iadb.org. This paper reflects the authors' own opinions and does not reflect IADB's official positions. All errors are our own. E-mail: diegove@iadb.org.

1 Introduction

A large literature argues that factor misallocation can explain cross-country differences in output and income, and further, improving the allocation of resources within-country has the potential to unlock economic growth (see, for example, Banerjee and Duflo (2005); Restuccia and Rogerson (2008); Alfaro et al. (2008); Hsieh and Klenow (2009); Bartelsman et al. (2013); Foster et al. (2006)). This argument relies on two pillars. First, there must be substantial heterogeneity in productivity across firms and entrepreneurs. Second, market frictions must impede inputs from flowing to the most productive firms. As a result, marginal products are not equalized across firms, and a move towards equalizing marginal products could lead to large gains in output.

Obtaining empirical support for these premises is challenging, as it entails tackling issues of measurement error in different margins (Restuccia and Rogerson, 2017; Gollin and Udry, 2019; Atkin et al., 2019; Foster et al., 2016). One popular approach is to analyze the cross-sectional relationship between firm-level marginal products and productivity as, in the absence of misallocation, marginal products should be equalized across firms regardless of productivity.[1] In practice, this approach faces important limitations. First, researchers observe neither productivity nor marginal products. Without independent sources of variation to estimate both objects, researchers use possibly noisy estimates of productivity to back out marginal products possibly leading to spurious correlations and overstating the degree of misallocation. Second, it is unclear if cross-sectional relations are stable over time or driven by transitory productivity shocks. Third, without variation in frictions in labor or capital markets, the sources of misallocation are unclear, and hence the ability to draw precise policy implications is limited.

In this paper, we overcome these issues by exploiting 10 years of panel data combined with quasi-experimental variation in the supply of credit to analyze misallocation in the context of small firms in Thailand. In particular, we ask whether cross-sectional heterogeneity in productivity predicts either the allocation of credit to entrepreneurs or explains dispersion in the returns to credit. We begin by developing a simple model of constrained households to show that, when credit constraints are binding, the shadow price of capital–i.e., the marginal returns of relaxing credit constraints– is an increasing function of household productivity. We then empirically test this prediction by first exploiting two different sources of variation to estimate productivity and the returns to credit, and

[1]A similar approach is used in settings that fit models of monopolistic competition. In such cases researchers test whether physical productivity–i.e., TFPQ as in Hsieh and Klenow (2009)– is correlated with revenue productivity (TFPR). In these cases, researchers first estimate TFPR and use it to back out TFPQ.

then assessing whether productivity drives heterogeneity in the returns to credit.

Our analysis requires three crucial components. First, it requires a credible measure of productivity. One well-known problem when estimating production functions is that investment decisions may be endogenous to unobserved time-varying shocks to productivity (Olley and Pakes, 1996). We propose a novel method in which we use data on entrepreneur beliefs about future business conditions to proxy for firm-specific productivity shocks. Our beliefs-based method is particularly attractive for settings like ours where households are likely credit constrained.[2] Second, we need exogenous variation to estimate the marginal returns to credit. For this, we use quasi-experimental variation in the exposure to the Million Baht Program, one of the largest credit-expansion programs of its kind.[3] We follow Kaboski and Townsend (2012), who exploit the fact that each program village received the same amount of funds from the central government to lend to local households, independent of village size. Thus we can compare villages before versus after the implementation of the program, by per-capita program resources. Finally, we require detailed panel data with enough pre-intervention observations to estimate productivity, and enough post-interventions to capture stable relationships between productivity and returns. Here, we use the Townsend Thai Project panel (Townsend, 2007b,a), which follows 960 households from 64 villages during five pre- and post-intervention years, and includes information on assets, inputs, revenues and profits for all household businesses.

Armed with these three tools, we first use the pre-program data to recover estimates of firm-level productivity for all potential borrowers. We then compute average pre-program productivity for each entrepreneur and combine our estimates with the cross-village variation in village size and the program's rollout to test for productivity-based heterogeneity in the effects of the credit expansion. As the rollout of the program was orthogonal to measured productivity, our empirical strategy prevents measurement error in productivity from driving dispersion in the returns to increased access to credit.

Consistent with Kaboski and Townsend (2012), we find that indeed, smaller villages experience a large increase in short-term credit following the implementation of the program, relative to the pre-

[2]Shenoy (2017a) argues that the assumptions typically made by control function approaches to estimating production functions (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015) are likely unsuitable when there are binding credit constraints. These methods traditionally use investment or intermediate inputs to proxy for unobserved changes in productivity, but households may not be able to adjust intermediate inputs in response to productivity shocks when they are constrained.

[3]Starting 2001, the Thai government disbursed THB one million to each of the 77,000 participating villages (approximately USD 24,000 at 2001 exchange rates). The total program resources account for approximately USD 1.8 billion and reached over 95% of the total villages in Thailand.

period. We also find that the allocation of program credit is not detectably different for high- versus low-productivity households. In other words, for this community-driven credit product, credit was not disproportionately directed toward the more productive households. One implication is that potential heterogeneity in downstream outcomes is unlikely to be explained by differential access to program credit.

While credit does not flow disproportionately to higher-productivity households, we nevertheless find strong patterns of heterogeneity by baseline productivity on business outcomes. First, we find no detectable impacts of the program on household income or business profits for low productivity households. However, the picture is quite different for high productivity households, which experience increases in total household income coming largely from household enterprise profits. These strong differences persist during the five post-intervention years; thus, they are unlikely to be driven by transitory productivity shocks. As in Banerjee and Duflo (2014), this is evidence that high-productivity households were indeed credit constrained before the program. Moreover, the increase in profitability comes almost entirely from non-agricultural businesses rather than farm-related activities. One interpretation is that in the Thai context, credit constraints aren't as binding for agricultural businesses, perhaps due to differences in collateralizability of farm versus non-farm assets or due to pre-existing targeted agricultural lending programs.[4]

Next, we show that productivity-based heterogeneity in the effects of the program is even larger when we analyze the subsample of households with preexisting non-agricultural businesses. Thus, our results are unlikely to be driven by entrance of new businesses. We also find evidence that for high-productivity households, program credit crowds in other types of borrowing. While high and low productivity households obtain similar amounts of village fund credit, total short-term borrowing increases more for high-productivity households relative to low-productivity households. Consequently, we find that among owners of preexisting non-agricultural businesses, high-productivity households are better able to use the village credit to increase profits. This increase in profitability appears to be driven by an immediate increase in assets, rather than increased inventories and wage expenses.

We show that these results are robust in two ways. First, we use an alternate fixed effects-based approach to estimate pre-program productivity and show that the results are qualitatively quite similar. Second, one limitation of our dataset is that it does not track total labor inputs for

[4] Agriculture-oriented lenders are prominent in the context of rural Thailand. For instance, before the program's implementation, the Bank for Agriculture and Agricultural Cooperatives (BAAC) provided agricultural loans in all the sample villages.

family businesses. We thus present several approaches to impute total labor inputs to use in our estimation of household productivity.[5] We show that our main findings are robust to estimating a production function in per-capita terms, and to the inclusion of the number of family and paid workers as a measure of labor in the production function estimation. Additionally, we show that the results are robust to using investment data to correct for potential measurement error in capital (Collard-Wexler and De Loecker, 2016).

In order to interpret the magnitude of these effects, we use the rollout of the program to instrument for total short-term credit, and compute the effects of an additional THB of total short-term credit on household profits assuming that the program rollout only affected profits through the receipt of credit.[6] With this caveat in mind, we find that profits increased by THB 1.47 per additional THB of total short-term credit in the case of high-productivity entrepreneurs. In the case of incumbent non-agricultural businesses, the top-third most productive entrepreneurs exhibit annual returns to credit of the order of THB 2.9 per one additional THB of credit.[7] Such returns are above and beyond the average interest rates of program loans (7%), and of loans from other formal financial institutions (9%). Our estimates are consistent with evidence of large annual returns to cash/asset grants in Mexico (McKenzie and Woodruff, 2008) and high returns to cash grants for entrepreneurs with high-growth potential in India (Hussam et al., 2017).

Overall, our results suggest misallocation in credit markets. High-productivity entrepreneurs exhibited high returns to credit, and yet they did not obtain more credit. Thus, there may be gains from reallocating program credit. We use our quasi-experimental estimates to quantify the gains from reallocating program credit from lower- to higher-productivity households, while holding the total amount of credit in the village constant. We find that reallocating resources towards the most productive households would yield marginal returns to credit that are three times as large as the ones associated to the actual allocation, and that are 20% less disperse. In turn, simply reallocating program resources could magnify the average village-level output by 2.9 to 10.3%. Thus, simply improving the program's screening process may lead to substantial aggregate output gains.

This paper makes a number of contributions. First, it contributes to the literature studying the role of financial frictions on business performance in developing countries. While model-based

[5]Unfortunately, the Townsend Thai Project annual data does not track total labor inputs in household businesses (i.e., time use by business activity). It only contains measures of the number of workers hired for non-agricultural businesses and the number of households members whose main occupation is to work in household businesses.

[6]To minimize potential violations to this assumption associated to general equilibrium effects, we focus on estimates covering only the first and second years after the program rollout.

[7]These effects correspond to specification using winsorized data. Using raw data we obtain even higher estimates of the order of THB 2.87 per additional THB of total short-term credit.

assessments emphasize the importance of heterogeneity in productivity across entrepreneurs (Buera et al., 2012, 2015, 2017) and the consequences of financial frictions in terms of misallocation (Midrigan and Xu, 2014), experimental and quasi-experimental studies of the effects of credit expansion programs do not seem to include such distinction.[8] For instance, previous studies documented non-transformative average effects on business profits (Banerjee et al., 2015; Meager, 2019), and that there is heterogeneity based on business size (Crepon et al., 2015) and pre-program business ownership (Banerjee et al., 2015). Our results provide novel evidence showing that the most-productive households indeed exhibit high returns to credit, and that productivity predicts larger returns even within key subpopulations. However, we find no detectable correlation between program borrowing and household TFP, and our results are consistent with the ex ante credit constraints binding more for high productivity entrepreneurs. One implication is that improved screening and targeting could magnify the impacts of credit expansions. This could potentially entail improvements in externally identifying entrepreneurs (see Fafchamps and Woodruff (2017), Hussam et al. (2017), and Mckenzie and Sansone (2017)). Alternatively, financial institutions could try to design better screening mechanisms for self-targeting (Beaman et al., 2014).

Second, we also contribute to the empirical literature quantifying factor misallocation in developing countries (Restuccia and Rogerson, 2017; Hsieh and Klenow, 2009). Our results document evidence of misallocation in credit markets by providing a novel methodological approach. As opposed to previous empirical studies based on cross-sectional comparisons using one or few rounds of data,[9] our empirical strategy leverages on 10 years of panel data and two independent sources of variation to reduce the influence of measurement-error in productivity on the dispersion of marginal returns. We also contribute by documenting gains of 2.9-10% in village-level output from simply reallocating program credit. Although substantial, these gains are modest relative to gains from fully reallocating factors across farms (240% in the case of farmers in Malawi (Restuccia and Rogerson, 2017)) and firms in developing countries (100-115% in firms in India and China (Hsieh and Klenow, 2009)). However, achieving such large gains through policy requires institutional and political assumptions that are unlikely to hold in developing countries. Our less-ambitious empirical exercise suggests that there could be substantial gains from simply improving the screening process and targeting methods in already existing credit-market interventions.

[8] There is large body of research on the effects of micro-credit expansion programs in several settings (see Banerjee et al. (2015) for a review of recent studies) or government-funded village fund programs (Kaboski and Townsend (2012) in Thailand and Cai et al. (2017) in China).

[9] Examples of these include Restuccia and Rogerson (2017); Alfaro et al. (2008); Hsieh and Klenow (2009); Restuccia and Rogerson (2008), and Foster et al. (2006), among others

Finally, in the context of the Million Baht Program, previous studies showed that the program led to increases in consumption without changes in business outcomes (Kaboski and Townsend, 2012), to rather small reductions in capital misallocation among rice farmers (Shenoy, 2017b), and that a cash-transfer scheme would have been more cost-effective (Kaboski and Townsend, 2011). The evidence in this paper suggests that financial frictions preventing program credit from reaching high-productivity entrepreneurs may explain the lack of transformative results. Indeed, as the program was fully managed by local committees, the evidence on this paper is consistent with evidence showing that the program resources were disproportionately delivered to households with connections to the village chief (Vera-Cossio, 2018).

The body of the paper proceeds as follows. Section 2 details the empirical context and the data. Section 3 presents a simple framework of credit supply expansions under credit constraints and also outlines our production function estimation methodology. Section 5 documents the core first stage and reduced form results, while Section 6 provides IV estimates of the returns to credit and the returns to fixed capital. Section 7 performs a counterfactual analysis to quantify the potential gains from reallocation. Finally, Section 8 concludes.

2 Context and Data

We study the heterogeneous impacts of the Million Baht Program on household income and profits in the 64 villages of the Townsend Thai Project (Townsend, 2007b,a). Under the Million Baht program, the Thai government disbursed approximately USD 1.8 billion to 77,000 villages starting in 2001.[10] Our empirical strategy is based on the work of Kaboski and Townsend (2012), hereinafter KT, and exploits the unique implementation of the program to facilitate identification of its causal effects. Notably, the government disbursed exactly THB 1,000,000 to each village regardless of size, wealth or location (approximately USD 24,000 at 2001 exchange rates).[11] As such, inhabitants of small villages stood to receive more credit, on average, than residents of larger villages. In general, most of the credit was lent on a short-term (less than or equal to 12 months) basis, and because any funds repaid to the village fund committees were meant to be used to finance follow-on lending activities, the program could be viewed as a permanent supply shock to local short-term credit.[12]

[10]See Kaboski and Townsend (2012) for a detailed description of the program.

[11]Subject to each village successfully forming a village fund committee, the body which would ultimately manage the funds and make credit decisions along with loan collections.

[12]However, by 2004 several village fund committees had gone bankrupt due to mismanagement or default outbreaks, spurred by powerful members of the village.

Kaboski and Townsend (2012) provide quasi-experimental evidence of the effects of the program on household consumption and productive activities. Concretely, they document that short-term borrowing increased , crowding in credit from other lenders in the village, and leading to sizeable effects on consumption (increases of 1.7 THB per THB injected by the program). While most models of credit frictions would predict an increase in business investment and profits, the average effects of the program on productive activities are rather small.[13] In this paper, we build on this previous work by tackling the question of misallocation and asking whether the absence of any average effects on businesses are evidence for misallocation.

We focus on the Thai context for three reasons: First, the Thai Million Baht program provides quasi-experimental variation in the timing and size of the program to identify the effects of the program on household outcomes. Second, cross-village variation in the size of the program allows us to capture enough heterogeneity in household productive characteristics among program borrowers; in small villages which receive large per-capita program funds, both high and low productivity households may borrow. Third, the implementation of the program overlaps with the availability of a long-panel dataset, the Townsend Thai Project, which records extremely detailed household records for 960 households from 64 villages in 4 Thai provinces. The nature of the data is unique in its comprehensiveness and panel length, which allows us to exploit the detailed, repeated nature of the household observations to implement modern panel-data methods to characterize households in terms of pre-program productivity and other productive characteristics.

Table 1 presents summary statistics for the study sample. Two important characteristics are worth emphasizing. First, household economic performance involves a variety of economic activities. While on average, higher shares of household operating income correspond to farming and wage work outside the household, 35% of households have an off-farm business. Second, even before the program, access to credit was common. Over two-thirds of households borrowed either from institutional or informal lenders. Moreover, 50% of households report having an outstanding loan with institutional lenders such as the state-owned Bank of Agriculture and Agricultural Cooperatives (BAAC), commercial banks and other local cooperatives or village organizations.

[13] The absence of large, detectable average effects on profits and incomes is consistent with the broader microfinance literature (Banerjee et al., 2015).

3 A simple theoretical framework

In this section, we propose a simple theoretical framework to characterize the households who are best able to convert increased credit supply into business profits. We argue that in the presence of credit constraints, cross-household variation in the marginal return to capital - i.e., the shadow price from relaxing the budget constraint - is mainly driven by variation in total factor productivity (TFP). In order to illustrate this point, we start by analyzing a simple static profit-maximizing problem of a household or business facing a liquidity constraint.

Households are different in terms of total factor productivity ($TFP = A_i$), and combine K and labor L to produce output Y. Consider a Cobb-Douglas production function ($Y_i = A_i K_i^{\alpha_K} L_i^{\alpha_L}$),[14] then each household maximizes profits subject to a budget constraint:

$$\max_{K_i, L_i} A_i K_i^{\alpha_K} L_i^{\alpha_L} - p_K K_i - p_L L_i \tag{1}$$

subject to

$$p_K K_i + p_L L_i \leq B_i \tag{2}$$

Where B_i denotes the total budget available to household i, and includes both wealth and credit. We allow heterogeneity in this dimension to capture differences in wealth as well as access to credit across households. Input prices (p_K, p_L) are normalized with respect to the price of output. Let λ_i denote the LaGrange multiplier associated with the budget constraint (2). Thus, λ_i represents the shadow value of a marginal increase in household i's budget (B_i): the marginal return to capital. Thus, if credit expansion programs effectively modify the availability of resources B_i, then heterogeneity in λ_i captures heterogeneity in the ability of a household to benefit from increases in the supply of credit.

Combining the first order conditions corresponding to the choice of each input, it is possible to show that an optimal solution implies:

$$A_i B_i^{\alpha_K + \alpha_L - 1} \kappa = 1 + \lambda_i \tag{3}$$

As κ is strictly positive, λ_i is an increasing function of total factor productivity (A_i).[15] Moreover,

[14] The theoretical predictions highlighted in this sections do not depend on the number of inputs and hold for concave production functions.

[15] $\kappa = \left(\frac{1}{\alpha_K + \alpha_L}\right)^{\alpha_K + \alpha_L - 1} \left(\frac{\alpha_K}{p_K}\right)^{\alpha_K} \left(\frac{\alpha_L}{p_L}\right)^{\alpha_L}$

with decreasing returns to scale ($\alpha_K + \alpha_L < 1$), λ_i is decreasing in B_i. In words, households benefit more from relaxing the budget constraint if productivity is high and if wealth or credit availability are low. In the context of a technology with constant returns to scale, the budget constraint is irrelevant, and only heterogeneity in TFP drives heterogeneity in the shadow value of capital.

4 Empirical strategy

Kaboski and Townsend (2012) exploit variation in the timing and size of the program to estimate its causal effects on productive outcomes. In particular, they compare changes in outcomes before and after 2001 corresponding to villages with high per-capita expected credit supply (or high inverse village size, $invHH$) to those with low per-capita expected credit supply (or low inverse village size). This approach would lead to the causal identification of the effects of the program under the assumption that there were not time-varying shocks that differently affected small and large villages, and could potentially be related to outcomes. The authors argue that the spatial distribution of village size is as if random and validate the identification assumptions with numerous robustness checks. We build on their empirical approach by analyzing the heterogeneous effects of the program motivated by our theoretical framework.

Our aim is to understand if households with higher λ_i do in fact benefit more from the increase in the supply of credit induced by the Million Baht Village Fund program. Let $\lambda_{i,n}$ be the household's shadow value of capital for household i in village n, corresponding to the baseline periods. While we do not observe $\lambda_{i,n}$, our theoretical framework suggests that baseline productivity $A_{i,n,t}$ captures important variation in the returns to capital. Thus, in our empirical analysis we aim to estimate the heterogeneous reduced-form effects of the program following:

$$\begin{aligned} y_{i,n,t} =& \delta_1 invHH_n \times Post_t + \delta_2 invHH_n \times Post_t \times \text{High A}_{i,n} + X_{i,n,t}\Gamma \\ &+ \delta_3 \text{High A}_{i,n} + \theta_t \times \text{High A}_{i,n} + \theta_t + \theta_n + e_{n,t} \end{aligned} \tag{4}$$

Here, n indexes the village, t indexes the year, and i indexes households. High $\text{A}_{i,n}$ is an indicator that identifies households in the top-third of the TFP distribution, within each village. We mainly focus on rankings rather than levels to attenuate potential measurement error as we estimate $A_{i,n}$ (see Section 4.1). $Post_t$ is an indicator that identifies post-program years (2002-2006). We allow for A-specific time trends and include a $1 \times I$ vector of covariates $X_{i,n,t}$ (including household composition, age, and education), village (θ_n) and year fixed effects (θ_t). The coefficients of interest are δ_1, $\delta_1 + \delta_2$, and δ_2; they represent the reduced-form effects of the program for households in the bottom-two thirds of the productivity distribution ($HighA = 0$), high-productivity households, and treatment effect heterogeneity between high and lower productivity households, respectively.

We are also interested in assessing the dynamics of the effects of the program. In section 5, we report estimates corresponding to the following flexible differences-in-differences specification in equation (5), which is separately estimated for high and low productivity households.

$$y_{i,n,t} = \sum_{\tau=1997,\tau\neq 2001}^{\tau=2006} \delta_\tau invHH_n \times \mathbf{I}[t=\tau] + X_{i,n,t}\Gamma + \theta_t + \theta_n + e_{n,t} \tag{5}$$

In this case, the parameters of interest are δ_τ. They denote differences in the outcome variable between villages with high and low per capita program resources in period τ relative to the same differences in 2001, the year preceding the full implementation of the program. This exercise is useful to graphically examine potential violations to the parallel trends assumption which is necessary for causal interpretation in differences-in-differences designs.

4.1 Production function estimation

Our analysis involves the measurement of baseline productivity $A_{i,n}$ for each potential borrower, which typically requires the estimation of a production function. We model log value added ($va_{i,t}$), aggregated across all household enterprises,[16] as a function of the stock of fixed capital $k_{i,t}$,[17] productivity shocks which are observed by the household but not by the researcher $\omega_{it} = log(A_{it})$, and unexpected shocks to production ($\epsilon_{i,t}$) which are neither known by the household nor by the researcher.[18]

$$va_{i,t} = \beta_0 + \beta_k k_{i,t} + \omega_{it} + \epsilon_{i,t} \tag{6}$$

We are interested in estimating ω_{it} for each household, which represents variation in value-added conditional on capital.[19] That is, we aim to capture differences across households in their ability to

[16]Enterprise activities include cultivation, livestock, production of livestock produce and off-farm family business. Value added is measured as total revenues net of the cost from input usage, other than capital and labor. For instance, we subtract the value of fertilizer, seeds, feed, merchandise and fuel (among others) from total gross household revenues.

[17]The stock of capital is measured as the stock of fixed assets corresponding to farm and non-farm businesses.

[18]We restrict the analysis to a value-added function for ease of exposition of our method. This is an advantage with regard to home-produced goods, which may serve as inputs for the production of other goods. A value added approach prevents double counting.

[19]We use a value-added function over a gross revenue function as households may have different sources of income

generate value added, holding constant their capital endowments. We note that ideally, Equation (6) would also include labor inputs on the right hand side in addition to capital. However, unfortunately, we do not have detailed data regarding labor hours. In our main analysis, we estimate ω_{it} using Equation (6), considering only capital inputs. However, we present robustness checks based on estimates that use the number of workers hired for off-farm businesses and the number of adults in the households as a proxy for labor inputs (see Section 5.3).

We allow productivity to evolve following two sources of variation: foreseen variation based on previous realizations (e.g., $\omega_{i,t-1}$) and unforeseen shocks to productivity $\zeta_{i,t}$. The empirical challenge is to consistently estimate β_k, which is essential to back out ω_{it}. In order to do so, we need to tackle two potential problems. First, households may adjust capital to respond to unforeseen shocks to production $\epsilon_{i,t}$ –i.e., spoilage– and shocks to productivity $\zeta_{i,t}$ such as unexpected favorable business opportunities. Second, households may optimally decide their investment decisions in order to accommodate foreseen variation in productivity ω_{it} (Olley and Pakes, 1996). Both sources of endogeneity may lead to biased OLS estimates of β_k. Ideally, we would rely on household-level experimental variation in the stock of fixed capital to compute β_k. While such a source of variation is not available in our context, the richness and length of our panel dataset allow us to go a long way in reducing these concerns.

In order to tackle the first problem, we define the stock of capital available at the beginning of period t as the stock of capital reported in the survey wave $t-1$. By doing so, we focus on a predetermined measure of capital such that $\mathbf{E}[k_{i,t}\epsilon_{i,t}] = 0$ and $\mathbf{E}[k_{i,t}\zeta_{i,t}] = 0$. This approach is consistent with models in which there is time to build related to productive capital (Kydland and Prescott, 1982) and with evidence of lumpy investments in Thai villages (Samphantharak and Townsend, 2010). Tackling the second problem requires controlling for unobserved variation in $\omega_{i,t}$ which is correlated with capital choices. We propose two approaches that rely on different identification assumptions to overcome this issue.

4.1.1 Fixed-effects approach

In the fixed effects approach, we assume that variation in productivity is explained by a time-invariant component which is correlated with capital decisions, year-specific aggregate shocks, and a time-variant unforeseen shock which is experienced after households choose capital –i.e., $\omega_{it} =$

and use output from one occupation as inputs for another. For instance, a farmer may produce some crops for sale but may use part of the harvest for feed for its livestock. Without a systematic accounting process, a gross revenue approach could lead to double accounting.

$\bar{\omega}_i + \omega_t + \zeta_{it}$, with $\mathbf{E}[\bar{\omega}_i k_{it}] \neq 0$ and $\mathbf{E}[\zeta_{it} k_{it}] = 0$. This specification allows us to estimate (6) through a fixed-effects approach using the 5 years preceding the program (1997-2001)[20] and use within-village rankings of the estimated $\hat{\bar{\omega}}_i$ to estimate equations (4) and (5).

While simple, this approach has two limitations. First, by not allowing the foreseen part of productivity to evolve over time, the fixed-effects approach rules out models in which households may accumulate knowledge or develop abilities which may allow them to more efficiently use capital in future periods. If the latter models are the main drivers of households behavior, then the fixed-effects approach may fail to fully account for the relation between capital and productivity. Second, even if a fixed-effects model is a good description of the true data-generating process, the identification of β_k will rely on within household-variation in capital, which may be troublesome in contexts in which investment is lumpy and there is measurement error in capital. In such cases, fixed-effects estimates of productivity may end up absorbing most of the variation in the stock of capital.

4.1.2 Control function approach

A less restrictive approach for estimating β_k relies on the use of proxy variables in order to control for variation in productivity (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015). By doing so, this approach allows productivity ω_{it} to vary over time and across households. Typically, the control-function approach uses variation in the demand for intermediate inputs to proxy for variation in productivity, which involves assuming that there is a strict monotonic relation between the demand of intermediate inputs and productivity. Thus, as long as firms can afford to modify intermediate inputs to accommodate productivity shocks, the control-function approach would yield consistent estimates of β_k. While appealing, the approach is not well-suited to settings of limited access to credit: liquidity-constrained households/firms may not be able to freely adjust intermediate inputs in order to accommodate productivity shocks, and thus variation in intermediate inputs may not fully capture variation in productivity (Shenoy, 2017a).[21]

In this paper, we propose a simple modification to the control-function approach that overcomes some of the problems highlighted by Shenoy (2017a) while still taking advantage of the benefits of

[20] Concretely, we estimate the following specification through OLS: $va_{it} = \bar{\omega}_i + \beta_k k_{it} + \delta_t + u_{it}$. $\bar{\omega}_i$ represents household-specific indicators and $\delta_t$represents year fixed effects. We then use the OLS coefficients associated to the household-specific indicators as estimates of productivity: $\hat{\bar{\omega}}_i$.

[21] Shenoy (2017a) proposes the use of dynamic panel methods that would be based on weaker assumptions regarding optimal firm behavior. However, relaxing such assumptions as in Blundell and Bond (2000), comes at the cost of imposing functional forms to the productivity process (typically, assuming that productivity follows an AR(1) process). Moreover, the implementation of such models requires long time series in order to avoid problems with precision.

the control function approach. Namely, we use household beliefs about future business conditions to proxy for variation in foreseen productivity. While credit constraints may prevent households from adjusting intermediate inputs to accommodate productivity shocks, they are less likely to prevent household from adjusting their beliefs. We view this insight as a contribution to the literature in cases where credit constraints are likely to bind.

While households observe productivity and we don't, the Townsend Thai survey includes questions about household forecasts of future profits. We postulate that household's beliefs about business conditions in period t ($b_{i,t}$) are a function of capital (observable to the researcher) and productivity (unobservable to the researcher): $b_{i,t} = b(k_{i,t}, \omega_{i,t})$. Thus, our ability to effectively use variation in $b_{i,t}$ to proxy for variation in $\omega_{i,t}$ relies on the idea that if we observed different beliefs across households with similar stocks of capital, it should be the case that households with more positive beliefs are also households with higher productivity. If households fully incorporate variation in productivity into their beliefs in a frictionless way, then beliefs are a strict monotonic function of productivity. Under this assumption, it is thus possible to invert the relation between beliefs and productivity and write down ω as a function of household beliefs and capital ($\omega_{i,t} = b^{-1}(k_{i,t}, b_{i,t})$).

Under these assumptions, our estimation procedure is similar to the two-stage approach proposed by Levinsohn and Petrin (2003) and can be easily extended to Ackerberg et al. (2015)'s approach to allow frictions in other inputs. We focus on the former for simplicity. First, we use third-order polynomials of $k_{i,t}$ and $b_{i,t}$ to semi-parametrically recover variation in value added that is explained by capital and household beliefs:

$$\hat{va}_{i,t} = \sum_{j=0}^{3}\sum_{l=0}^{3} \hat{\delta}_{jl} k_{i,t}^{j} b_{i,t}^{l} \tag{7}$$

Second, for a given initial value of β_k, we can recover estimates of productivity shocks:

$$\hat{\omega}_{i,t}(\beta_k) = \hat{va}_{i,t} - \beta_k k_{i,t} \tag{8}$$

Next, we allow non-parametric persistence in productivity by assuming ω follows a first-order Markov process ($\omega_{it} = \mathbf{E}[\omega_{i,t}|\omega_{i,t-1}] + \zeta_{i,t}$), and estimate $\mathbf{E}[\hat{\omega}_{i,t}|\hat{\omega}_{i,t-1}]$ by regressing $\hat{\omega}_{i,t}(\beta_k)$ on a third-order polynomial of the previous realization of the shock ($\hat{\omega}_{i,t-1}(\beta_k)$). Finally, β_k^* is is chosen to minimize the sum of squared residuals:

$$\min_{\beta_k^*} \sum_t \sum_i \left(va_{it} - \beta_k^* k_{it} - E\left[\hat{\omega}_{i,t}(\beta_k)|\hat{\omega}_{i,t-1}(\beta_k)\right]\right)^2 \tag{9}$$

We implement this procedure using the pre-program sample only. A more formal discussion of the identification assumptions and the estimation process are detailed in Appendix Section (C.1). At the end of the procedure, we average the estimates $\hat{\omega}_{i,t}(\beta_k^*)$ over the pre-intervention periods and generate within-village rankings of household productivity. We then use these rankings to analyze heterogeneity in the effects of the program.

Note that our approach relies on the same moment conditions corresponding to Levinsohn and Petrin (2003)– i.e.,$\mathbf{E}[\hat{\zeta}_{i,t}|k_{i,t}] = 0$,[22] but it uses a different source of variation to compute the sample analog of such conditions. Based on the idea that agents smoothly respond to productivity shocks by modifying the demand of intermediate inputs, the traditional control-function approach uses variation in intermediate inputs and capital to recover foreseen productivity $(E\left[\hat{\omega}_{it}|\hat{\omega}_{it-1}\right])$.[23]

In contrast, our approach uses variation in household beliefs and capital to proxy for the foreseen part of productivity and makes no assumption regarding the existence or not of liquidity constraints.

Our approach is not free of assumptions. First, it requires that household beliefs capture meaningful variation in value added, conditional on capital. Appendix Table A1 reports within-village correlations between household value added and income forecasts, with and without including the stock of capital as a predictor. Reassuringly, household forecasts are significant predictors of value added. Second, our approach assumes that there is a strict monotonic relation between household beliefs and productivity. This requires that households adjust their beliefs in the same direction of foreseen productivity shocks.[24]

We implement our empirical strategy using household projections of profits at time t, which were measured at the end of period $t-1$.[25] In order to account for differences across households

[22]Because we assume that capital is predetermined with respect to production shocks $(\epsilon_{i,t})$ and to unforeseen innovations in productivity $\zeta_{i,t}$, identification is achieved under the following moment condition $\mathbf{E}[\zeta_{i,t}|k_{i,t}] = 0$ in which $\zeta_{i,t} = va_{i,t} - \mathbf{E}[\omega_{i,t}|\omega_{i,t-1}] - \beta_k k_{i,t}$

[23]Traditional proxy variables are materials or electricity. The control function approach observes, that demand for intermediate goods, m_{it}, can be expressed as a function of the current capital stock and productivity, $m_{it} = m_{it}(k_{it}, \omega_{it})$. Under some assumptions, mainly a strictly monotonic relation between m and ω, the demand function can be inverted yielding $\omega_{it} = m^{-1}(K_{i,t}, m_{i,t})$

[24]One clear limitation of this assumption is that it rules out models of cognitive rigidities in the formation in beliefs (Handel and Schwartzstein, 2018).

[25]Note that because we measure beliefs about $t+1$ at the end of period t, we assume that such beliefs fully capture the part of productivity in $t+1$ that is correlated with input use. If this assumption fails then beliefs only capture the foreseen part of productivity $(E[\omega_{i,t+1}|\omega_{i,t}])$ instead of the actual realization $(\omega_{i,t+1})$. However, Appendix Section C.1 shows that identification is not compromised when capital is predetermined, which is the case here.

in the way in which they form beliefs as well as scale and volatility, we follow Ahlin and Townsend (2007) and use household forecasts under different scenarios to recover subjective beliefs about a successful business. Concretely, we exploit the fact that the survey collects information regarding income projections in *a)* a regular scenario, *b)* an adverse scenario and *c)* a good scenario. We normalize household beliefs by dividing the difference in projected income between a regular and a bad scenario by the difference in projected income between the good and bad scenario.[26]

Table 2 reports estimates of β_k under different methods and provides summary statistics of $\hat{\omega}$, averaged across the 5 pre-program periods, which is our main measurement of household productivity A_i. While the fixed-effects approach achieves low estimates of β_k and larger estimates of productivity than the control-function approach, the implied within-village productivity rankings are similar across both methods: Appendix Table A2 reports correlations between percentile rankings of control-function estimates of productivity (dependent variables) and percentile rankings of fixed-effects estimates of productivity with and without fixed effects (regressor). It shows that both productivity measures are highly correlated. Importantly, Appendix Table A3 shows that our productivity estimates are correlated with household characteristics that are usually associated with higher productivity, such as education. This pattern holds for both estimates of productivity: (the fixed-effects and proxy-variable approach) and suggest that our productivity estimates capture meaningful economic attributes. Throughout the rest of the paper, we present evidence based on estimates from both approaches and rely on results that are robust across both measurement strategies.

5 Reduced-form results

5.1 First Stage: Effects on program and total short-term credit

We begin by asking whether baseline productivity captures heterogeneity in program borrowing and total short-term borrowing. For the sake of consistency with previous studies (KT), we focus our analysis on observations from 1997 to 2006, covering 5 years of pre and post-program data. Two notes regarding estimation and inference are worth discussing. First, because household outcomes such as income, profits and earnings are likely to exhibit outliers, we winsorize each outcome to the top 1% of the full sample distribution. However, we do report results using untrimmed raw data as

[26]More formally, we define beliefs as the probability of observing high profits as: $b_{i,t+1|t} = \frac{\pi_{\{i,t+1|t\}}}{\pi^g_{\{i,t+1|t\}} - \pi^a_{i,t+1|t}}$. Where π denotes profits.

robustness in the appendix. Second, we conduct inference based on block-bootstrapped standard errors at the village level that incorporate both the estimation of the within-village productivity rankings and the estimation of equation (4) in each bootstrap replication.[27]

Figure 1 presents flexible difference-in-differences estimates of the effect of the rollout of the program on program credit (Panel A) and total short-term credit (Panel B), for high and low productivity households (based on the proxy-variable approach). Consistent with Kaboski and Townsend (2012), we find that indeed, villages with large inverse village sizes experience a large increase in short-term program credit following the implementation of the program, relative to the baseline years. These increases are associated with an average loan size of THB 16,000 for compliers (USD 360 at 2001 exchange rate).

We also find that baseline productivity is not predictive of program borrowing. Panel A from Figure 1 shows that differences in program participation in small villages (more per capita resources) with respect to large villages (less per capita program resources) are orthogonal to productivity. Panel A in Table 3 shows that there are no TFP-based differences in program access either. These results are unlikely to be driven by smaller villages delivering credit to marginal borrowers with lower TFP as we do find that average borrower TFP is not correlated with village size (See Appendix Table A4). Interestingly, this strong "first-stage" for both high and lower productivity households suggests that potential heterogeneity in downstream outcomes is unlikely to be driven by differences in access to program credit.

To analyze whether the program crowded in or crowded out other sources of credit, Panel B from Figure 1 shows the reduced-form effects of the program for high and low productivity businesses. Instead of crowding out other sources of credit, the program appears to have crowded-in other types of credit. The figure shows point estimates that are larger than those associated with program credit (Panel A). While this result holds for high and low productivity households pooled together, the point estimates are larger in the case of high-productivity households, suggesting that the rollout of the program caused them to borrow more from other lenders. One implication is that the failure of the program to provide more credit to higher productivity households is unlikely to be driven by lack of demand. Households with businesses in the top-third of the baseline TFP distribution also borrowed from other sources, likely at higher rates of interest, as program credit was subsidized. However, these differences are not significant on average (see Column 4 from Panel A in Table 3).

This pattern suggests that there was some degree of misallocation as resources were not sys-

[27] Regression tables also present village-clustered standard errors for reference.

tematically delivered to the most productive households (top third of the TFP distribution), and is consistent with evidence of allocative frictions in the program (Vera-Cossio, 2018). However, finding that credit does not flow towards the most productive households does not necessarily imply misallocation in credit markets. For instance, it is possible that returns to credit were already equalized across households before the program. Thus, it is important to quantify heterogeneity in the returns to credit in order to test for misallocation.

5.2 Heterogeneous returns to credit and productivity

In this section, we empirically test whether heterogeneity in pre-program productivity predicts heterogeneity in the returns to credit. We begin by providing a graphical analysis of the effects of the credit expansion induced by the program on income and enterprise profits. The top-left panel of Figure 2 shows that total income did not change significantly for lower-productivity households due to the program, but increased substantially in the case of high-productivity entrepreneurs. To complement the graphical evidence, Panel A from Table 4 presents reduced-form estimates corresponding to the specification in equation (4), which capture the effect of an extra per-capita THB of credit in a given village on household outcomes.[28] Column (1) documents statistically significant differences in the effects of the program on income between high and low productivity households ($p < 0.05$). Columns (3) and (4) show that while there is not substantial heterogeneity in the effects on wage income, productivity-based heterogeneity is large in the case of the effects of the program on profits from household enterprises ($p < 0.1$). These patterns are reassuring because productivity is expected to be more predictive of higher effects in businesses rather than labor earnings. Panel B from Table 4 show qualitatively similar results based on fixed-effects estimates of baseline productivity.[29] However, Appendix Table A5 shows that there is no heterogeneity in input spending.

To analyze the main source of this increase, we also look at the effects of the program on farm profits and non-agricultural business profits. We find no evidence of heterogeneous effects on farm profits (see bottom-left panel of Figure 2). However, we observe quite a different pattern in the case of non-agricultural businesses. The top-right panel of Figure 2 shows that non-agricultural business profits increased dramatically for high-productivity households and not so in the case of

[28] We divided the point estimates from equation (4) by 1,000,000 to provide a THB-to-THB interpretation.

[29] Interestingly, Panel B shows significant increases in wage income for high-A households and not so for lower-A households. This result is consistent with evidence of increases in wages due to the program (Kaboski and Townsend, 2012).

lower-productivity households. Note that while the precision of the estimates decays over time, the relation between baseline productivity and the effects of the program on profits is quite stable. Thus, the TFP-based dispersion in the effects of the credit-expansion program is unlikely to be driven by transitory shocks. Column (8) from Panel A on Table 4 shows that the differences in the reduced-form treatment effects are also statistically significant ($p < 0.10$). Appendix Figure A1 shows that these patterns are qualitatively similar if we use the fixed-effects approach to recover baseline household productivity, and Panel B in Table 4 shows that this alternative approach yields magnitudes that are similar though less precisely estimated than those from our benchmark specification.

These results show that non-agricultural family businesses drive the effects of the program on household profits. One interpretation is that in the rural Thai context, credit constraints are not as binding for agricultural businesses, perhaps due to differences in collateralizability of farm versus non-farm assets or due to preexisting credit options targeting agricultural businesses. Indeed, over one-half of the households in the sample had access to institutional credit at baseline (see Table 1), mainly through the Bank of Agriculture and Agricultural Cooperatives (BAAC) but also through other agriculture-oriented lenders such as production credit groups (PCGs), and cooperatives. The results suggest that in the Thai context, a targeted policy oriented at alleviating constraints for non-agricultural businesses would have complemented the preexisting government-led agricultural programs.

We also analyze whether there is productivity-based heterogeneity in the effects of the program on consumption. Column (3) from Appendix Table (A6) shows that total consumption increases similarly in the case of both high-productivity and low productivity households. We also fail to detect significant heterogeneity on food spending and spending on durables (vehicle and dwelling repairs). One explanation is that while low-productivity households may have borrowed to finance consumption, high-productivity households may have borrowed to generate income and used part of the increases in income to finance consumption. The results are similar using the fixed-effects approach.

Overall, the our estimates reveal a strong, stable relationship between baseline productivity and the returns to credit. Through the lens of our simple theoretical framework, these results suggest misallocation in credit markets. Despite higher returns for higher productivity households, program credit did not flow to the most productive households.

Next, we explore the extent to which the evidence of higher effects for high-productivity households are driven by the creation or expansion of non-agricultural businesses. Appendix Table (A7) shows that there is not significant heterogeneity in the reduced-form effects of the program on the number of non-agricultural businesses and the number of non-agricultural businesses that were started less than one year prior to the survey. These results suggest that increased credit might have been used to boost preexisting, well-established businesses.

We empirically test this premise by applying our framework to the sub-sample of households with non-agricultural businesses that were operating during the pre-intervention periods.[30] We begin by visually analyzing the heterogenous effects of the program on borrowing for this subpopulation. Column (1) from Panel A in Table 5 show that baseline productivity does not predict higher borrowing from the program. However, Column (2) shows that productivity predicts higher reduced-form effects of the program on total short-term credit (p<0.10). Although this result is noisier when we use fixed-effect productivity estimates, the patterns go in the same directions (See Column (2) in Panel B). One interpretation is that while some high-productivity households were able to borrow from the program, they were not able to fully satisfy their needs for liquidity and ended up borrowing from other sources of credit as well.

Next, we analyze whether there was productivity-based heterogeneity in the effects of the program on profits from preexisting non-agricultural household enterprises. Figure 3 shows that profits increased substantially for high-productivity households, but not so for low-productivity households in this subsample. Column 3 from Table 5 complements the graphical evidence by showing that there is significant productivity-based heterogeneity in the program effects on profits ($p < 0.05$). Put together, the results are consistent with the misallocation hypothesis. Resources were not targeted at the most productive and as a result we observe large dispersion in the effects of the program based on baseline TFP.

Our results seem to be driven by high-TFP households increasing non-agricultural business assets ($p < 0.10$) and not to changes in inventories (see Columns (4) and (6)). Thus, high-productivity households seem to have used the increased credit supply to scale up their non-agricultural businesses. There is also significant heterogeneity on wage-labor spending, probably triggered by the large business expansions. However, in the case of high-productivity households, the effects on wage-labor spending (0.3, $p < 0.05$) are not substantial and are relatively small with respect to the effects on assets (8.2, $p < 0.01$). These results are robust to estimating productivity following the fixed-effects approach (see Panel B from Table 5).

[30] We define business owners as households who hold business assets in the period preceding the program.

Figure 4 plots flexible difference-in-difference estimates for high and low productivity owners of pre-existing businesses. While low-productivity households do not increase business assets due to the program, high-productivity households start scaling up their businesses as early as 2001, the year preceding the full rollout of the program. The increase in assets in the case of high-tfp households precedes the increases in business profits due to the program. This pattern is consistent with time-to-build models and suggest that the returns from business expansions are not be immediate.

One possible explanation for observing increases in assets as early as 2001 is related to the early announcement of the program. While the implementation date in each village was unknown to villagers, the program was one of the flagship campaign policies of the Thai Rak Thai (TRT) party (Montesano, 2001). Thus following the victory of the TRT in February of 2001, it is possible that entrepreneurs already incorporated this information in their investment decisions.[31] Although anticipation is possible, it is unlikely that the effects are entirely explained by anticipation. Figure 4 shows that program effects increase over time.

To analyze the extent to which our results are mostly driven by early reactions in 2001, we drop 2001 from the estimating sample and compute reduced-form effects of the program on non-agricultural business assets. Reassuringly, Appendix Figure A2 shows that this exercise yields estimates of the program effects on assets that are even larger than those from our main specification.

5.3 Robustness

Sensitivity to trimming potential outliers. Our main results are based on trimmed outcome variables which were top coded with respect to the 99th percentile of the respective outcome distribution. Appendix Tables A8 and A9 present reduced-form effects of the program using raw, untrimmed outcome data. As expected, both tables show that point estimates tend to be larger but noisier than the ones corresponding to our main specification.

Sensitivity to using productivity rankings. Our main results highlight differences between households who belong to the top third and the bottom two thirds of the productivity distribution in each village. Appendix Tables A14 and A14 show that the results are robust to using the productivity percentile rank itself and suggest that our results are not dependent on how we group households in terms of productivity.

Accounting for labor. One limitation of our empirical analysis is that, due to data constraints,

[31] For instance, providers may sell assets on credit with the idea of recovering part of the value of assets once the program resources are disbursed.

our productivity estimates do not account for the role of labor and only capture variation in output conditional on the stock of capital. As a result, high-productivity households ($HighA$) are the ones that would generate more value added given a certain amount of productive capital, but may not be the ones that would generate more value-added holding constant *both* capital and labor. However, our estimates of productivity would still capture economically meaningful variation in contexts in which the effects of micro-credit programs on household profits are not likely to be driven by adjustments in labor markets. Our results suggest that such a scenario is likely to fit the Thai context.[32]

We report two robustness analyses that try to account for labor using proxies.[33] First, we replicate our analysis estimating the production function in per-capita values in order to account for household size which could be correlated with labor. Second, we replicate the control-function approach including the number of household members who reported working in household production as their main occupation plus the number of hired workers for household non-agricultual businesses as a proxy for labor.[34] Appendix Figures A3 and A4 replicate our main results using these approaches. Though noisier, the patterns are still similar to those corresponding to our main empirical approach.

Accounting for measurement error in capital. It is possible that capital is measured with error, leading to attenuation bias, and over-estimating productivity for capital-intensive households. In order to test the sensitivity of our main results to measurement error in capital, we follow Collard-Wexler and De Loecker (2016) and estimate a value-added production function using expenditures on fixed capital goods in period $t-1$ to instrument for current capital both in the first and second stage of our estimation procedure (see Online Appendix Section(C.4) for details regarding the estimation procedure). This approach leads to higher estimates of β_k even after including labor, however the estimates are noisy as investment can be lumpy in the Thai context. Despite these issues, reassuringly panels (e) and (f) from Figures A3 and A4 show that the results are not qualitatively different to our main estimates and confirm our main results: high-productivity households were better able to convert credit into profits.

[32] Kaboski and Townsend (2012) fail to find average effects on household spending in labor and provide suggestive evidence of impacts on the probability of investment in agricultural assets.

[33] Ideally, we would want to observe information regarding time use. In particular, we would need information on the number of hours allocated to household production by household members and the number of work hours by hired labor. Unfortunately we only observe the number of household members that report mainly working in household enterprises and the number of hired workers for non-agricultural business.

[34] We estimate a slightly modified version of our main approach using the two-stage procedure described by Ackerberg et al. (2015). See Online Appendix Section(C.2) a detailed description of the procedure.

6 IV estimates of the returns to credit

While the reduced-form estimates are important to test the presence of productivity-based heterogeneity, we also provide IV estimates of the local average treatment effect (LATE) of an additional THB of credit on profits corresponding to households who were induced to borrow more due to the program. This approach provides an approximation of the baht-to-baht relationship between total short-term credit and household productive outcomes for program borrowers.

We slightly modify the approach used by Kaboski and Townsend (2012) by using the variation induced by the timing and relative size of the program to instrument for total short-term credit as opposed to program credit only. We chose that specification because we found evidence suggesting that the program crowded in other sources of credit. [35] We then estimate the effects of short-term credit on household outcomes using the following specification:

$$\begin{aligned} y_{i,n,t} =& \beta_1 STCR_{i,n,t} + \beta_2 \text{High A}_{i,n} \times STCR_{i,n,t} + X_{i,n,t}\Gamma \\ &+ \beta_3 \text{High A}_{i,n} + \theta_t \times \text{High A}_{i,n} + \phi_n + \phi_t + \epsilon_{i,n,t} \end{aligned} \tag{10}$$

with first stage:

$$STCR_{i,n,t} = \sum_{\tau=2002}^{\tau=2003} \delta_\tau invHH_n \times \mathbf{I}[t=\tau] + X_{i,n,t}\Sigma + \theta_t + \theta_n + e_{n,t} \tag{11}$$

Here, the parameters of interest are β_1, which captures the LATE of short-term credit on business profits for low-productivity households, β_2 which captures the differential effect of credit between high and low productivity households, and $\beta_1 + \beta_2$ which captures the LATE for high-productivity households. Note that we also need to instrument for High $\text{A}_{i,n} \times STCR_{i,n,t}$. Because High $\text{A}_{i,n}$ is predetermined, we simply construct the first stage by pre-multiplying all terms in the standard first stage by High $\text{A}_{i,n}$. In the structural equation, this yields two endogenous regressors ($STCR_{i,n,t}$ and High $\text{A}_{i,n} \times STCR_{i,n,t}$) and two sets of instruments.

We focus the analysis on the first two years following the rollout of the program for three reasons. First, most experimental evidence is based on outcomes measured between one and two

[35] Moreover, potential responses in local credit markets are likely to occur in this setting. For instance, (Kinnan and Townsend, 2012) show that households rely on indirect access to formal credit to smooth consumption and investment decisions.

years from the initial intervention. Second, the reduced-form analysis presented in the previous section suggests that heterogeneity is more precisely estimated during the first couple of years following the introduction of the program. Finally, the exclusion restriction –i.e., the program only affected household outcomes through short-term credit– is less likely to hold for a longer time horizon. For instance, households may reinvest resources and general equilibrium effects are more likely to kick in as suggested by Buera et al. (2012); indeed, using data a five-year post-program time span, Kaboski and Townsend (2012) detect increases in wages due to the program. Considering these caveats, we emphasize that results corresponding to IV estimates are rather suggestive but still useful as they constitute a tool to compare the approximated financial returns to an extra unit of credit with other estimates in the literature.

We find that our IV estimates imply sizable returns to credit for high-productivity households. First, we focus on the full sample. Table 6 reports IV estimates on household income and profits. Columns (1) and (2) from Panel A show that in the case of high-productivity households, income increases by THB 1.4 to 2.8 per additional THB of total credit, depending on whether we use trimmed data or raw data respectively. These estimates are significantly different than those corresponding to lower-productivity households. Columns (5) and (6) report large effects on profits in the case of high-productivity households which imply annual returns to credit of 100-250% (See bottom rows form Panel A) and rather small and insignificant negative returns for low-productivity households. In fact, we do find negative significant effects of short-term credit on total income in the case of low-productivity households. One interpretation is that given that selection into the program was not a function of productivity, the program resources ended up financing fruitless projects in the case of lower-productivity households.

Second, we find even higher returns to credit when we focus on high-productivity owners of preexisting non-agricultural businesses. The bottom panel of Table 7 shows that the effects on business profits are on the order of THB 2.9 to 5 per additional THB of total credit. Though noisier, the results are robust to including up to five years following the introduction of the program (see Appendix tables A16 and A17).

The point estimates suggest effects that are similar to the effects found by Crepon et al. (2015) in Morocco (2.4). Relative to the literature estimating the returns to cash grants, the returns to credit for Thai high-productivity business owners are as high as those of entrepreneurs who were identified as being of "high-growth potential" by their peers in India–3.3 increase in annual profits per additional rupee from grants (Hussam et al., 2017),[36] and annual returns to cash grants of 2.3-3.9 for Mexican firms(McKenzie and Woodruff, 2008).[37]

Table 7 also shows that, in the case of high-productivity households, there are neither meaningful effects on inventories nor expenditures on wage work, but that there are substantial increases in business assets in the order of THB 4 to 8 per additional THB of credit. One potential explanation to such magnitudes is that, consistent with the idea of the program crowding in other sources of credit, the results suggest that households may have also used cash holdings to complement credit in financing a lumpy investment. Though noisier, the results are robust to including up to five years following the introduction of the program (see Appendix tables A16 and A17).

Estimates of returns to fixed capital. Our empirical exercises documents large returns to additional credit in the case of households with preexisting non-agricultural businesses, and that these returns are mostly driven by increases in non-agricultural assets. Given such findings, it is natural to assess the profitability of such investments.

Under the assumption that credit only affected profits through changes in fixed capital, we use the variation in credit supply induced by the program to quantify the returns to non-agricultural fixed capital. Two pieces of evidence suggest the validity of this assumption. First, our previous results showed that non-wage expenses did not significantly increase due to the program. Second, we also find that wage expenses did increase significantly but not substantially as the effects of credit on wage expenses only represents 3% of the effect of credit on assets.[38] However, we do acknowledge that credit could have also modified the use of unpaid household labor, or could have been spent on improving productivity or on non-tangible inputs which are not measured in our data. With these caveats in mind we conduct some back-of-the-envelope calculations to approximate the returns to fixed capital in the Thai context.

To do so, we simply divide our estimates of the effect of credit on profits by our estimates of the effect of credit on business assets, for the subsample of households with pre-exisitng non-agricultural businesses. Our point estimates suggest that profits increased by THB 2.9 per additional THB of credit and that assets increased by THB 4.6 per additional THB of credit (see the bottom panel in Table 7).We find that annual profits increased by THB 0.63 per additional THB of fixed assets:

[36] Hussam et al. (2017) document returns to capital grants as high as 28% monthly, which multiplied by 12 represent 330%

[37] McKenzie and Woodruff (2008) show that monthly profits increase between 292-487 pesos after receiving 1,500 pesos in cash grants.

[38] Table 7 shows that non-agricultural wage expenses increased by THB 0.18 per additional THB of credit. In contrast, non-agricultural assets increased by THB 4.6 per additional THB of credit.

an annual rate of return of 63% per annum (5.2% per month) . These estimates are substantially higher than the interest rates charged by the existing lenders in the Thai context: 7% per annum in the case of program loans, 12% for loans from BAAC bank and 22% for loans from informal lenders. Moreover, the estimates are consistent with other estimates of the returns to fixed capital in the literature: 39.6% annual for pre-existing firms winning a business plan competition in Nigeria (McKenzie, 2017), or estimates as large as 66-70% annual for the case of SMEs in Sri-Lanka (de Mel et al., 2008).[39]

Overall, the results highlight the existence of high-returns to capital for non-agricultural businesses, though only for high-productivity households with consolidated enterprises. One important implication of these results is that the success of public efforts in expanding access to credit is bounded by the ability of policy makers to effectively deliver resources to high-productivity households. Had there been less misallocation, the population-level effects of the program could have been substantially higher.

7 Potential gains from reallocation

So far, we have documented a large of degree of dispersion in the returns to credit that is driven by variation in pre-program household productivity. Following Hsieh and Klenow (2009), we interpret the dispersion in the returns to credit as evidence of misallocation in credit markets. To quantify the degree of misallocation, we compare the observed dispersion in the returns to credit with the degree of dispersion that would arise under a benchmark allocation of program credit. We do so by simulating a counterfactual allocation that delivers more program credit to higher-productivity households, while fixing the total amount of program credit and program borrowers in the village.

We proceed in four steps. First, within each village, we rank households in terms of pre-intervention productivity. Second, we construct rankings of average program borrowing during the first two years of the program. We then allocate larger amounts of program credit based on village productivity rankings. For example, if the entrepreneur with the highest amount of program loans in the village obtained on average THB 15000, then our counterfactual allocation would deliver THB 15000 to the household with the highest productivity in the village. Third, we use our reduced-form estimates from section 5 to compute the amount of total short-term credit under the counterfactual allocation. This step is important as we documented evidence of program credit crowding in short-

[39]We obtain these values by multiplying the monthly returns reported in each study by 12.

term credit from other sources. Finally, we combine our IV estimates of the effect of total short-term credit on profits (β_1) and the interaction effect with productivity (β_2) to obtain the distribution of marginal returns to credit ($mrc = \beta_1 + \beta_2 \hat{A}_i$) for the pool of borrowers under the actual and the counterfactual allocation.

We begin by providing graphical evidence of the gains from reallocation. Figure 5 depicts the distribution of marginal returns to credit for actual and counterfactual borrowers, normalized with respect to the observed village mean. The figure shows that a move towards allocating credit to higher productivity entrepreneurs would maginify the marginal returns to credit, relative to the actual allocation. Reallocating resources based on pre-intervention productivity shifts the distribution of returns to the right. Importantly, the dispersion of the distribution of marginal returns to credit reduces dramatically.

Table 8 quantifies the results from reallocation. Panel A shows that redirecting credit towards the most productive entrepreneurs would increase the average return to credit from THB 0.14 per additional THB of short-term credit, in the case of the set of program borrowers, to THB 0.67 per additional THB of credit,in the case of counterfactual borrowers. Relative to the actual pool of program borrowers, reallocation would also reduce the standard deviation of the returns to credit by 18%, and the ratio of returns associated to the top and bottom 25% of the distribution by 32%. Consistent with our results suggesting that misallocation is concentrated in the subpopulation of non-agricultural incumbent businesses, Panel B of Table 8 shows that reallocating resources to higher-TFP households would lead to even larger increases in mean marginal returns (400%), and larger reductions in dispersion (21% decrease in the standard deviation and 27% decrease in the interquartile ratio), with respect to the actual allocation. Overall, the results suggest that targeting credit at the most productive households could have magnified the effects of the Million Baht program.

7.1 Gains in village-level output

In this section we quantify the aggregate (village level) gains from reallocating program resources to the most productive households under three regimes. We first focus on village-level output gains largely coming from reallocating resources from low- to high-productivity firms within each sector (agricultural and non-agricultural) in each village. We do so as our results that the dispersion in returns to credit is driven by firms from the non-agricultural sector. Second, we reallocate resources across firms within each village, regardless of the sector. This regime allows us to quantify gains from switching program credit from lower productivity agricultural firms to higher productivity non-agricultural firms in the same village. Third, we compute village-level gains from reallocating resources across all firms in our sample, regardless of their sector or location. The former two regimes preserve the total amount of program credit and program borrowers in each village, while the latter regime relaxes the key feature of the program—THB 1 million for all villages— while holding constant the aggregate program portfolio.

We quantify counterfactual changes in village-level output in the following way. For each regime, we reallocate program credit based on productivity rankings within the relevant category as described in Section 7. Once program credit is reallocated, we compute counter-factual total borrowing $STCR^C_{i,t}$ as follows:

$$STCR^C_{i,t} = STCR_{i,t} + (\frac{\hat{\delta}^{stcr}_1}{\hat{\delta}^{vf}_1} + \frac{\hat{\delta}^{stcr}_2}{\hat{\delta}^{vf}_2} \times HighA_i) * (vf^C_{i,t} - vf_{i,t})$$

Where δ_1,δ_2 denote the reduced-form effects of the program's rollout (by higher- and lower-productivity) on total short-term credit and program credit. This procedure accounts for potential increases (decreases) in total borrowing due to changes in program borrowing induced by the counterfactual exercise. However, it is unclear if high-productivity households would have demanded more non-program credit had they obtained enough credit from the program. To consider such possibility we also report results from computing counterfactual short-term credit as $STCR^C_{i,t} = STCR_{i,t} - vf_{i,t} + vf^C_{i,t}$, that is assuming that the reallocation of program credit fully satisfies each household's demand for credit.

Next, we use our IV estimates of the effect of short-term credit on firm-level assets(K) and value-added(Y) to quantify counterfactuals for higher- and lower-productivity firms.

$$K^C_{i,t} = K_{i,t} + (\hat{\beta}^k_1 + \hat{\beta}^k_2 \times HighA_i) * (STCR^C_{i,t} - STCR_{i,t})$$
$$Y^C_{i,t} = Y_{i,t} + (\hat{\beta}^y_1 + \hat{\beta}^y_2 \times HighA_i) * (STCR^C_{i,t} - STCR_{i,t})$$

Finally, we aggregate at the village level adding actual and counterfactual measures of output (value added) and the stock of capital across sample firms in each village. We then compute gains in from reallocation as:

$$
\tilde{Y}_v = \frac{\sum_i^{N_v} Y_{i,v}^C}{\sum_i^{N_v} Y_{i,v}} - 1
$$

$$
\tilde{K}_v = \frac{\sum_i^{N_v} K_{i,v}^C}{\sum_i^{N_v} K_{i,v}} - 1
$$

One advantage of our approach is that it uses quasi-experimental variation in credit supply to compute returns to credit, and thus identification of the parameters used in the counterfactual analysis do not rely on structural assumptions. However, such advantage comes at the cost of implicitly assuming that the marginal return to an extra unit of total credit is constant with respect to total borrowing. With this caveat in mind, we proceed to analyze aggregate village-level gains in output and capital. Table 9 reports gains from reallocation under different counterfactual regimes. It reports average village gains by village size, and population weighted aggregate gains. In each panel, the first two columns present results after adjusting total short-term credit for potential crowding in, while the last two columns present results without adjusting for crowding in. As each scenario relies on different assumptions, they help us provide upper and lower bounds of the gains from reallocation.

Panel A shows that, reallocating program credit from lower- to higher-productivity firms within sectors and villages, on average, would increase village-level output between 10-18%. Likewise, aggregate capital would increase between 3-14%. Note that we obtain larger gains when we allow firms to borrow more from non-program credit after the reallocation exercise (columns 1 and 2) than in the case in which we assume that reallocating resources fully satisfies the entrepreneur's demand for credit. The gains from reallocation vary substantially by village size. For instance, the gains seem to be driven by smaller villages in which there are more per-capita resources to reallocate. Consequently, the village-size adjusted average gains from within sector reallocation are modest as they are in the range of 0.6 to 3.8%. Panel B shows that reallocating program credit within villages allowing for cross-sector reallocation leads to even larger gains in aggregate output. The village-size adjusted average output gains are of the order of 2.9 to 10.3%. As in Panel A, the gains from reallocation are larger for smaller villages.

One important feature of the program is that the Thai government decided to allocate $ THB 1 Million per village, regardless of village size. This constraint may have led to further distortions in credit markets. In Panel C, we quantify the gains from reallocating credit from lower- to higher-productivity entrepreneurs regardless of their location, and thus allowing village-level program resources to vary across villages. Panel C shows that the population-weighted average gains are of the order of 6.6 to 51%, depending on the assumptions regarding crowding in of non-program credit after reallocation. Note that as opposed to the case of the within village reallocation regimes in Panels A and B, gains seem to be concentrated in larger villages that received relatively smaller endowments of per-household program credit. One implication of this finding is that ignoring spatial productivity differences in credit expansion policies may come at the costs of substantial costs.

8 Concluding remarks

We use the context of one of the largest microfinance lending programs to provide two main results. First, we show that program borrowing was not a function of productivity. Thus, some high-return entrepreneurs ended up not obtaining credit but other less profitable businesses did. Second, we document a large degree of TFP-based heterogeneity in the effects of the program on business profits. Together, these two pieces of evidence suggest misallocation. Heterogeneity is stronger for entrepreneurs with pre-existing businesses and implies high returns to credit for high-TFP entrepreneurs. Such returns are similar to returns to cash-grant programs for SMEs in developing countries (de Mel et al., 2008; McKenzie and Woodruff, 2008; Hussam et al., 2017). Put together, our results show that allocative frictions in credit markets may impede the flow of capital to the most productive firms and that there is substantial heterogeneity in the returns to capital across family enterprises.

Our analysis requires a suitable method to estimate household productivity. Popular methods for estimating production functions assume that firms can freely adjust inputs in response to TFP shocks (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015), which is unlikely in the case of credit-constrained businesses (Shenoy, 2017a). We propose a novel implementation of the control-function approach using beliefs about future profits as a proxy for TFP shocks, which makes no assumptions regarding how investment and inputs are adjusted and is suitable to a context with potential credit market frictions.

Our results have important policy implications. They suggest that improved screening and targeting could greatly magnify the impacts of credit expansions. While we document high-returns to credit for high productivity households, we also find that program borrowing was orthogonal to baseline productivity. Thus the policy challenge involves effectively targeting high-productivity entrepreneurs. Overcoming this challenge could potentially entail improvements in externally identifying entrepreneurs (see Fafchamps and Woodruff (2017), Hussam et al. (2017), and Mckenzie and Sansone (2017)). The screening mechanism is also likely to matter. For instance, in the case of the Million Baht program, Vera-Cossio (2018) shows that the allocation of program credit was heavily influenced by connections to local leaders. In contrast, Beaman et al. (2014) show that relying only on interest rates as a screening device may allow high-return households to self-select into credit.

References

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

Ahlin, C. and R. Townsend (2007). Selection into and across credit contracts: Theory and field research. *Journal of Econometrics 136*, 665–98.

Alfaro, L., A. Charlton, and F. Kanczuk (2008). Plant size distribution and cross-country income differences. *NBER International Seminar on Macroeconomics 5*(1), 243–272.

Atkin, D., A. K. Khandelwal, and A. Osman (2019, May). Measuring productivity: Lessons from tailored surveys and productivity benchmarking. *AEA Papers and Proceedings 109*, 444–49.

Banerjee, A., E. Breza, E. Duflo, and C. Kinnan (2015). Do credit constraints limit entrepreneurship? heterogeneity in the returns to microfinance. *Working Paper*.

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

Banerjee, A. V. and E. Duflo (2005). Growth theory through the lens of development economics. *Handbook of economic growth 1*, 473–552.

Banerjee, A. V. and E. Duflo (2014). Do firms want to borrow more? testing credit constraints using a directed lending program. *The Review of Economic Studies 81*(2), 572–607.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2013). Cross-country differences in productivity: The role of allocation and selection. *American Economic Review 103*(1), 305–34.

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

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

Buera, F. J., J. P. Kaboski, and Y. Shin (2012, March). The macroeconomics of microfinance. Working Paper 17905, National Bureau of Economic Research.

Buera, F. J., J. P. Kaboski, and Y. Shin (2015). Entrepreneurship and financial frictions: A macrodevelopment perspective. *Annual Review of Economics 7*(1), 409–436.

Buera, F. J., J. P. Kaboski, and Y. Shin (2017). Taking Stock of the Evidence on Micro-Financial Interventions. In *The Economics of Poverty Traps*, NBER Chapters. National Bureau of Economic Research, Inc.

Cai, S., A. Park, and S. Wang (2017). Microfinance can raise incomes. Working paper.

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.

Crepon, B., F. Devoto, E. Duflo, and W. Pariente (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.

de Mel, S., D. McKenzie, and C. Woodruff (2008). Returns to capital in microenterprises: evidence from a field experiment. *Quarterly Journal of Economics 123*, 1329–1372.

Fafchamps, M. and C. Woodruff (2017). Identifying gazelles: Expert panels vs. surveys as a means to identify firms with rapid growth potential. *World Bank Economic Review 31*(3), 670–686.

Foster, L., C. Grim, J. Haltiwanger, and Z. Wolf (2016, May). Firm-level dispersion in productivity: Is the devil in the details? *American Economic Review 106*(5), 95–98.

Foster, L., J. Haltiwanger, and C. J. Krizan (2006). Market selection, reallocation, and restructuring in the us retail trade sector in the 1990s. *The Review of Economics and Statistics 88*(4), 748–758.

Gollin, D. and C. R. Udry (2019, January). Heterogeneity, measurement error and misallocation: Evidence from african agriculture. Working Paper 25440, National Bureau of Economic Research.

Handel, B. and J. Schwartzstein (2018, February). Frictions or mental gaps: What's behind the information we (don't) use and when do we care? *Journal of Economic Perspectives 32*(1), 155–78.

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.

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

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

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

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

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies 70*(2), 317–341.

McKenzie, D. (2017). Identifying and spurring high-growth entrepreneurship: experimental evidence from a business plan competition. *American Economic Review 107*(8), 2278–2307.

McKenzie, D. and C. Woodruff (2008). Experimental evidence on returns to capital and access to finance in mexico. *World Bank Economic Review 22*(3), 457–482.

Mckenzie, D. J. and D. Sansone (2017, December). Man vs. machine in predicting successful entrepreneurs : evidence from a business plan competition in Nigeria. Policy Research Working Paper Series 8271, The World Bank.

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.

Midrigan, V. and D. Y. Xu (2014, February). Finance and misallocation: Evidence from plant-level data. *American Economic Review 104*(2), 422–58.

Montesano, M. J. (2001). Thailand in 2000 ¡em¿shifting politics, dragging economy, troubled border¡/em¿. *Asian Survey 41*(1), 171–180.

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

Restuccia, D. and R. Rogerson (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic dynamics 11*(4), 707–720.

Restuccia, D. and R. Rogerson (2017, August). The causes and costs of misallocation. *Journal of Economic Perspectives 31*(3), 151–74.

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

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

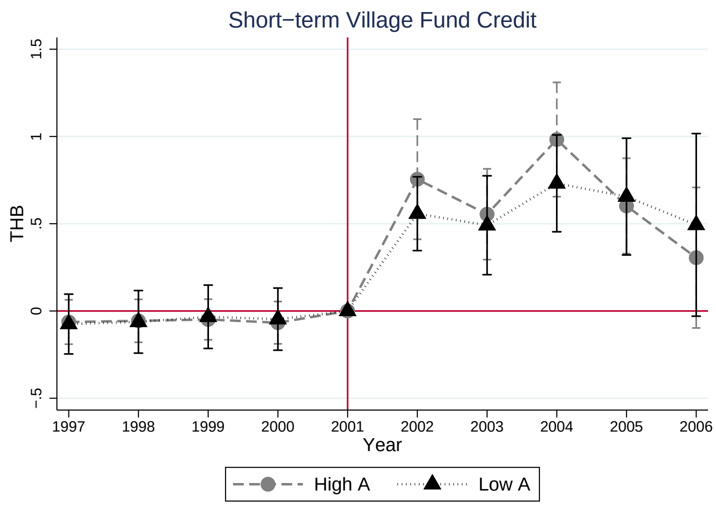
Shenoy, A. (2017b). Market failures and misallocation. *Journal of Development Economics 128*(C), 65–80.

Townsend, R. M. (2007a). Townsend Thai project household annual resurvey, 1998-2006. [Computer file]. 2nd Release. Chicago, IL: Social Sciences Computing Services, the University of Chicago [Producer & distributor].

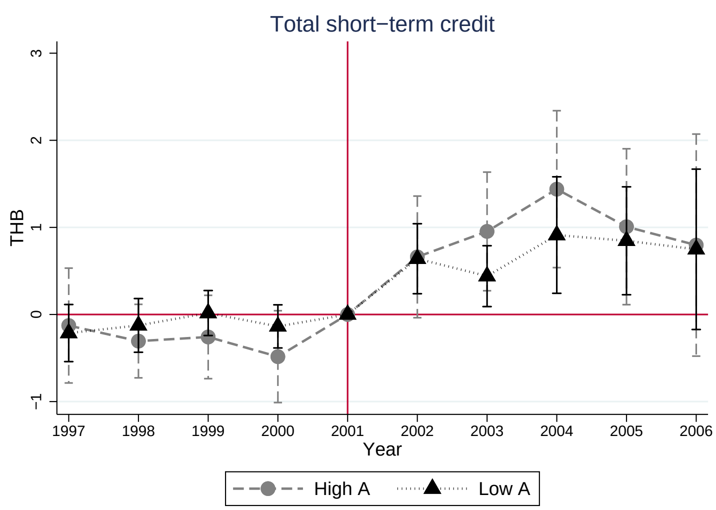
Townsend, R. M. (2007b). Townsend Thai project initial household survey, 1997. [Computer file]. 2nd Release. Chicago, IL: Social Sciences Computing Services, the University of Chicago [Producer & distributor].

Vera-Cossio, D. (2018). Targeting credit through community members. Working paper, University of California, San Diego.

9 Figures



(a) Program Credit



(b) Total Short-term Credit

Figure 1: Effects on short-term credit

Note: The figure depicts flexible difference-in-differences estimates corresponding to the specification in (5). Each dot represents differences in program borrowing between households from villages with high and low per-capita program funds, for each year, with respect to the year of the announcement of the program (2001). Each coefficient has been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome (in THB). High A: household belongs to the top-third of the baseline productivity distribution in each village. Low A: household belongs to the bottom two-thirds of the baseline productivity distribution in each village. Productivity estimates correspond to the control-function approach using household beliefs about profits as a proxy variable. 95% confidence intervals are computed based on standard errors, which are clustered at the village level to account for the empirical design.

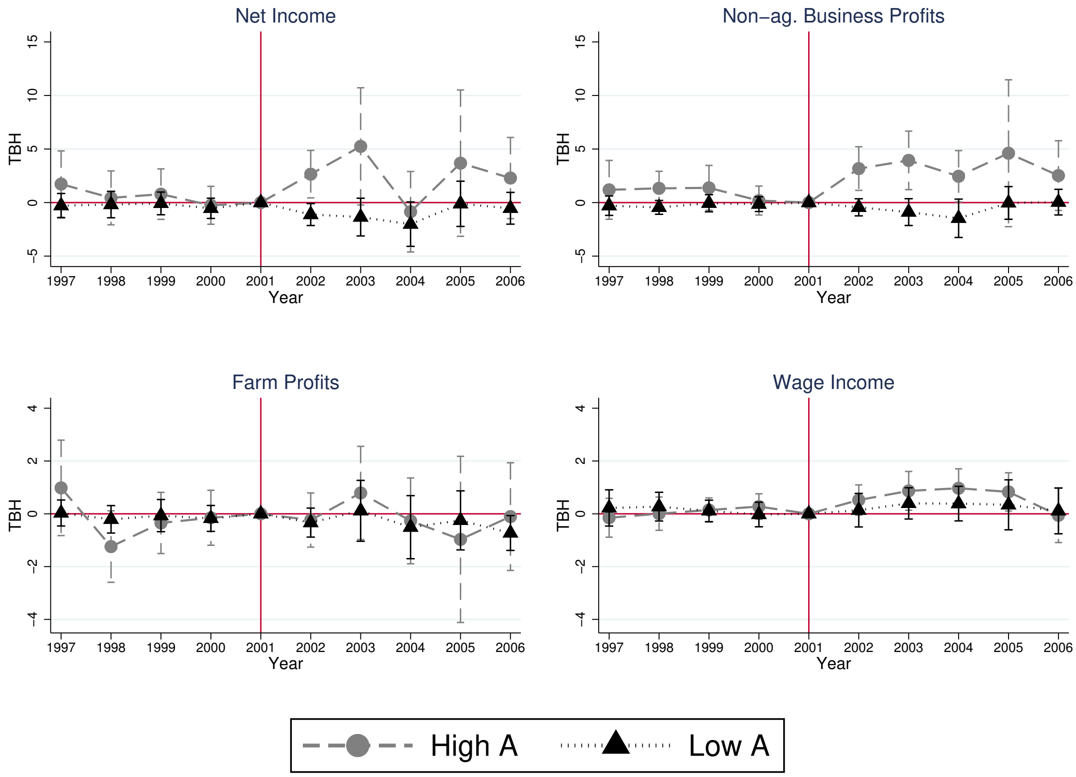


Figure 2: Reduced-form effects on household income - Proxy-variable approach

Note: The figure depicts flexible difference-in-differences estimates corresponding to the specification in (5). Each dot represents differences in program borrowing between households from villages with high and low per-capita program funds, for each year, with respect to the year of the announcement of the program (2001). Each coefficient has been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome (in THB). High A: household belongs to the top-third of the baseline productivity distribution in each village. Low A: household belongs to the bottom two-thirds of the baseline productivity distribution in each village. Productivity estimates correspond to the control-function approach using household beliefs about profits as a proxy variable. 95% confidence intervals are computed based on standard errors, which are clustered at the village level to account for the empirical design.

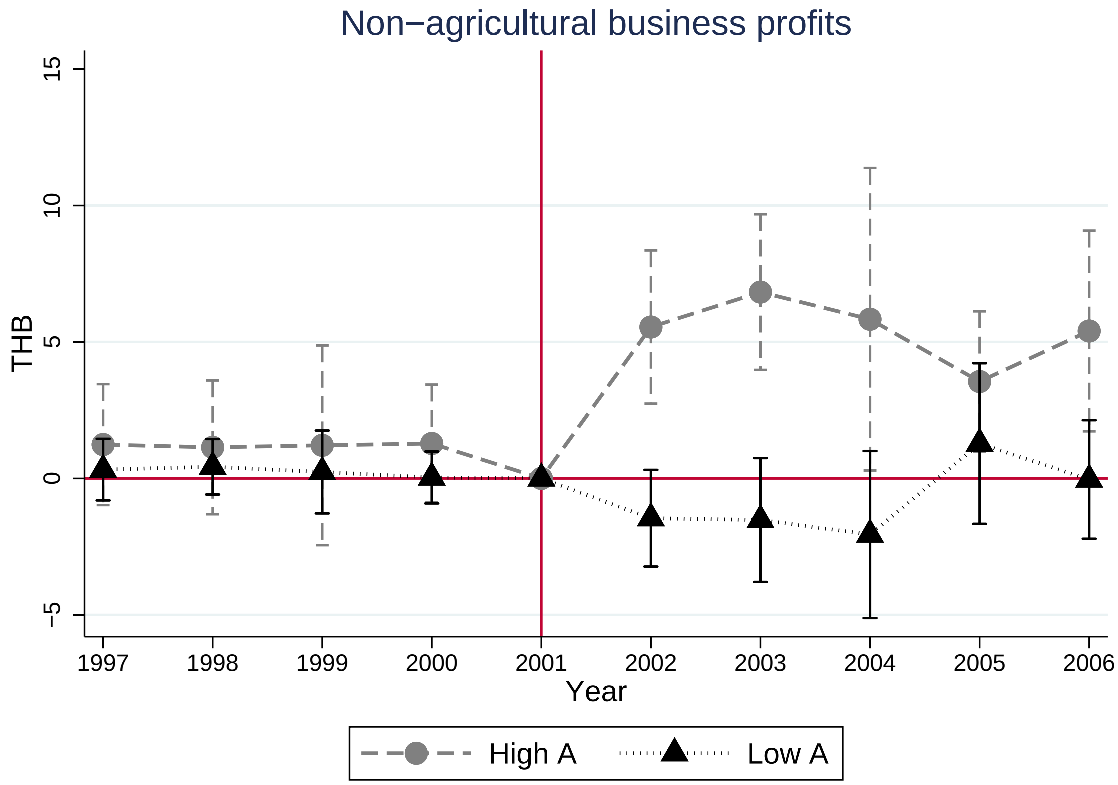


Figure 3: Reduced-form effects on non-agricultural business profits (preexisting non-agricultural businesses)
Note: The figure depicts flexible estimates corresponding to the specification in (5).Each dot represents differences in program borrowing between households from villages with high and low per-capita program funds, for each year with respect to the year of the announcement of the program (2001). Each coefficient has been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on off-farm business profits (measured in baht). The effects are estimated over a sample of 230 households who reported holding business assets during the year preceding the rollout of the program. High A: household belongs to the top-third of the baseline productivity distribution in each village. Low A: household belongs to the bottom two-thirds of the baseline productivity distribution in each village. Productivity estimates correspond to the control function approach using household beliefs about profits as a proxy variable. 95 % confidence intervals are computed based on standard errors, which are clustered at the village level to account for the empirical design. The dependent variable is winsorized with respect to the top 1% of the distribution.

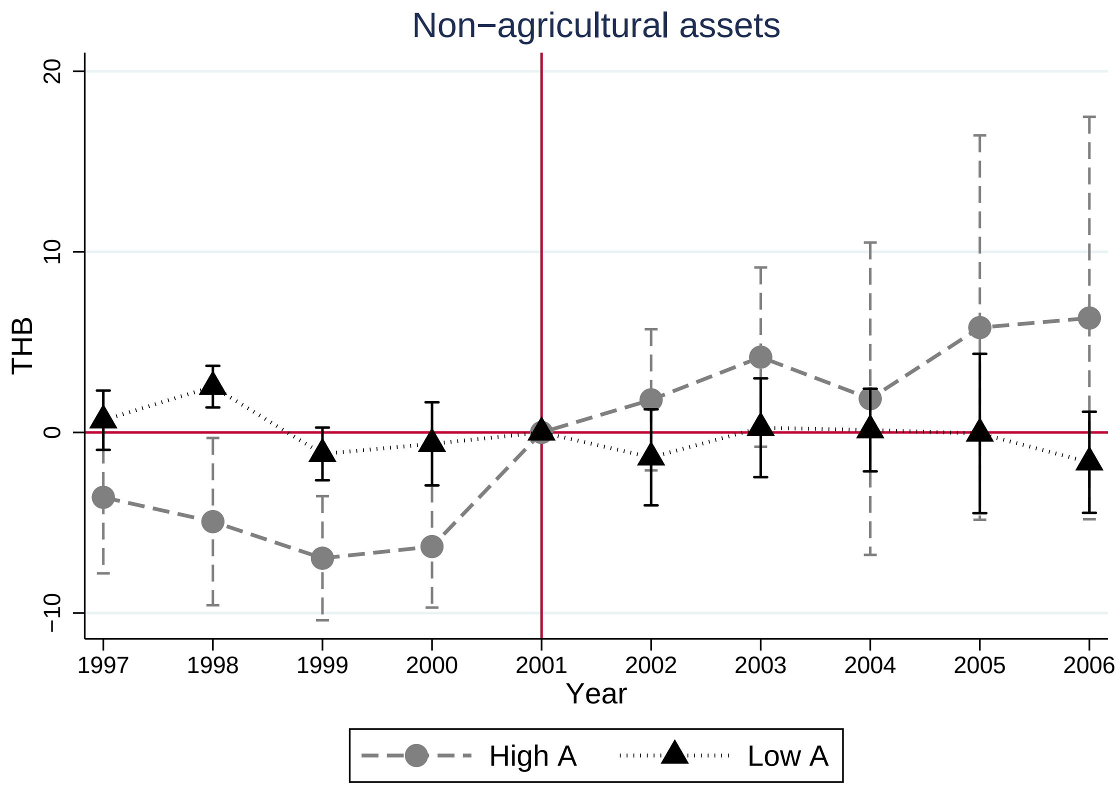


Figure 4: Reduced-form effects on assets (preexisting non-agricultural businesses)

Note: The figure depicts flexible estimates corresponding to the specification in (5). Each dot represents differences in program borrowing between households from villages with high and low per-capita program funds, for each year with respect to the year of the announcement of the program (2001). Each coefficient has been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the value of off-farm business assets (measured in baht). The effects are estimated over a sample of 230 households who reported holding business assets during the year preceding the rollout of the program. High A: household belongs to the top-third of the baseline productivity distribution in each village. Low A: household belongs to the bottom two-thirds of the baseline productivity distribution in each village. Productivity estimates correspond to the control function approach using household beliefs about profits as a proxy variable. 95 % confidence intervals are computed based on standard errors, which are clustered at the village level to account for the empirical design. The dependent variable is winsorized with respect to the top 1% of the distribution.

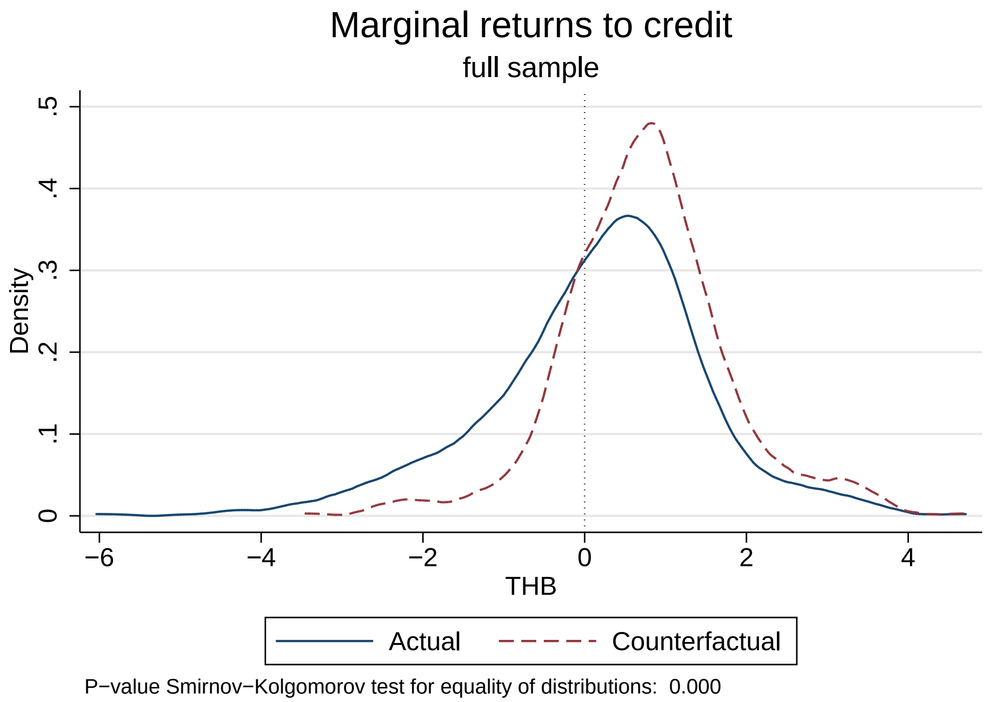


Figure 5: Distribution of actual and counterfactual marginal returns to credit

Note: The figure depicts the distribution of the marginal returns to credit for actual program borrowers ("Actual"), and the set of borrowers who would obtain credit under an allocation based on within village productivity rankings ("Counterfactual"). Both distributions are centered with respect to the village mean of the observed distribution.

10 Tables

Table 1: Baseline Summary Statistics

Variable	N	Mean	SD
Household head is a male	4423	0.74	0.44
Age (household head)	4423	52.85	13.43
Years of schooling (household head)	4343	6.04	3.18
Number of household members	4603	4.44	2.06
Farm (share of net operating income)	4291	0.5	2.29
Fish/shrimp (share of net operating income)	4291	-0.03	2.4
Off-farm business (share of net operating income)	4291	0.1	0.82
Wage income (share of net operating income)	4291	0.43	1.88
Number of household off-farm businesses	4423	0.35	0.55
Household opened a new business (past 12 months)	4603	0.04	0.2
Net per-capita income (THB)	4423	21299	34592
Per-capita consumption spending (THB)	4423	12046	35602
Household borrows (institution or informal)	4603	0.78	0.41
Household borrows from formal/quasi-formal sources of credit	4603	0.56	0.5
Household borrows from informal sources of credit	4603	0.49	0.5

Note: The table presents summary statistics corresponding to the study sample and survey waves preceding the program (1997-2001). Farm activities include cultivation of several crops as well as produce from livestock. Institutional credit includes credit from commercial banks, BAAC (the state-owned bank) and other quasi-formal sources of credit such as cooperatives, and village-credit groups. Exchange rate THB/USD (2001) : 44.51.

Table 2: Estimates of value-added production functions

Panel A: Capital elasticities			
	(1)	(2)	(3)
	OLS	FE	Control function
β_k	0.360***	0.015	0.380***
	(0.025)	(0.036)	(0.050)
Observations	2,622	2,622	2,622
R-squared	0.137	0.004	
# of Households	835	835	835
Panel B: Productivity characteristics			
	OLS	FE	Control function
Persistence ω			0.238***
			(0.032)
Mean ω	5.15	9.67	4.98
SD ω	1.02	1.17	0.06

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

Standard errors in parentheses

Note: The table presents estimates of the elasticity of value-added with respect to capital β_k obtained using the 5 survey waves preceding the program (1997-2001). Column (1) presents OLS estimates for reference, columns (2)-(3) present estimates computed through the fixed-effects approach and the control-function approach, respectively. The bottom panel presents summary statistics for the estimates of log-productivity ($\omega = log(A)$). Standard errors corresponding to the control-function approach are computed using block bootstrap with 1000 iterations.

Table 3: First stage: Program effects on borrowing

Panel A: Proxy-variable approach

	Prob. of borrowing		*Short-term credit (THB)*	
	VF	Total	VF	Total
	(1)	(2)	(3)	(4)
Post X Inv HH X High Productivity	-0.0369	-0.0227	-0.0313	0.288
	(0.0325)	(0.0417)	(0.0731)	(0.265)
	[0.033]	[0.036]	[0.056]	[0.316]
Post X Inv HH	0.172***	0.0944**	0.591***	0.793***
	(0.0550)	(0.0363)	(0.122)	(0.215)
	[0.055]	[0.036]	[0.106]	[0.304]
RF effect - High Productivity	0.135**	0.0716**	0.560***	1.081***
SE	(0.052)	(0.038)	(0.102)	(0.307)
SE (bootstrap)	[0.058]	[0.041]	[0.114]	[0.346]
Observations	8659	8659	8659	8659
Number of households	922	922	922	922
R-Squared	0.613	0.457	0.601	0.537

Panel B: Fixed-effects approach

	Prob. of borrowing		*Short-term credit (THB)*	
	VF	Total	VF	Total
	(1)	(2)	(3)	(4)
Post X Inv HH X High Productivity	-0.128***	-0.0828*	-0.294***	0.0145
	(0.0376)	(0.0426)	(0.0886)	(0.272)
	[0.048]	[0.053]	[0.119]	[0.471]
Post X Inv HH	0.199***	0.115***	0.668***	0.877***
	(0.0475)	(0.0398)	(0.0851)	(0.181)
	[0.054]	[0.044]	[0.096]	[0.245]
RF effect - High Productivity	0.0706	0.0324	0.374***	0.892*
SE	(0.056)	(0.03)	(0.142)	(0.348)
SE (bootstrap)	[0.055]	[0.04]	[0.141]	[0.517]
Observations	8659	8659	8659	8659
Number of households	922	922	922	922
R-Squared	0.615	0.457	0.604	0.539

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

Note: The table reports the reduced-form estimates of the effects of the program as a function of productivity estimated through the proxy-variable approach (Panel A) and the fixed-effects approach (Panel B). Each column reports coefficients from regressions following the specification on equation 4. Post X Inv HH captures the reduced-form effect of the program for households belonging to the bottom-two thirds of the baseline productivity distribution. Post X Inv HH X High Productivity report the difference in the reduced form effects of the program between households in the top-third and bottom-two-thirds of the baseline productivity distribution. Reduced-form effects for high-productivity households are computed by adding up the coefficients on Post X Inv HH and Post X Inv HH X High Productivity. Standard errors, in parentheses, are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are also presented in brackets and are computed using 500 bootstrap samples and clustered at the village level. Short-term credit involves loans with a term shorter than a year.

Table 4: Reduced-form effects of the program on household income and profits

	Household Income	Wage Income	Profits	Farm Profits	Shrimp/Fish Profits	Non-ag Bus.Profits
Panel A: Proxy-variable approach						
	(1)	(2)	(3)	(4)	(5)	(6)
Post X Inv HH X High Productivity	2.103**	0.411	1.346*	0.315	0.0219	1.01*
	(0.993)	(0.375)	(0.505)	(0.357)	(0.0241)	(0.404)
	[0.871]	[0.443]	[0.705]	[0.333]	[0.049]	[0.575]
Post X Inv HH	-0.901**	0.0871	-0.655**	-0.346	-0.0196	-0.289*
	(0.380)	(0.203)	(0.262)	(0.208)	(0.0339)	(0.280)
	[0.381]	[0.223]	[0.236]	[0.208]	[0.044]	[0.15]
RF effect - High Productivity	1.202	0.498	0.691	-0.0312	0.002	0.720
SE	(0.823)	(0.294)	(0.491)	(0.319)	(0.016)	(0.364)
SE (bootstrap)	[0.752]	[0.339]	[0.635]	[0.307]	[0.028]	[0.552]
Observations	8659	8659	8659	8659	8659	8659
Number of households	922	922	922	922	922	922
R-Squared	0.575	0.714	0.445	0.421	0.161	0.408
Panel B: Fixed-effects approach						
	Household Income	Wage Income	Profits	Farm Profits	Shrimp/Fish Profits	Non-ag Bus.Profits
	(1)	(2)	(3)	(4)	(5)	(6)
Post X Inv HH X High Productivity	1.443	0.623*	0.582	-0.0545	-0.0435	0.680
	(0.976)	(0.352)	(0.565)	(0.396)	(0.0629)	(0.411)
	[1.266]	[0.387]	[0.607]	[0.435]	[0.069]	[0.529]
Post X Inv HH	-0.528	0.0542	-0.317	-0.183	0.00167	-0.136
	(0.357)	(0.181)	(0.288)	(0.206)	(0.0320)	(0.279)
	[0.396]	[0.191]	[0.245]	[0.209]	[0.015]	[0.177]
RF effect - High Productivity	0.915	0.677	0.265	-0.237	-0.0419	0.544
SE	(0.868)	(0.301)	(0.542)	(0.369)	(0.051)	(0.381)
SE (bootstrap)	[1.116]	[0.343]	[0.614]	[0.409]	[0.075]	[0.533]
Observations	8659	8659	8659	8659	8659	8659
Number of households	922	922	922	922	922	922
R-Squared	0.575	0.715	0.444	0.421	0.161	0.408

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

Note: The table reports the reduced-form estimates of the effects of the program as a function of productivity estimated through the proxy-variable approach (Panel A) and the fixed-effects approach (Panel B). All dependent variables are winsorized with respect to the top 1% of the full sample distribution. The coefficients have been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome. Each column reports coefficients from regressions following the specification on equation 4. Post X Inv HH captures the reduced-form effect of the program for households belonging to the bottom-two thirds of the baseline productivity distribution. Post X Inv HH X High Productivity report the difference in the reduced form effects of the program between households in the top-third and bottom-two-thirds of the baseline productivity distribution. Reduced-form effects for high-productivity households are computed by adding up the coefficients on Post X Inv HH and Post X Inv HH X High Productivity. Standard errors, in parentheses, are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are also presented in brackets and are computed using 500 bootstrap samples and clustered at the village level. Household profits include farm, fishing and shrimping and non-agricultural business profits. Farm profits include profits from agriculture and livestock. All dependent variables are measured in THB. Exchange rate THB/USD (2001) : 44.51.

Table 5: Reduced-form effects of the program on preexisting non-agricultural businesses

Panel A: Proxy-variable approach

	(1) VF credit	(2) Total credit	(3) Profits	(4) Non-wage Expenses	(5) Wage Expenses	(6) Assets
Post X Inv HH X High Productivity	-0.0476	1.102*	5.680***	1.406	0.294**	9.241***
	(0.150)	(0.570)	(1.559)	(4.416)	(0.125)	(2.751)
	[0.144]	[0.656]	[1.79]	[3.998]	[0.139]	[2.711]
Post X Inv HH	0.566***	0.624	-1.009	-1.570	-0.0204	-0.964
	(0.187)	(0.426)	(0.766)	(1.245)	(0.0443)	(0.794)
	[0.203]	[0.571]	[0.652]	[1.78]	[0.076]	[1.772]
Effect-High Productivity	0.518***	1.726**	4.671**	-0.165	0.273**	8.277***
SE	(0.156)	(0.592)	(1.446)	(4.303)	(0.123)	(3.137)
SE (bootstrap)	[0.183]	[0.58]	[1.691]	[3.915]	[0.121]	[2.677]
Baseline mean (DV)	11.75	19149.5	31903.0	92164.4	3362.5	92236.6
Observations	2190	2190	2190	2190	2190	2190
Number of households	229	229	229	229	229	229
R-Squared	0.595	0.522	0.391	0.548	0.525	0.645

Panel B: Fixed-effects approach

	(1) VF credit	(2) Total credit	(3) Profits	(4) Non-wage Expenses	(5) Wage Expenses	(6) Assets
Post X Inv HH X High Productivity	-0.407***	0.328	4.042**	0.396	0.350**	6.046**
	(0.141)	(0.632)	(1.602)	(4.983)	(0.135)	(2.561)
	[0.17]	[0.689]	[1.496]	[3.973]	[0.152]	[3.405]
Post X Inv HH	0.685***	0.930***	-0.294	-1.224	-0.0328	0.454
	(0.152)	(0.287)	(0.760)	(0.904)	(0.0442)	(1.698)
	[0.232]	[0.403]	[0.692]	[1.577]	[0.071]	[1.544]
Effect High Productivity	0.278	1.257	3.748	-0.829	0.318	6.500
SE	(0.229)	(0.801)	(1.567)	(4.847)	(0.132)	(2.491)
SE (bootstrap)	[0.199]	[0.785]	[1.296]	[3.835]	[0.138]	[3.377]
Observations	2190	2190	2190	2190	2190	2190
Number of households	229	229	229	229	229	229
R-Squared	0.597	0.523	0.386	0.547	0.523	0.640

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

Note: The table reports the reduced-form estimates of the effects of the program as a function of productivity estimated through the proxy-variable approach (Panel A) and the fixed-effects approach (Panel B). All dependent variables are winsorized with respect to the top 1% of the full sample distribution. The coefficients have been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome. Each column reports coefficients from regressions following the specification on equation 4. Post X Inv HH captures the reduced-form effect of the program for households belonging to the bottom-two terciles of the baseline productivity distribution. Post X Inv HH X High Productivity report the difference in the reduced form effects of the program between households in the top-third and bottom-two-thirds of the baseline productivity distribution. Reduced-form effects for high-productivity households are computed by adding up the coefficients on Post X Inv HH and Post X Inv HH X High Productivity. Standard errors, in parentheses, are clustered at the village level (64 clusters) to account for the quasi-experimental design. PBootstrapped standard errors are also presented in brackets and are computed using 500 bootstrap samples and clustered at the village level. The estimating sample only includes households who reported owning non-agricultural business assets the year preceding the rollout of the program. All dependent variables are measured in THB. Exchange rate THB/USD (2001) : 44.51.

Table 6: IV effects of total credit on income and profits (all households)

Panel A: IV estimates of the effect of credit on household income (in THB)

	(1)	(2)	(3)	(4)	(5)	(6)
	Household Income		Wage Income		Total Profits	
	Raw	Winsorized	Raw	Winsorized	Raw	Winsorized
Total Short Term Credit * High Productivity	4.317***	2.708***	0.336	0.248	3.346***	1.784***
	(1.768)	(1.076)	(0.432)	(0.403)	(1.301)	(0.595)
	[1.12]	[0.69]	[0.431]	[0.43]	[0.956]	[0.508]
Total Short Term Credit	-1.449*	-1.238*	0.202	0.193	-0.802	-0.748
	(0.833)	(0.683)	(0.264)	(0.254)	(0.688)	(0.464)
	[0.79]	[0.562]	[0.352]	[0.349]	[0.626]	[0.372]
Effect- High Productivity	2.868***	1.47***	0.54	0.44	2.544***	1.04
SE	(1.26)	(0.33)	(1.02)	(0.53)	(0.16)	(0.86)
SE bootstrap	[0.79]	[0.24]	[0.81]	[0.34]	[0.14]	[0.69]

Panel B: IV estimates of the effect of credit on Profits by source (in THB)

	(1)	(2)	(3)	(4)	(5)	(6)
	Farm		Fishing/Shrimping		Off-farm Business	
	Raw	Winsorized	Raw	Winsorized	Raw	Winsorized
Total Short Term Credit * High Productivity	0.519	0.215	-0.119	0.0210	2.946***	1.548***
	(0.330)	(0.250)	(0.117)	(0.0322)	(1.229)	(0.575)
	[0.45]	[0.34]	[0.094]	[0.057]	[0.999]	[0.44]
Total Short Term Credit	-0.0445	-0.208	0.00790	-0.0179	-0.766	-0.522
	(0.414)	(0.239)	(0.108)	(0.0682)	(0.568)	(0.404)
	[0.437]	[0.339]	[0.085]	[0.055]	[0.597]	[0.247]
Effect- High Productivity	0.47	0.01	-0.11	0.00	2.18***	1.025***
SE	(0.53)	(0.16)	(0.86)	(0.7)	(0.27)	(0.36)
SE bootstrap	[0.34]	[0.14]	[0.69]	[0.45]	[0.23]	[0.36]
First-stage F-stat: Short Term Credit	10.68					
First-Stage F-stat: Interaction	11.92					
Observations	6117	6117	6117	6117	6117	6117
Number of households	911	911	911	911	911	911

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

Note: The table reports instrumental-variables estimates of the effects of total short-term credit as a function of productivity estimated through the proxy-variable approach following the specification in equation 10. Panel A presents the effects of total short-term credit on income by source and Panel B presents effects of total short-term credit on profits, by type of activity. Odd-numbered columns report IV coefficients after truncating the dependent variable at the top 1%. "Total Short Term Credit" denotes the effects of short-term credit on the outcome of interest for households from bottom two thirds of the baseline productivity distribution. "Total Short Term Credit X High Productivity" denotes differences in the effects of short-term credit between high and low productivity households. "Effect- High Productiviy" is computed by adding the coeffcients of "Total Short Term Credit"' and "Total Short Term Credit X High Productivity ", and represents the effects of short-term credit for households from top-third of the baseline productivity distribution. Standard errors, in parentheses, are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are presented in brackets and are computed using 500 bootstrap samples and clustered at the village level. Short-term credit involves program loans with a term shorter than a year and has been top coded with respect to the 99th percentile for precision. Household profits include farm, fishing and shrimping and off-farm business profits. All dependent variables are measured in THB. Exchange rate THB/USD (2001) : 44.51.

Table 7: IV effect of total credit on preexisting non-agricultural businesses

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Profits		Non-wage Expenses		Wage Expenses		Assets	
	Raw	Winsorized	Raw	Winsorized	Raw	Winsorized	Raw	Winsorized
Total Short Term Credit * High Productivity	8.525**	6.596**	0.819	2.317	0.219	0.287**	6.692	5.760*
	(3.625)	(2.596)	(2.918)	(2.417)	(0.301)	(0.134)	(5.051)	(2.968)
	[3.844]	[2.061]	[3.136]	[2.543]	[0.315]	[0.133]	[5.478]	[2.862]
Total Short Term Credit	-2.696	-3.621*	-2.916	-2.962	-0.0742	-0.106	1.501	-1.079
	(2.810)	(2.175)	(3.729)	(2.994)	(0.252)	(0.113)	(2.409)	(1.347)
	[2.744]	[1.483]	[3.202]	[2.495]	[0.295]	[0.114]	[3.02]	[2.027]
Effect- High Productivity	5.829**	2.975**	-2.10	-0.65	0.14	0.181**	8.19	4.681**
SE	(1.45)	(0.88)	(2.23)	(1.9)	(0.18)	(0.06)	(5.25)	(2.44)
SE bootstrap	[2.07]	[1.16]	[2.51]	[1.999]	[0.21]	[0.07]	[5.09]	[1.95]
First-stage F-stat: Short-term credit	4.74						0	
First-Stage F-stat: Interaction	9.05							
Observations	1544	1544	1544	1544	1544	1544	1544	1544
Number of households	227	227	227	227	227	227	227	227

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

Note: The table reports instrumental-variables estimates of the effects of total short-term credit as a function of productivity estimated through the proxy-variable approach following the specification in equation 10. The estimating sample includes households with preexisting non-agricultural businesses. Odd-numbered columns report IV coefficients after truncating the dependent variable at the top 1%. 'Total Short Term Credit" denotes the effects of short-term credit on the outcome of interest for households from bottom two thirds of the baseline productivity distribution. "Total Short Term Credit X High Productivity" denotes differences in the effects of short-term credit between high and low productivity households. "Effect-High Productiviy" is computed by adding the coeffcients of "Total Short Term Credit"' and "Total Short Term Credit X High Productivity ", and represents the effects of short-term credit for households from top-third of the baseline productivity distribution. Standard errors, in parentheses, are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrap standard errors are also reported in brackets and are computed using 500 bootstrap samples and clustered at the village level. Short-term credit involves program loans with a term shorter than a year. Business ownership is defined as owning non-agricultural business assets the year preceding the rollout of the program. All dependent variables are measured in THB. Exchange rate THB/USD (2001) : 44.51.

Table 8: Gains in returns to credit from allocating resources to high-TFP entrepreneurs

Panel A: All businesses

	Returns to credit (THB)		
	Actual (A)	Counterfactual (C)	Gains ($\frac{C-A}{A}$) %
Mean	0.14	0.67	368%
Median	0.21	0.77	256%
Dispersion			
SD	1.07	0.87	-18%
75-25	1.90	1.29	-32%
90-10	2.05	1.86	-9%

Panel B: Incumbent non-agricultural businesses

	Returns to credit (THB)		
	Actual (A)	Counterfactual (C)	Gains ($\frac{C-A}{A}$) %
Mean	0.50	2.51	403%
Median	0.84	3.04	263%
Dispersion			
SD	4.82	3.80	-21%
75-25	7.96	5.80	-27%
90-10	13.68	10.20	-25%

Note: The Table reports summary statistics of returns to credit for actual and counterfactual program borrowers, and gains from reallocating resources based on within village TFP rankings. Panel A reports results for the full sample, while Panel B reports results for the sub-sample of non-agricultural businesses that operated before the rollout of the Million Baht Program.

Table 9: Village-level gains from program credit reallocation

Panel A: Within sectors and villages

	With crowd in		Without crowd in	
	Output	Capital	Output	Capital
	(1)	(2)	(3)	(4)
Average gain	18.7%	14.4%	10.6%	3.6%
By terciles of village size				
Bottom	29.6%	20.2%	16.1%	6.3%
Medium	14.2%	13.0%	8.6%	1.8%
Top	10.9%	9.5%	6.5%	2.4%
Average gain (weighted by village size)	3.8%	3.7%	0.6%	-2.6%

Panel B: Across sectors but within villages

	With crowd in		Without crowd in	
	Output	Capital	Output	Capital
Average gain	28.4%	21.2%	14.2%	5.3%
By terciles of village size				
Bottom	39.7%	29.1%	18.7%	8.6%
Medium	22.7%	16.4%	14.5%	3.0%
Top	21.6%	17.0%	9.0%	3.9%
Average gain (weighted by village size)	10.3%	7.8%	2.9%	-1.6%

Panel C: Across sectors and villages

	With crowd in		Without crowd in	
	Output	Capital	Output	Capital
Average gain	35.8%	28.1%	16.1%	7.2%
By terciles of village size				
Bottom	28.9%	19.0%	21.6%	7.2%
Medium	33.9%	25.7%	15.2%	5.9%
Top	45.3%	40.3%	10.8%	8.5%
Average gain (weighted by village size)	51.3%	47.4%	6.6%	7.3%

Note: The table reports statistics corresponding to percentage changes in village-level output and capital due to reallocation by terciles of village size. Panel A reports gains from reallocating program credit across firms in each sector and in each village with respect to the observed values. Panel B reports gains from reallocating program credit across all firms in each village with respect to the observed values. Panel C reports gains from reallocating program credit across all firms in the sample with respect to the observed values. Gains from reallocation: $X_v^C/X_v - 1$ where X_v^C denotes village-level variables under each counter-factual scenario, and X_v corresponds to the observed variable of interest. Statistics are performed after trimming the top and bottom 5% values of village-level actual and counterfactual values.

APPENDIX

A Supportive evidence and robustness checks

Table A1: Correlation between beliefs and value-added

	(1)	(2)	(3)	(4)
VARIABLES		log Value Added		
Log beliefs	0.0530***	0.0296**	0.0512***	0.0334**
	(0.0170)	(0.0143)	(0.0167)	(0.0150)
Observations	1,915	1,911	1,915	1,911
R-squared	0.010	0.137	0.155	0.240
Control for capital	No	Yes	No	Yes
Village F.E.	No	No	Yes	Yes

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

Note: The table presents estimates of a regression of value-added in logs on household beliefs regarding current profits for several specifications. Standard errors are clustered at the household level to account for serial correlation. Beliefs are measured as the self-reported household projected income for period t predicted in $t-1$.

Table A2: Correlation between within-village productivity rankings

	(1)	**(2)**
	Percentile ranking - Proxy-variable Method	
Percentile ranking - Fixed Effects Method	0.470***	0.499***
	(0.0315)	(0.0341)
Constant	0.331***	0.316***
	(0.0187)	(0.0165)
Observations	821	821
R-squared	0.208	0.266
Village FE	NO	YES

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

Note: The table presents correlations between the within-village productivity rankings obtained by the proxy-variable method and the fixed effects method. Standard errors are clustered at the village level.

Table A3: Correlates of baseline productivity and demographic characteristics

	Proxy-variable			Fixed-effects		
	(1)	(2)	(3)	(4)	(5)	(6)
	Productivity	Productivity Rank	High-productivity	Productivity	Productivity Rank	High-productivity
Number of Adult Males in Household	0.0626**	0.0199*	0.0159	0.158***	0.0335***	0.0516***
	(0.0305)	(0.0104)	(0.0185)	(0.0417)	(0.0111)	(0.0179)
Number of Adult Females in Household	0.259***	0.0923***	0.126***	0.271***	0.0432***	0.0673***
	(0.0333)	(0.0118)	(0.0235)	(0.0505)	(0.0129)	(0.0219)
Number of Children in Household	-0.0638***	-0.00754	0.00103	0.0295	0.00666	-0.00554
	(0.0239)	(0.00795)	(0.0136)	(0.0285)	(0.00743)	(0.0129)
Dummy: Male Head of Household	0.146*	0.0438	0.0515	0.188*	0.0448	0.0394
	(0.0742)	(0.0270)	(0.0426)	(0.106)	(0.0280)	(0.0425)
Head's main occupation: Farm (agriculture/livestock)	0.254***	0.0447*	0.0965**	0.371***	0.0572**	0.136***
	(0.0732)	(0.0243)	(0.0408)	(0.123)	(0.0260)	(0.0471)
Number of Businessowners in Household	0.340***	0.0591***	0.0588*	0.370***	0.0618***	0.0718**
	(0.0637)	(0.0215)	(0.0334)	(0.100)	(0.0197)	(0.0322)
Age of Head of Household	-0.00481*	0.0000670	0.00229*	-0.0166***	-0.00253***	-0.00334**
	(0.00245)	(0.000820)	(0.00136)	(0.00320)	(0.000716)	(0.00136)
Years of schooling - HH head	0.0253**	0.0122***	0.0163**	0.0357**	0.00649*	0.00961*
	(0.00975)	(0.00338)	(0.00622)	(0.0149)	(0.00327)	(0.00516)
Observations	886	886	886	811	811	811
R-Squared	0.257	0.134	0.0845	0.218	0.127	0.0893

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

Note: The table reports within village correlations between estimates of productivity and demographic characteristics. All regressions include village fixed-effects. Standard errors are clustered at the village level.

Table A4: Correlates of average borrower productivity and village size

	DV: Avg TFP for program borrowers			
	Proxy-variable method		Fixed-effects method	
	(1)	(2)	(3)	(4)
Village Size	0.0000		0.0001	
	-0.0001		-0.0001	
Log village size		-0.0367		0.00228
		(0.0627)		(0.0918)
Province FE	Yes	Yes	Yes	Yes
Observations	64	64	64	64
R-squared	0.511	0.514	0.189	0.185

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

Note: The table reports regressions of borrower productivity on village size. Standard errors are clustered at the village level.

Table A5: Reduced-form effects of the program on input spending

	Wage Exp. (1)	Non-Wage Exp. (2)	Assets (value) (3)
Post X Inv HH X High Productivity	-0.156	-1.091	-0.468
	(0.179)	(2.037)	(21.54)
	[0.202]	[1.344]	[13.755]
Post X Inv HH	0.0582	-1.072*	-1.791
	(0.0400)	(0.638)	(3.996)
	[0.045]	[0.691]	[7.023]
Effect-High Productivity	-0.0981	-2.162	-2.259
SE	(0.17)	(2.02)	(23.41)
SE (bootstrap)	[0.2]	[1.61]	[18.11]
Observations	8659	8659	8659
Number of households	922	922	922
R-Squared	0.576	0.598	0.673

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

Note: The table reports the reduced-form estimates of the effects of the program on input spending as a function of productivity estimated through the proxy-variable approach (Panel A) and the fixed-effects approach (Panel B). Column (1) reports effects on food spending, Column (2) presents effects on spending on durables (vehicle and house repairs). Column (3) presents effects on total consumption. Each column reports coefficients from regressions following the specification on equation 4. Post X Inv HH captures the reduced-form effect of the program for households belonging to the bottom-two terciles of the baseline productivity distribution. Post X Inv HH X High Productivity report the difference in the reduced form effects of the program between households in the top-third and bottom-two-thirds of the baseline productivity distribution. Reduced-form effects for high-productivity households are computed by adding up the coefficients on Post X Inv HH and Post X Inv HH X High Productivity. Standard errors, in parentheses, are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are also presented in brackets and are computed using 500 bootstrap samples and clustered at the village level. Spending includes operations corresponding to all business activities. All dependent variables are measured in THB. Exchange rate THB/USD (2001) : 44.51.

Table A6: Reduced-form effects of the program on consumption.

Panel A: Proxy-variable approach

	(1) Food	(2) Durables	(3) Total
Post X Inv HH X High Productivity (δ_2)	0.0528	0.137	0.345
	(0.0654)	(0.152)	(0.722)
	[0.071]	[0.147]	[0.659]
Post X Inv HH (δ_1)	0.0324	0.100	0.601*
	(0.0441)	(0.105)	(0.340)
	[0.049]	[0.121]	[0.429]
RF effect - High Productivity ($\delta_1 + \delta_2$)	0.0852	0.237	0.946
SE	(0.0543)	(0.181)	(0.845)
SE (bootstrap)	[0.059]	[0.158]	[0.729]
Observations	8659	8659	8619
Number of households	922	922	922
R-Squared	0.528	0.221	0.569

Panel B: Fixed-effects approach

	(1) Food	(2) Durables	(3) Total
Post X Inv HH X High Productivity (δ_2)	0.0743	0.259	1.105
	(0.0613)	(0.216)	(0.765)
	[0.062]	[0.228]	[0.869]
Post X Inv HH (δ_1)	0.0310	0.0712	0.347
	(0.0448)	(0.144)	(0.402)
	[0.046]	[0.075]	[0.41]
RF effect - High Productivity ($\delta_1 + \delta_2$)	0.105	0.331	1.453*
SE	(0.0490)	(0.178)	(0.866)
SE (bootstrap)	[0.063]	[0.244]	[0.88]
Observations	8659	8659	8619
Number of households	922	922	922
R-Squared	0.528	0.222	0.570

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

Note: The table reports the reduced-form estimates of the effects of the program on food consumption as a function of productivity estimated through the proxy-variable approach (Panel A) and the fixed-effects approach (Panel B). Column (1) reports effects on food spending, Column (2) presents effects on spending on durables (vehicle and house repairs). Column (3) presents effects on total consumption. Each column reports coefficients from regressions following the specification on equation 4. Post X Inv HH captures the reduced-form effect of the program for households belonging to the bottom-two terciles of the baseline productivity distribution. Post X Inv HH X High Productivity report the difference in the reduced form effects of the program between households in the top-third and bottom-two-thirds of the baseline productivity distribution. Reduced-form effects for high-productivity households are computed by adding up the coefficients on Post X Inv HH and Post X Inv HH X High Productivity. Standard errors, in parentheses, are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are also presented in brackets and are computed using 500 bootstrap samples and clustered at the village level. All dependent variables are measured in THB. Exchange rate THB/USD (2001) : 44.51.

Table A7: Reduced-form effects of the program on business creation.

Panel A: Proxy-variable approach				
	(1) # of Non-ag Biz.	(2) New Non-ag Biz.	(3) # of Farm Biz.	(4) New Farm Biz.
Post X Inv HH X High Productivity (δ_2)	2.908	1.440	0.652	-0.0252
	(3.587)	(1.779)	(1.173)	(0.354)
	[3.461]	[1.554]	[1.423]	[0.458]
Post X Inv HH (δ_1)	0.727	0.0641	-0.522	-0.105
	(3.533)	(1.395)	(0.946)	(0.217)
	[4.105]	[1.209]	[1.007]	[0.314]
RF effect - High Productivity ($\delta_1 + \delta_2$)	3.636	1.504	0.129	-0.130
SE	(5.472)	(1.325)	(1.236)	(0.294)
SE (bootstrap)	[4.587]	[1.507]	[1.399]	[0.284]
Observations	8658	8658	8658	8658
Number of households	922	922	922	922
R-Squared	0.953	0.132	0.557	0.137

Panel B: Fixed-effects approach				
	(1) # of Non-ag Biz.	(2) New Non-ag Biz.	(3) # of Farm Biz.	(4) New Farm Biz.
Post X Inv HH X High Productivity (δ_2)	6.655	1.912	2.181	0.804**
	(6.997)	(1.618)	(1.329)	(0.361)
	[6.639]	[1.911]	[2.262]	[0.49]
Post X Inv HH (δ_1)	1.209	0.173	-0.960	-0.353
	(3.123)	(1.220)	(0.986)	(0.248)
	[2.842]	[1.131]	[0.841]	[0.243]
RF effect - High Productivity ($\delta_1 + \delta_2$)	7.864	2.085	1.221	0.451
SE	(8.214)	(1.472)	(1.326)	(0.259)
SE (bootstrap)	[6.841]	[1.732]	[2.248]	[0.424]
Observations	8658	8658	8658	8658
Number of households	922	922	922	922
R-Squared	0.953	0.132	0.137	0.557

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

Note: The table reports the reduced-form estimates of the effects of the program on business creation as a function of productivity estimated through the proxy-variable approach (Panel A) and the fixed-effects approach (Panel B). Each column reports coefficients from regressions following the specification on equation 4. Post X Inv HH captures the reduced-form effect of the program for households belonging to the bottom-two terciles of the baseline productivity distribution. Post X Inv HH X High Productivity report the difference in the reduced form effects of the program between households in the top-third and bottom-two-thirds of the baseline productivity distribution. Reduced-form effects for high-productivity households are computed by adding up the coefficients on Post X Inv HH and Post X Inv HH X High Productivity. Standard errors, in parentheses, are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are also presented in brackets and are computed using 500 bootstrap samples and clustered at the village level.

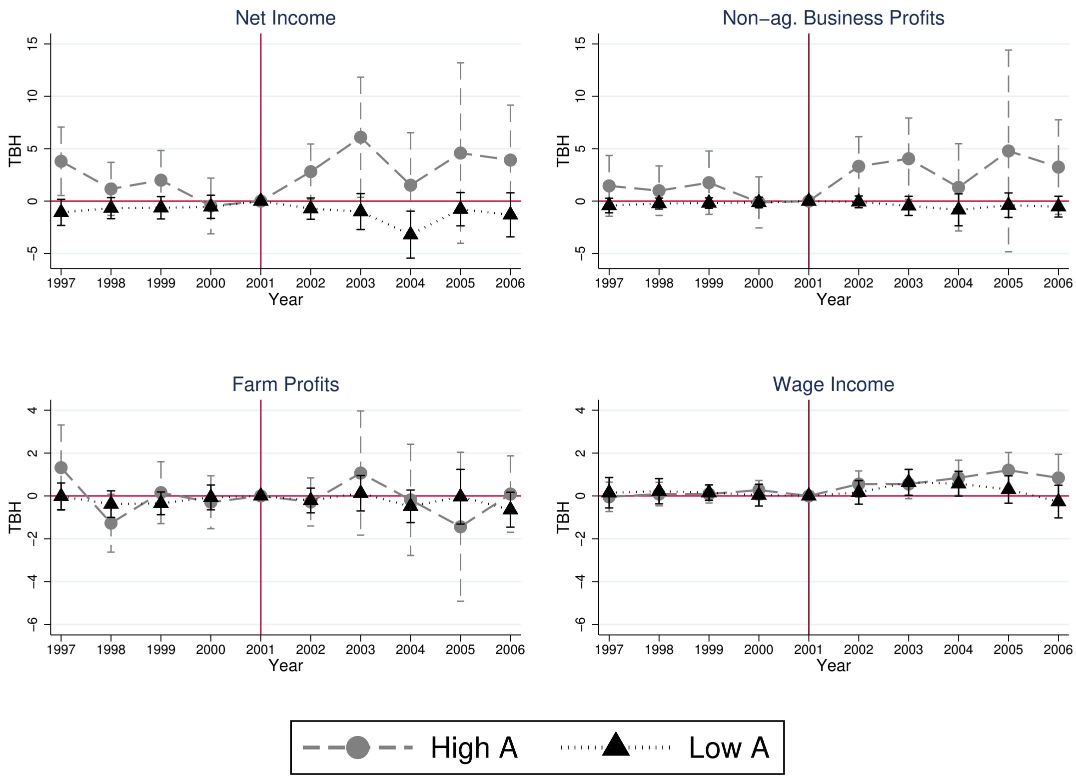


Figure A1: Effects of program rollout on household income - Fixed-effects approach

Note: The figure depicts flexible estimates corresponding to the specification in (5). Each dot represents differences in income between households from villages with high and low per-capita program funds, for each year with respect to the year preceding the announcement of the program (2001). Each coefficient has been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome. High A: household belongs to the top-third tercile of the baseline productivity distribution in each village. Low A: household belongs to the bottom two-thirds of the baseline productivity distribution in each village. Productivity estimates correspond to the control function approach using household beliefs about profits as a proxy variable. 95 % confidence intervals are computed based on standard errors, which are clustered at the village level to account for the empirical design.

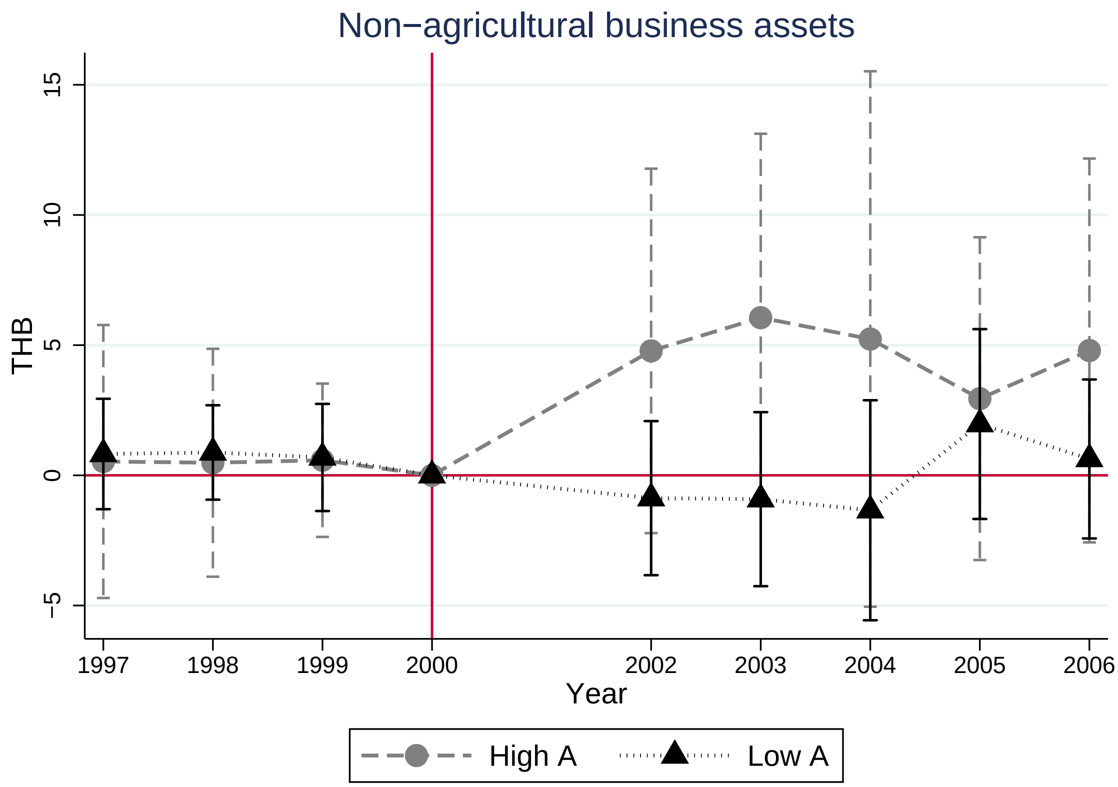


Figure A2: Reduced-form effects on assets (preexisting non-agricultural businesses)

Note: The figure depicts flexible estimates corresponding to the specification in (5). Each dot represents differences in program borrowing between households from villages with high and low per-capita program funds, for each year with respect to 2000. Observations corresponding to the 2001 wave were dropped from the estimating sample. Each coefficient has been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the value of off-farm business assets (measured in baht). The effects are estimated over a sample of 230 households who reported holding business assets during the year preceding the rollout of the program. High A: household belongs to the top-third of the baseline productivity distribution in each village. Low A: household belongs to the bottom two-thirds of the baseline productivity distribution in each village. Productivity estimates correspond to the control function approach using household beliefs about profits as a proxy variable. 95 % confidence intervals are computed based on standard errors, which are clustered at the village level to account for the empirical design. The dependent variable is winsorized with respect to the top 1% of the distribution.

Table A8: Reduced-form effects of the program on income and profits- Without winsorizing

	(1) VF short-term Credit	(2) Short-term Credit	(3) Household Income	(4) Wage Income	(5) Profits	(6) Farm Profits	(7) Shrimp/Fish Profits	(8) Business Profits
Panel A: Proxy-variable approach								
Post X Inv HH X High Productivity	0.0314	0.643	3.309**	0.471	3.286**	0.579	-0.109	2.816**
	(0.106)	(0.383)	(1.642)	(0.393)	(1.274)	(0.437)	(0.133)	(1.225)
	[0.082]	[0.445]	[1.499]	[0.464]	[1.387]	[0.435]	[0.162]	[1.283]
Post X Inv HH	0.612***	0.781**	-0.830	0.106	-0.601	-0.227	-0.0391	-0.335
	(0.124)	(0.282)	(0.541)	(0.209)	(0.465)	(0.287)	(0.0845)	(0.431)
	[0.108]	[0.365]	[0.526]	[0.229]	[0.483]	[0.281]	[0.081]	[0.412]
Effect for High Productivity	0.644***	1.424**	2.478*	0.577	2.685*	0.352	-0.148	2.481*
SE	(0.067)	(0.485)	(1.421)	(0.329)	(1.301)	(0.459)	(0.189)	(1.122)
SE (bootstrap)	[0.08]	[0.549]	[1.349]	[0.378]	[1.387]	[0.475]	[0.23]	[1.183]
Observations	8659	8659	8659	8659	8659	8659	8659	8659
Number of households	922	922	922	922	922	922	922	922
R-Squared	0.586	0.383	0.516	0.729	0.479	0.252	0.301	0.425
Panel B: Fixed-effects approach								
	(1) VF short-term Credit	(2) Short-term Credit	(3) Household Income	(4) Wage Income	(5) Profits	(6) Farm Profits	(7) Shrimp/Fish Profits	(8) Business Profits
Post X Inv HH X High Productivity	-0.239***	0.0243	3.282	0.568	2.678	0.328	-0.172	2.521
	(0.0697)	(0.509)	(1.782)	(0.381)	(1.528)	(0.607)	(0.205)	(1.528)
	[0.11]	[0.837]	[2.195]	[0.385]	[1.882]	[0.767]	[0.412]	[1.788]
Post X Inv HH	0.696***	0.983***	-0.555	0.114	-0.149	-0.0676	-0.0147	-0.0672
	(0.0827)	(0.228)	(0.426)	(0.227)	(0.382)	(0.256)	(0.0691)	(0.341)
	[0.096]	[0.282]	[0.462]	[0.224]	[0.333]	[0.272]	[0.028]	[0.239]
Effect for High Productivity	0.457***	1.007	2.727	0.682**	2.529	0.261	-0.186	2.454
SE	(0.101)	(0.624)	(1.683)	(0.3)	(1.594)	(0.641)	(0.238)	(1.462)
SE (bootstrap)	[0.107]	[0.921]	[2.035]	[0.342]	[1.932]	[0.763]	[0.421]	[1.763]
Observations	8659	8659	8659	8659	8659	8659	8659	8659
Number of households	922	922	922	922	922	922	922	922
R-Squared	0.588	0.387	0.516	0.729	0.479	0.252	0.302	0.426

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

Note: The table reports the reduced-form estimates of the effects of the program as a function of productivity estimated through the proxy-variable approach (Panel A) and the fixed-effects approach (Panel B). Reduced-form effects for high-productivity households are computed by adding up the coefficients on Post X Inv HH and Post X Inv HH X High Productivity. The coefficients have been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome. Standard errors are clustered at the village level (64 clusters) to account for the quasi-experimental design. Panel A also presents bootstrap standard errors in brackets, which are computed using 500 bootstrap samples and clustered at the village level.

Table A9: Reduced-form effects of the program on non-agricultural preexisting businesses- Without winsorizing

Panel A: Proxy-variable approach

	(1) VF short-term credit	(2) Total short-term credit	(3) Profits	(4) Non-wage Expenses	(5) Wage Expenses	(6) Assets
Post X Inv HH X High Productivity	-0.00998	2.148*	11.21**	1.138	0.480	16.35*
	(0.167)	(0.990)	(4.401)	(5.938)	(0.352)	(8.814)
	[0.16]	[1.055]	[4.559]	[5.079]	[0.332]	[9.032]
Post X Inv HH	0.577***	0.606	-0.487	-3.751	-0.0840	-0.304
	(0.194)	(0.516)	(1.403)	(2.424)	(0.120)	(1.144)
	[0.21]	[0.712]	[1.984]	[2.631]	[0.202]	[2.572]
Effect-High Productivity	0.567***	2.754**	10.72**	-2.613	0.396	16.04*
SE	(0.135)	(1.154)	(3.937)	(5.335)	(0.368)	(9.403)
SE (bootstrap)	[0.17]	[1.172]	[3.898]	[4.816]	[0.35]	[9.183]
Baseline mean (DV)	12	21237	50974	109802	7292	119204
Observations	2190	2190	2190	2190	2190	2190
Number of households	229	229	229	229	229	229
R-Squared	0.590	0.509	0.439	0.468	0.470	0.589

Panel B: Fixed-effects approach

	(1) VF short-term credit	(2) Total short-term credit	(3) Profits	(4) Non-wage Expenses	(5) Wage Expenses	(6) Assets
Post X Inv HH X High Productivity	-0.434***	1.211	11.61***	-3.524	0.409	13.52
	(0.145)	(1.152)	(3.689)	(6.485)	(0.490)	(9.362)
	[0.177]	[1.106]	[3.615]	[5.712]	[0.607]	[9.075]
Post X Inv HH	0.719***	1.012***	-0.159	-2.055	-0.031	1.161
	(0.144)	(0.364)	(1.103)	(1.343)	(0.108)	(2.308)
	[0.227]	[0.514]	[1.497]	[2.392]	[0.23]	[2.293]
Effect High Productivity	0.285	2.223	11.45	-5.578	0.379	14.68
SE	(0.231)	(1.394)	(3.812)	(6.237)	(0.462)	(9.6)
SE (bootstrap)	[0.202]	[1.34]	[3.804]	[5.146]	[0.515]	[9.331]
Observations	2190	2190	2190	2190	2190	2190
Number of households	229	229	229	229	229	229
R-Squared	0.592	0.511	0.439	0.469	0.471	0.588

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

Note: The table reports the reduced-form estimates of the effects of the program as a function of productivity estimated through the proxy-variable approach (Panel A) and the fixed-effects approach (Panel B). Reduced-form effects for high-productivity households are computed by adding up the coefficients on Post X Inv HH and Post X Inv HH X High Productivity. The coefficients have been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome. Standard errors are clustered at the village level (64 clusters) to account for the quasi-experimental design. Panel A also presents bootstrap standard errors in brackets, which are computed using 500 bootstrap samples and clustered at the village level. The estimating sample only includes households who reported owning business assets the year preceding the rollout of the program. The estimating sample only includes households who reported owning non-agricultural business assets the year preceding the rollout of the program.

Table A10: Reduced-form effects on selected outcomes- Excluding 2001

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Raw data				Truncated top 1%	
	VF short-term Credit	Short-term Credit	Household Income	Profits	Off-farm Business profits	Household Income	Profits	Off-farm Business profits
Post X Inv HH X High Productivity	0.0240	0.547	3.589**	3.358**	2.516*	2.389***	1.441*	0.827
	(0.102)	(0.423)	(1.653)	(1.466)	(1.280)	(0.921)	(0.587)	(0.452)
	[0.075]	[0.516]	[1.586]	[1.601]	[1.382]	[0.894]	[0.818]	[0.68]
Post X Inv HH	0.600***	0.813***	-0.849	-0.474	-0.212	-0.958**	-0.588**	-0.235
	(0.122)	(0.277)	(0.569)	(0.499)	(0.447)	(0.386)	(0.287)	(0.282)
	[0.106]	[0.361]	[0.559]	[0.556]	[0.456]	[0.393]	[0.287]	[0.173]
Effect for High Productivity	0.624***	1.360**	2.740*	2.884*	2.305*	1.430**	0.853	0.592
SE	(0.072)	(0.514)	(1.439)	(1.501)	(1.161)	(0.781)	(0.543)	(0.417)
SE (bootstrap)	[0.084]	[0.603]	[1.414]	[1.589]	[1.237]	[0.774]	[0.696]	[0.619]
Observations	7764	7764	7764	7764	7764	7764	7764	7764
Number of households	922	922	922	922	922	922	922	922
R-Squared	0.593	0.385	0.530	0.495	0.438	0.572	0.438	0.404

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

Note: The table presents reduced-form estimates of the program dropping 2001 from the estimating sample. Productivity is estimated through the proxy-variable approach. The coefficients have been scaled down by 1,000,000 in order to capture the effect of an additional per capita THB in each village on the corresponding outcome. Standard errors are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are reported in brackets and are computed using 500 bootstrap samples and clustered at the village level.

Table A11: Reduced-form effects on selected outcomes for households with preexisting non-agricultural businesses - Excluding 2001

	(1)	(2)	(3)	(4)	(5)	(6)
		Raw data			Truncated top 1 %	
	VF short-term credit	Total short-term credit	Profits	Assets	Profits	Assets
Post X Inv HH X High Productivity	-0.0321	2.184**	10.41**	19.36*	5.349**	10.98***
	(0.155)	(0.953)	(4.646)	(10.00)	(2.015)	(3.087)
	[0.152]	[1.038]	[4.922]	[10.095]	[2.122]	[3.078]
Post X Inv HH	0.580***	0.648	-0.168	-0.789	-0.928	-1.428
	(0.182)	(0.462)	(1.508)	(1.582)	(0.786)	(1.107)
	[0.206]	[0.63]	[2.175]	[3.169]	[0.734]	[2.285]
Effect-High Productivity	0.548***	2.832***	10.25***	18.57**	4.421**	9.553***
SE	(0.14)	(1.08)	(4.14)	(10.83)	(1.87)	(3.72)
SE (bootstrap)	[0.17]	[1.16]	[4.12]	[10.43]	[1.94]	[3.02]
Baseline mean (DV)	0	17644.6	65887.3	125723.4	31519.2	86039.0
Observations	1963	1963	1963	1963	1963	1963
Number of households	229	229	229	229	229	229
R-Squared	0.596	0.513	0.452	0.568	0.389	0.631

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

Note: The table presents reduced-form estimates of the program dropping 2001 from the estimating sample. Productivity is estimated through the proxy-variable approach. The coefficients have been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome. Standard errors are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are reported in brackets and are computed using 500 bootstrap samples and clustered at the village level. The estimating sample only includes households who reported owning non-agricultural business assets the year preceding the rollout of the program.

Table A12: Reduced-form effects on selected outcomes- Excluding villages with pre-program village funds

Panel A: Proxy-variable approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Raw data				Truncated top 1%	
	VF short-term Credit	Short-term Credit	Household Income	Profits	Off-farm Business profits	Household Income	Profits	Off-farm Business profits
Post X Inv HH X High Productivity	0.0524	0.920	1.641	3.047*	2.592	1.052	1.550***	1.142**
	(0.137)	(0.582)	(2.082)	(1.757)	(1.726)	(1.172)	(0.551)	(0.541)
	[0.122]	[0.704]	[2.027]	[1.869]	[1.861]	[1.01]	[0.707]	[0.677]
Post X Inv HH	0.651***	1.018***	-0.0598	-0.299	-0.156	-0.445	-0.725**	-0.507*
	(0.169)	(0.369)	(0.602)	(0.632)	(0.542)	(0.426)	(0.336)	(0.302)
	[0.163]	[0.5]	[0.701]	[0.716]	[0.591]	[0.486]	[0.272]	[0.212]
Effect for High Productivity	0.704	1.938	1.581	2.748	2.436	0.606	0.826	0.635
SE	(0.082)	(0.728)	(2)	(1.961)	(1.675)	(1.084)	(0.486)	(0.447)
SE (bootstrap)	[0.113]	[0.885]	[1.991]	[2.159]	[1.843]	[0.989]	[0.725]	[0.659]
Observations	6838	6838	6838	6838	6838	6838	6838	6838
Number of households	729	729	729	729	729	729	729	729
R-Squared	0.576	0.457	0.445	0.406	0.367	0.564	0.466	0.415

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

Note: The table reports the reduced-form estimates of the effects of the program as a function of productivity estimated through the proxy-variable approach. The sample excludes the 10 villages with pre-program village funds. The coefficients have been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome. Standard errors are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are reported in brackets and are computed using 500 bootstrap samples and clustered at the village level.

Table A13: Reduced-form effects on selected outcomes for households with preexiting non-agricultural businesses - Excluding villages with pre-program village funds

	(1)	(2)	(3)	(4)	(5)	(6)
		Raw data			Truncated top 1 %	
	VF short-term credit	Total short-term credit	Profits	Assets	Profits	Assets
Post X Inv HH X High Productivity	0.0120	1.786	9.768*	21.53*	4.796***	11.09***
	(0.215)	(1.151)	(5.688)	(11.13)	(1.622)	(3.314)
	[0.242]	[1.495]	[6.596]	[12.875]	[2.355]	[3.628]
Post X Inv HH	0.540**	0.940	0.690	1.595	-1.402	0.209
	(0.235)	(0.779)	(2.114)	(1.359)	(0.989)	(0.688)
	[0.294]	[0.974]	[3.226]	[3.225]	[0.929]	[2.112]
Effect-High Productivity	0.55	2.73	10.46	23.12	3.39	11.30
SE	(0.17)	(1.54)	(5.54)	(11.85)	(1.26)	(3.57)
SE (bootstrap)	[0.27]	[1.8]	[6.12]	[12.98]	[2.15]	[3.45]
Baseline mean (DV)	0	23127	54689	123911	32260	91420
Observations	1679	1679	1679	1679	1679	1679
Number of households	175	175	175	175	175	175
R-Squared	0.570	0.509	0.372	0.583	0.388	0.632

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

Note: The table reports the reduced-form estimates of the effects of the program as a function of productivity estimated through the proxy-variable approach. The sample excludes the 10 villages with pre-program village funds. The coefficients have been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome. Standard errors are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are reported in brackets and are computed using 500 bootstrap samples and clustered at the village level. The estimating sample only includes households who reported owning business assets the year preceding the rollout of the program.

Table A14: Reduced-form effects on selected outcomes- Productivity ranking

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Raw data				Truncated top 1%	
	VF short-term Credit	Short-term Credit	Household Income	Profits	Off-farm Business profits	Household Income	Profits	Off-farm profits
Post X Inv HH X Productivity rank	-0.00354	1.017	4.422	5.058***	4.083*	3.134	2.745*	2.063
	(0.145)	(0.679)	(2.796)	(1.946)	(1.708)	(1.869)	(1.235)	(0.920)
	[0.148]	[0.801]	[3.02]	[2.633]	[2.433]	[1.967]	[1.596]	[1.325]
Post X Inv HH	0.625***	0.464	-1.795	-1.908	-1.313	-1.661	-1.526	-0.926
	(0.139)	(0.308)	(1.356)	(0.904)	(0.737)	(0.939)	(0.562)	(0.407)
	[0.142]	[0.354]	[1.297]	[1.031]	[0.971]	[0.932]	[0.689]	[0.528]
Observations	8578	8578	8578	8578	8578	8578	8578	8578
Number of households	910	910	910	910	910	910	910	910
R-Squared	0.587	0.383	0.517	0.479	0.425	0.577	0.444	0.405

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

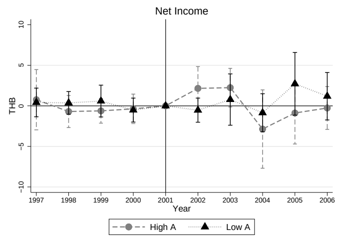
Note: The table reports the reduced-form estimates of the effects of the program as a function of productivity estimated through the proxy-variable approach. The coefficients have been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome. Standard errors are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are reported in brackets and are computed using 500 bootstrap samples and clustered at the village level.

Table A15: Reduced-form effects on selected outcomes for households with preexisting non-agricultural businesses- Productivity ranking

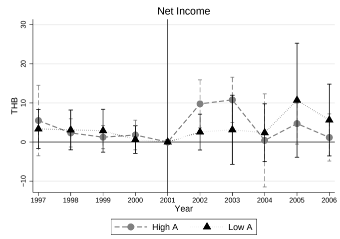
	(1)	(2)	(3)	(4)	(5)	(6)
		Raw data			Truncated top 1 %	
	VF short-term credit	Total short-term credit	Profits	Assets	Profits	Assets
Post X Inv HH X Productivity rank	-0.239	2.745	18.58***	27.83*	10.53***	15.55***
	(0.252)	(1.921)	(6.815)	(15.97)	(2.847)	(5.650)
	[0.278]	[1.908]	[8.376]	[15.381]	[3.768]	[5.007]
Post X Inv HH	0.693***	-0.288	-6.674*	-10.59*	-4.812***	-6.685**
	(0.161)	(0.955)	(3.382)	(6.224)	(1.299)	(3.002)
	[0.272]	[0.779]	[3.788]	[5.485]	[1.528]	[2.604]
Baseline mean (DV)	11.75	21237.0	50973.5	119203.8	31903.0	92236.6
Observations	2161	2161	2161	2161	2161	2161
Number of households	225	225	225	225	225	225
R-Squared	0.589	0.509	0.440	0.588	0.390	0.631

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

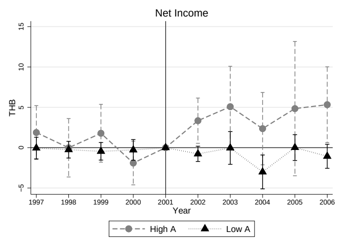
Note: The table reports the reduced-form estimates of the effects of the program as a function of productivity estimated through the proxy-variable approach. The coefficients have been scaled down by 1,000,000 in order to capture the effect of an additional per-capita THB in each village on the corresponding outcome. Standard errors are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are reported in brackets and are computed using 500 bootstrap samples and clustered at the village level. The estimating sample only includes households who reported owning business assets the year preceding the rollout of the program.



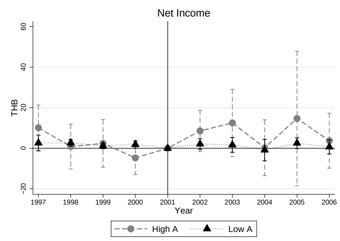
(a) Full sample - Percapita model



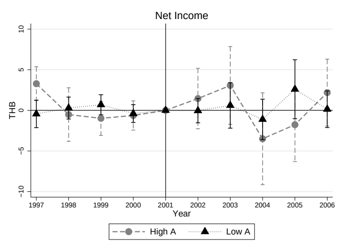
(b) Preexisting non-ag businesses - Per capita model



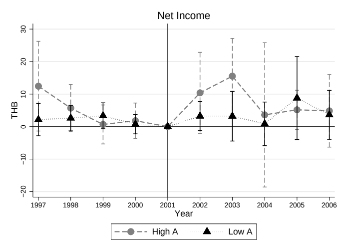
(c) Full sample - Including number of workers



(d) Preexisting non-ag businesses - Including number of workers



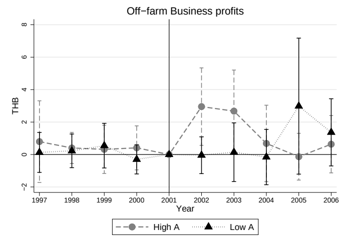
(e) Full sample - Including number of workers & M.E. in k



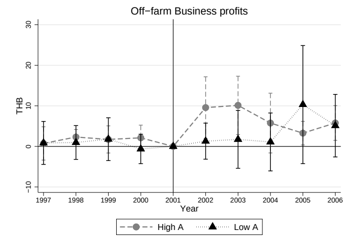
(f) Preexisting non-ag businesses - Including number of workers & M.E. in k

Figure A3: Effects of program rollout on household income by alternative measures of productivity

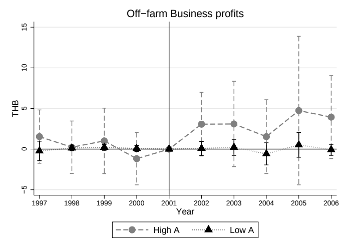
Note: The figure depicts flexible estimates corresponding to the specification in (5). Each dot represents differences in income between households from villages with high and low per capita program funds, for each year with respect to the year of the of the announcement of the program (2001). Each coefficient has been scaled down by 1,000,000 in order to capture the effect of an additional per capita THB in each village on the corresponding outcome. All dependent variables are winsorized with respect to the top 1% of the full sample distribution. High A: household belongs to the top-third 33% of the baseline productivity distribution in each village. Low A: household belongs to the bottom two-thirds of the baseline productivity distribution in each village. Productivity estimates correspond to the control function approach using household beliefs about profits as a proxy variable. 95% confidence intervals are computed based on standard errors, which are clustered at the village level to account for the empirical design.



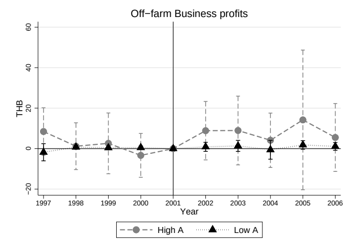
(a) Full sample - Per capita Model



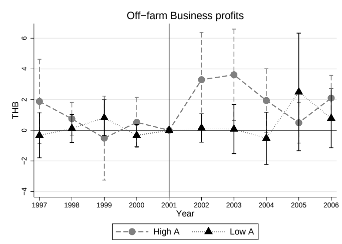
(b) Preexisting non-ag businesses - Per-capita Model



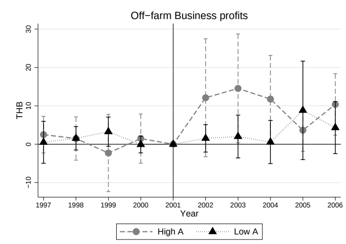
(c) Full sample - Including number of workers



(d) Preexisting non-ag businesses - Including number of workers



(e) Full sample - Including number of workers & M.E. in k



(f) Preexisting non-ag businesses - Including number of workers & M.E. in k

Figure A4: Effects of program rollout on business profits by alternative measures of productivity

Note: The figure depicts flexible estimates corresponding to the specification in (5). Each dot represents differences in income between households from villages with high and low per capita program funds, for each year with respect to the year in which the program was announced (2001). Each coefficient has been scaled down by 1,000,000 in order to capture the effect of an additional per capita THB in each village on the corresponding outcome. All dependent variables are winsorized with respect to the top 1% of the full sample distribution. High A: household belongs to the top-third of the baseline productivity distribution in each village. Low A: households belong to the bottom two-thirds of the baseline productivity distribution in each village. Productivity estimates correspond to the control function approach using household beliefs about profits as a proxy variable. 95 % confidence intervals are computed based on standard errors, which are clustered at the village level to account for the empirical design.

Table A16: IV effects of total credit on income and profits - 5 post-program years

Panel A: IV estimates of the effect of credit on household income (full sample)

	(1)	(2)	(3)	(4)	(5)	(6)
	Household Income		Wage Income		Total Profits	
	Raw	Winsorized	Raw	Winsorized	Raw	Winsorized
Total Short Term Credit * High Productivity	2.292*	1.658**	0.341	0.314	2.535**	1.211***
	(1.345)	(0.839)	(0.399)	(0.380)	(1.022)	(0.399)
	[1.239]	[0.727]	[0.404]	[0.406]	[1.062]	[0.580]
Total Short Term Credit	-0.932	-0.978**	0.147	0.123	-0.778	-0.825***
	(0.625)	(0.476)	(0.279)	(0.273)	(0.503)	(0.308)
	[0.879]	[0.564]	[0.357]	[0.352]	[0.738]	[0.416]
Effect- High Productivity	1.361*	0.681	0.488**	0.436	1.757**	0.386
SE	1.060	0.616	0.243	0.211	0.936	0.352
SE bootstrap	[0.778]	[0.455]	[0.245]	[0.226]	[0.777]	[0.423]

Panel B: IV estimates of the effect of credit on Profits by source (full sample)

	(1)	(2)	(3)	(4)	(5)	(6)
	Farm		Fishing/Shrimping		Off-farm Business	
	Raw	Winsorized	Raw	Winsorized	Raw	Winsorized
Total Short Term Credit * High Productivity	0.471	0.418	-0.0849	0.0239	2.149*	0.769**
	(0.359)	(0.301)	(0.0955)	(0.0291)	(1.076)	(0.387)
	[0.499]	[0.394]	[0.097]	[0.060]	[1.125]	[0.519]
Total Short Term Credit	-0.332	-0.495**	-0.0291	-0.0153	-0.417	-0.315
	(0.322)	(0.245)	(0.0836)	(0.0285)	(0.501)	(0.323)
	[0.494]	[0.408]	[0.090]	[0.059]	[0.759]	[0.285]
Effect- High Productivity	0.139	-0.0766	-0.114	0.00860	1.732**	0.454
SE	0.309	0.244	0.147	0.00694	0.869	0.283
SE bootstrap	[0.334]	[0.208]	[0.151]	[0.013]	[0.679]	[0.373]
First-stage F-stat: Short Term Credit	5.403					
First-Stage F-stat: Interaction	5.126					
Observations	8650					
Number of households	914					

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

Note: The table reports the instrumental-variables estimates of the effects of total short-term credit as a function of productivity estimated through the proxy-variable approach. The estimating sample includes 5 pre and post program years. Panel A presents the effects on income by source and Panel B presents effects on profits by type of activity. Odd-numbered columns report IV coefficients after truncating the dependent variable at the top 1%. Standard errors, in parentheses, are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrapped standard errors are presented in brackets and are computed using 500 bootstrap samples and clustered at the village level. Short-term credit involves program loans with a term shorter than a year and has been top coded with respect to the 99th percentile for precision. Household profits include farm, fishing and shrimping and off-farm business profits.

Table A17: IV effects of total credit on preexiting non-agricultural businesses - 5 post-program years

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Profits		Non-wage Expenses		Wage Expenses		Assets	
	Raw	Winsorized	Raw	Winsorized	Raw	Winsorized	Raw	Winsorized
Total Short Term Credit * High Productivity	6.331**	4.029***	3.648	1.239	0.38	0.152	7.009*	4.688***
	(2.263)	(1.187)	(3.044)	(2.209)	(0.196)	(0.0712)	(4.186)	(1.833)
	[2.847]	[1.579]	[3.377]	[2.364]	[0.279]	[0.103]	[3.858]	[1.744]
Total Short Term Credit	-2.174	-2.200**	-4.509	-1.012	-0.144	-0.0293	0.180	-0.941
	(1.857)	(1.062)	(3.199)	(1.817)	(0.203)	(0.0692)	(1.077)	(1.165)
	[2.534]	[1.391]	[3.054]	[1.699]	[0.292]	[0.094]	[2.106]	[1.566]
Effect- High Productivity	4.157***	1.830**	-0.861	0.227	0.236	0.123**	7.188**	3.747***
SE	1.502	0.559	2.250	1.777	0.146	0.0433	4.026	1.500
SE bootstrap	[1.326]	[0.802]	[2.537]	[1.764]	[0.175]	[0.057]	[3.117]	[0.991]
First-stage F-stat: Short-term credit	4.902							
First-Stage F-stat: Interaction	6.306							
Observations	2188							
Number of households	228							

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

Note: The table reports instrumental-variables estimates of the effects of total short-term credit as a function of productivity estimated through the proxy-variable approach. The estimating sample includes 5 pre and post program years. Odd-numbered columns report IV coefficients after truncating the dependent variable at the top 1%. Standard errors, in parentheses, are clustered at the village level (64 clusters) to account for the quasi-experimental design. Bootstrap standard errors are also reported in brackets and are computed using 500 bootstrap samples and clustered at the village level. Short-term credit involves program loans with a term shorter than a year. The estimating sample includes household with pre-existing businesses only. All variables are measured in THB.

ONLINE APPENDIX

B Variable definition

B.1 Productivity estimates

1. **Value added:** We compute value added as the difference between gross total revenues (across all household businesses) and total spending in wages (across all household businesses). We do not consider revenues derived from wage work provision of household members to other households, nor transportation spending on these activities.

2. **Capital:** Capital is measured as the self-reported total value of the stock of fixed assets for all household businesses. This excludes household assets such as appliances or other durable goods.

3. **Beliefs:** We construct our measure of beliefs as: $b = \frac{\text{Profits regular} - \text{Profits pessimistic}}{\text{Profits optimistic} - \text{Profits pessimistic}}$ We use households self-reported projections of total income for the next year (in THB). Households report forecasts in three scenarios: a pessimistic scenario, an optimistic scenario, and a regular scenario.

4. **Labor:** We approximate labor as the sum of the following components:

   - Total number of out-of-household workers hired for a wage. This only includes off-farm businesses.
   - Total number of out-of-household workers that provide unpaid labor. This only includes off-farm businesses.
   - Number of household members that report working in a family business as their main occupation. This includes all family enterprises.

B.2 Outcomes

1. **Short-term credit:** Total amount of credit, which is obtained by each household from any lender. This measure only includes loans with repayment periods under 12 months.

2. **Income:** We measure total household income as the sum of profits from household enterprises, net earnings (after taxes), donations/transfers from other households, and government transfers.

3. **Profits:** We measure profits as gross revenues (sales) net of operations costs such as wages (when available) and non-labor spending. We do not consider the shadow wage for household workers nor the costs associated to labor-sharing schemes. It excludes income derived from providing labor to other households.

4. **Labor costs (non-agricultural business only):** Total spending in wage workers.

5. **Consumption:** Total consumption expenditure. It is computed as the weighted sum of consumption expenditure across several categories of goods. The weights represent the relative weight of such categories in the Thai Socio Economic Surveys.

C Identification and Estimation details.

In this section, we provide a more technical description of the identification assumptions of our method. We also discuss the steps to implement our method. We focus on a value-added function with only one predetermined input for the sake of simplicity. We also describe extensions to accommodate potential issues that are likely to arise in empirical work such as issues with the measurement of the proxy variable, the introduction of non-predetermined inputs and measurement error in capital.

C.1 Baseline model: using beliefs as proxy variable with predetermined regressors.

Consider the following value added function in which we assume that log value-added ($va_{i,t}$) is a function of log productivity $\omega_{i,t}$, log capital $k_{i,t}$, and shocks to production $\epsilon_{i,t}$.

$$va_{i,t} = \omega_{i,t} + \beta_k k_{i,t} + \epsilon_{i,t}$$

We assume that capital is pre-determined with respect to $\epsilon_{i,t}$ –i.e., households pick capital without observing shocks to production. However, capital is chosen based on $\omega_{i,t}$. Thus estimating β_k through OLS will yield biased estimates of the contribution of capital to value-added.

Assumption 1: We assume that ω follows a first-order Markov process:

$$\omega_{i,t} = \mathbb{E}[\omega_{i,t}|\omega_{i,t-1}] + \zeta_{i,t}$$

The previous structure suggests that households make choices based on foreseen variation in productivity ($\mathbb{E}[\omega_{i,t}|\omega_{i,t-1}]$) and in response to unforeseen productivity ($\zeta_{i,t}$). This structure is assumed in traditional control-function models such as Levinsohn and Petrin (2003); Olley and Pakes (1996) and Ackerberg et al. (2015). Similarly, our approach proxies for productivity using a variable that, conditional on input use, is a strict monotonic function of $\omega_{i,t}$. However, we do not rely on first-order conditions in the choice of intermediate inputs to derive such relation, since a one-to-one relation between productivity and intermediate inputs may not hold in the presence of credit constraints (Shenoy, 2017a). Instead, we use beliefs about future profits to proxy for productivity. The rationale for this choice is simple: household may use information that is not availble to the researcher–i.e., productivity–to construct forecasts about future profits, which are observed by the researcher.

More formally, we assume that beliefs about profits in period $b_{i,t}$ are a function of input use and productivity;

$$b_{i,t} = b(\omega_{i,t}, k_{i,t})$$

Assumption 2: There is a strict monotonic relationship between productivity and household beliefs, conditional on input use.

One consequence of the previous assumption is that it is possible to invert b and hence write down productivity as a function of beliefs and capital: $\omega_{i,t} = b^{-1}(b_{i,t}, k_{i,t})$.

Combining assumptions 1 and 2, it is possible to achieve identification of β_k. First, using Assumption 2, we plug in $\omega_{i,t} = b^{-1}(b_{i,t}, k_{i,t})$ into the value-added production function to obtain:

$$va_{i,t} = b^{-1}(b_{i,t}, k_{i,t}) + \beta_k k_{i,t} + \epsilon_{i,t} \tag{A1}$$

Or

$$va_{i,t} = \phi(b_{i,t}, k_{i,t}) + \epsilon_{i,t}$$

Where $\phi_{i,t} = b^{-1}(b_{i,t}, k_{i,t}) + \beta_k k_{i,t}$ is an unknown function of beliefs and capital. Note that because household beliefs and capital are pre-determined with respect to $\epsilon_{i,t}$, $\phi_{i,t}$ is identified. Thus, it is possible to use non-parametric or semi-parametric methods to estimate $\hat{\phi}_{i,t}$. In this case, $\hat{\phi}_{i,t}$ captures variation in value added explained by productivity and input use. Note that while Assumption 2 allows us to identify $\phi_{i,t}$, Assumption 2 is not sufficient to identify β_k. The intuition is simple; $\hat{\phi}$ captures the combined contribution of variation in capital to variation in value added. The former is composed of the contribution of capital to production (β_k) but also endogenous responses of capital to productivity, which in turn explains variation in value-added.

We now invoke Assumption 1 to exploit panel variation in $\omega_{i,t}$ to identify β_k. First, note that for a guess value $\tilde{\beta}_k$, we can use $\hat{\phi}_{i,t}$ to recover estimates of productivity $\hat{\omega}(\tilde{\beta}_k) = \hat{\phi}(b_{i,t}, k_{i,t}) - \tilde{\beta}_k k_{i,t}$. Using Assumption 1, it is possible to write down current productivity as an unknown function g of lagged productivity plus an unforeseen productivity shock $\hat{\zeta}$:

$$\hat{\omega}(\tilde{\beta}_k)_{i,t} = g(\hat{\omega}(\tilde{\beta}_k)_{i,t-1}) + \widehat{\zeta_{i,t}} \tag{A2}$$

Note that because $\hat{\omega}(\beta_k)_{i,t-1}$ is only a function of $b_{i,t-1}$ and $k_{i,t-1}$, which are predetermined with respect to the unforeseen shocks to productivity, the following moment condition is satisfied under Assumptions 1 and 2:

$$\mathbb{E}[\hat{\zeta}_{it}(\tilde{\beta}_k)|b_{i,t-1}, k_{i,t}] = 0$$

Thus it is possible to use GMM to recover:

$$\hat{\beta}_k = \text{argmin} \frac{1}{N} \sum_{i=1} \hat{\zeta}_{it}(\tilde{\beta}_k) \tag{A3}$$

C.1.1 Estimation

We estimate β_k following 5 steps:

1. We use third-order polynomials to approximate $\phi_{i,t}$:

$$va_{i,t} = \sum_{h=0}^{h=3}\sum_{j=0}^{j=3} \delta_{i,j} b_{i,t}^{h} k_{i,t}^{j} + \epsilon_{it}$$

2. We use $\hat{\phi}_{i,t} = \sum_{h=0}^{h=3}\sum_{j=0}^{j=3} \hat{\delta}_{i,j} b_{i,t}^{h} k_{i,t}^{j}$ and a candidate for β_k to compute $\hat{\omega}_{i,t} = \hat{\phi}_{i,t} - \tilde{\beta}_k k_{i,t}$. For this, we use an OLS regression of value-added on capital to obtain our first guess ($\tilde{\beta}_k$).

3. We estimate equation (A2) using a third-order polynomial to approximate g. We then compute the residuals $\hat{\zeta}_{i,t}(\tilde{\beta}_k)$ from the following regression:

$$\hat{\omega}_{i,t}(\tilde{\beta}_k) = \sum_{n=0}^{n=3} \theta_n \hat{\omega}_{i,t-1}^{n}(\tilde{\beta}_k) + \zeta_{i,t}$$

4. We iterate across different values for β_k in order to minimize:

$$\frac{1}{N}\sum_{i=1} \hat{\zeta}_{it}(\tilde{\beta}_k)$$

5. Finally, we use our estimates of $\hat{\beta}_k$ to recover estimates of productivity $\omega_{i,t} = \hat{\phi}_{i,t} - \hat{\beta}_k k_{i,t}$.

C.2 Extension: using beliefs as proxies with non-predetermined regressors

Our method can accommodate models in which other inputs such as labor $l_{i,t}$ which are likely to respond to unforeseen productivity shocks $\zeta_{i,t}$. If we assume that contemporary choices of inputs $l_{i,t}$ are correlated with $\zeta_{i,t}$, but previous input choices are not, then β_k and the factor elasticities corresponding to non-predetermined inputs (β_l) are identified based on the following moment condition:

$$\mathbb{E}[\hat{\zeta}_{it}(\tilde{\beta}_k \tilde{\beta}_l)|b_{i,t-1}, k_{i,t}, l_{i,t-1}] = 0$$

That means, that we can use lagged versions of labor to instrument for current labor. This result is analogous to the identification results of Ackerberg et al. (2015).

C.2.1 Estimation

The process is similar to the one without non-predetermined regressors with only two differences.

- First, $\hat{\phi}_{i,t} = \sum_{n=0}^{n=3}\sum_{i=0}^{i=3}\sum_{j=0}^{j=3}\hat{\delta}_{i,j}b_{i,t}^{i}k_{i,t}^{j}l_{i,t}^{n}$ in the first stage.

- Second, using two guess values$\tilde{\beta}_k$ and $\tilde{\beta}_l$, we proceed to compute $\hat{\omega}_{i,t} = \hat{\phi}_{i,t} - \tilde{\beta}_k k_{it} - \tilde{\beta}_l l_{i,t}$. We then estimate equation (A2) and computed the associated residuals to construct the sample analog of:

$$\mathbb{E}\left[\hat{\zeta}_{i,t}(\tilde{\beta})\begin{pmatrix} b_{i,t-1} \\ k_{i,t} \\ l_{i,t-1} \end{pmatrix}\right] = 0 \tag{A4}$$

Finally we use the previous set of moment conditions to recover GMM estimates of $\beta = \{\beta_k, \beta_l\}$ and productivity.

C.3 Extension: using beliefs measured in the previous period

One possible scenario is that the researcher does not obtain contemporaneous measures of beliefs but only beliefs about profits in period t which are constructed with information in period $t-1$. This is the case of our empirical application and we discuss the implications of this caveat for identification. For simplicity, we focus on the one-input case.

In this case, we only observe household beliefs about future profits based on the available information from the previous periods. Let $b_{i,t|t-1}$denote the beliefs about profits in t which are based on previous information. Because households may only observe the foreseen part of productivity ($\mathbb{E}[\omega_{i,t}|\omega_{i,t-1}]$) and not the unforeseen part of productivity ($\zeta_{i,t}$), we can not write beliefs as a function of $\omega_{i,t}$ as in our benchmark especificiation. However, we can write beliefs about profits in period t measured in $t-1$ as a function of the foreseen part of productivity and the available stock of capital :

$$b_{i,t|t-1} = \tilde{b}(\mathbb{E}(\omega_{i,t}|\omega_{i,t-1}), k_{i,t}) \tag{A5}$$

If, conditional on capital, there is a strict monotonic relation between expected productivity and $b_{i,t|t-1}$, then it is possible to write down the value added function as:

$$va_{i,t} = \tilde{b}^{-1}(b_{i,t|t-1}, k_{i,t}) + \beta_k k_{i,t} + \zeta_{i,t} + \epsilon_{i,t} \tag{A6}$$

The previous expression is very similar to equation (A1). The key difference is that the part of productivity that is unforseen to the farmer ($\zeta_{i,t}$) is not present in equation (A1)but shows up in (A6). Thus, our measure of beliefs captures productivity with measurement error, arising from the fact that we measure beliefs at the end of period $t-1$ which precedes the realization of the unforeseen shocks to productivity.

Note that in this case we can still write down value added as an unknown function of beliefs and capital and shocks:

$$va_{i,t} = \tilde{\phi}(b_{i,t|t-1}, k_{i,t}) + \zeta_{i,t} + \epsilon_{i,t} \tag{A7}$$

In this case, identification of $\tilde{\phi}_{i,t}$ requires that $b_{i,t|t-1}$ and $k_{i,t}$ are not correlated to $\zeta_{i,t} + \epsilon_{i,t}$. This could be a problem in terms of identification of ϕ, depending on the assumptions made about how households adjust inputs.

Predetermined regressors. In this setting, the former condition is satisfied as both beliefs and capital are measured at the end of period $t-1$ and are not correlated to $\zeta_{i,t} + \epsilon_{i,t}$. Thus, we can apply our framework with our measures of beliefs and still consistently estimate ϕ.

Regressors correlated with $\zeta_{i,t}$. It could be the case that household are able to partially modify inputs in the aftermath of unforeseen productivity shocks. If that is the case, then current measures of input would be correlated with $\zeta_{i,t}$ and ϕ in equation (A7) will not be identified. However, it is possible to use lagged versions of input to instrument for current input usage and consistently recover ϕ.

Once$\hat{\phi}$ is recovered, the rest of the procedure remains unchanged.

C.4 Extension: measurement error in capital

Collard-Wexler and De Loecker (2016) highlight the consequences of measurement error in capital in the context of the estimation of production functions. In this paper we are particularly interested in understanding the extent to which our results are robust to productivity estimates that account for potential measurement error in capital. Note that in the context of control-function methods, measurement error can affect identification β_k directly and indirectly through biased estimates of ϕ.

While $k_{i,t}$ may be measured with error arising from failure to recall the initial level of capital and assumptions regarding depreciation, we argue that investment spending in period $t-1$ should be highly predictive of the stock capital at t and is unlikely to suffer from measurement error related to imperfect recall and depreciation. Using this insight we apply Collard-Wexler and De Loecker (2016)'s procedure to correct for measurement error.

1. We approximate ϕ using a linear regression in which we use investment in $t-1$ ($i_{i,t-1}$) as an instrument for $k_{i,t}$ in the following regression:

$$va_{i,t} = \delta_0 + \delta_1 b_{i,t} + \delta_2 k_{i,t} + \delta_3 l_{i,t} + \epsilon_{it}$$

2. We use $\hat{\phi}_{i,t} = \hat{\delta}_0 + \hat{\delta}_1 b_{i,t} + \hat{\delta}_2 k_{i,t} + \hat{\delta}_3 l_{i,t}$ and candidates for β_l, β_k to compute $\hat{\omega}_{i,t} = \hat{\phi}_{i,t} - \tilde{\beta}_l l_{i,t} - \tilde{\beta}_k k_{i,t}$. Note that $\hat{\delta}_2$is estimated using lagged investment as an instrument for k, and that we assume a linear process as opposed to a semi-parametric approach as the latter would imply further assumptions regarding the use of investment as an instrument (Collard-Wexler and De Loecker, 2016).

3. We estimate equation (A2) using a an AR(1) process. Again, note that we assume a linear process as opposed to a semi-parametric approach as the latter would imply further assumptions regarding the use of investment as an instrument. We then compute the residuals $\hat{\zeta}_{i,t}(\tilde{\beta}_k)$ from the following regression:

$$\hat{\omega}_{i,t}(\tilde{\beta}_k) = \rho \hat{\omega}_{i,t-1}(\tilde{\beta}_k) + \zeta_{i,t}$$

4. We obtain the GMM estimates of β_l and β_k based on the following moment conditions:

$$\mathbb{E}\left[\hat{\zeta}_{i,t}(\tilde{\beta})\begin{pmatrix} b_{i,t-1} \\ i_{i,t-1} \\ l_{i,t-1} \end{pmatrix}\right] = 0 \tag{A8}$$

   Where $i_{i,t-1}$ denotes investment expenses of household i during the period $t-1$. In this case identification is based on a different moment condition which implies that lagged investment is uncorrelated with foreseen shocks to productivity.

5. Finally, we use our GMM estimates ($\hat{\beta}_l, \hat{\beta}_k$) to recover estimates of productivity $\omega_{i,t} = \hat{\phi}_{i,t} - \hat{\beta}_l l_{i,t} - \hat{\beta}_k k_{i,t}$.