Micro-enterprise saturation: Critical mass or overcrowding?*

Moritz Poll†

March 26, 2026

Latest version here.

Abstract

Should entrepreneurship support programs to extremely poor people be scaled by saturating a few villages, or maintaining low saturation and enrolling more villages at a higher cost? On the one hand, saturation may create a *critical mass* of entrepreneurs whose income generation and associated spending multipliers make each other viable. On the other hand, encouraging the start-up of many small firms in close proximity may be a recipe for *overcrowding* and artificially fierce competition in a small village economy. In a large randomized control trial in Malawi, I exogenously vary the saturation of an entrepreneurship support program to assess whether business performance increases or decreases in the share of treated neighbors. Business performance is tracked at fortnightly high frequency. Most participants start a retail business. Eighteen months after the start of the intervention, earnings take a 21% hit as saturation doubles and aggregate earnings are split among the larger number of supported entrepreneurs. About a third of the new retail businesses have gone out of business. Earnings losses and firm exits are of a comparable magnitude to implementer cost savings. The heightened competition does not lower consumer prices, but induces out-migration. I show that the sectoral concentration in retail is not driven by information frictions about other entrants' intended sectoral choice. Instead, respondents highlight a lack of hard skills.

KEYWORDS: Entrepreneurship, poverty alleviation, general equilibrium, randomized control trial, saturation design, spillovers, coordination

JEL CLASSIFICATION: O12, O14, O25, D7, D51, D58

*I owe deep gratitude to the participants who dedicated their time to this study; to countless local stakeholders and practitioners who shared their wisdom and advice; and of course to the staff of my implementing partner without whose tireless work this project would not have been possible. I want to thank Andrew Foster, Peter Hull, and Stelious Michalopoulos for their support and guidance. This project greatly benefited from discussions with Clare Balboni, Dan Björkegren, Dennis Egger, Doug Gollin, Johannes Haushofer, Reshma Hussam, Fedor Iskhakov, Joe Kaboski, Michael King, Cynthia Kinnan, Jessica Leight, Sara Lowes, Elisa Macchi, David McKenzie, Josh Merfeld, Jonathan Morduch, Paul Novosad, Matt Pecenco, Tommaso Porzio, Simon Quinn, Jonathan Robinson, Jon Roth, Bryce Steinberg, Hyun Soo Suh, Nick Tsivanidis, Davide Viviano, Maarten Voors, and Chris Woodruff. Generous funding from the Weiss Fund (BN-002), the National Science Foundation (#2315009), the Dovetail Impact Foundation, the Abdul Latif Jameel Poverty Action Lab (J-PAL GEA-25-03001 & LAI-25-03010 from UK Government International Development funding), the Livelihood Impact Fund, the Gates Foundation (BMG-23-10013-X3 via Innovations for Poverty Action), as well as the Bravo Center, Population Studies and Training Center, Watson Institute, and Nelson Center at Brown University is gratefully acknowledged. I also benefited from fellowships funded by the Watson Institute, Brown University Graduate School, the Population Studies and Training Center (under T32 NICHD grant and PSTC P2C NIH grant), and the Barrett Hazeltine Fellowship for Graduate Research in Entrepreneurship. This study was pre-registered at https://www.socialscienceregistry.org/trials/11789 and approved by Brown University IRB (#98) and Malawi's NCRSH (NCST/RTT/2/6).

†Brown University, Department of Economics. Email: moritz.poll@brown.edu. Website: www.moritz-poll.com

1 Introduction

Poverty alleviation in the absence of stable employment opportunities from the private or public sector relies on individual entrepreneurship by poor people themselves (Blattman and Ralston, 2015; Banerjee et al., 2025). Policies that support entrepreneurship in developing countries are growing in popularity, but economic theory does not offer a clear-cut prediction of their impact at scale. On the one hand, the reason we are seeing few firms in contexts of extreme poverty might be that firms are not individually profitable as they depend on the weak demand of other similarly poor people. It is possible that a large-scale support program would create a critical mass of entrepreneurs whose income generation and associated spending multipliers make each other viable. The logic of a big-push individual-level intervention (Balboni et al., 2022) would then carry over to the aggregate (local) economy (Rosenstein-Rodan, 1957). On the other hand, encouraging the start-up of many small firms in close proximity may be a recipe for overcrowding and artificially fierce competition in a small village economy (Hardy and Kagy, 2020). Failure rates for micro-enterprises are high, especially soon after launch (De Mel et al., 2012; Balboni et al., 2022; Misha et al., 2019) and anecdotally too many UPG-beneficiary households start similar businesses in close proximity (Bauchet et al., 2015). This would suggest tight upper bounds to how far these programs can be scaled – and more broadly to the approach of alleviating poverty bottom-up from the household level (Goldberg and Reed, 2023).

These two scenarios (critical mass versus overcrowding) motivate my study. They imply contrasting policy prescriptions for whether entrepreneurship should be fostered by saturating a few locations or scaling over a larger geographic area at low saturation. In a large randomized control trial in Malawi, I exogenously vary the program saturation of an entrepreneurship support training to assess whether business performance increases or decreases in the share of treated neighbors. Saturation at the household cluster level is randomly assigned to be either 50% or 100% of eligible households, corresponding to 18.5% and 37% of the population, respectively. I also test a simple coordination intervention encouraging households to coordinate with each other when choosing their business venture, rather than duplicating existing businesses. The entrepreneurship support program under investigation is the graduation model, first popularized by the Bangladeshi NGO BRAC, which combines monthly cash transfers, productive assets or business start-up grants, entrepreneurship training, formation of savings groups, and one-on-one life skills coaching.

The coordination intervention did not succeed in shifting entrepreneurs-to-be out of the retail sector. While there are weak indications of sectoral shifts in both the business plans that program participants submit as well as the livelihoods households report very shortly after the cash grant, these differences dissipate within just a few months. My study demonstrates that the barrier is not simply an information friction, but I provide evidence that the lack of hard skills confines these entrepreneurs in the petty trading sector.

Eighteen months after the start of the intervention, the saturation results are nuanced: In terms of earnings, the businesses in high and low saturation clusters initially track each other's performance closely over time, but show signs of divergence, with low saturation cluster entrepreneurs

earning at times up to twice as much. The cumulative earnings are not significantly different from one another, but the point estimate for the high-saturation group is about 21% lower than that of the low-saturation group. It is important to note that an implementer may prefer to implement at high saturation even if earnings were significantly lower as long as earnings losses do not offset savings from economies of scale. In the program I analyze, the cost savings at high saturation are about 30%. Firms in the high saturation clusters are significantly more likely to exit, with households reverting to casual labor. An even starker indication of overcrowding is that the higher saturation induces out-migration.

Ultra-poor graduation programs (UPGs), aimed at promoting entrepreneurial activity by people in extreme poverty, are the closest that development economics has come to a silver bullet for sustainably alleviating extreme poverty (Sulaiman et al., 2016; Wydick, 2020). While in-kind, conditional and unconditional cash transfers have been shown to provide effective poverty relief and invigorate micro-enterprises in the short run (Sulaiman et al., 2016; Baird et al., 2011; Blattman et al., 2014; Haushofer and Shapiro, 2016; Cahyadi et al., 2020; Brooks et al., 2022; Egger et al., 2022), in-kind transfers or strong-handed conditionalities may induce distortionary effects (Bryan et al., 2021), and the evidence on long-term effects is mixed at best and favors settings where households are above the international extreme poverty line to begin with (Gertler et al., 2012; De Mel et al., 2012; Aizer et al., 2016; Ikegami et al., 2017; Haushofer and Shapiro, 2018; Araujo et al., 2019; Millán et al., 2020; Aggarwal et al., 2022; Altındağ and O'Connell, 2023; Fiala et al., 2022). Skills training, coaching, or access to loans or savings facilities by themselves show few if any impacts (Blattman and Ralston, 2015; Berge et al., 2015; Valdivia, 2015; Bauchet et al., 2015; Chowdhury et al., 2017; Brooks et al., 2018; Sedlmayr et al., 2020; Giné and Mansuri, 2021; Leight et al., 2021, 2022). On the other hand, UPGs following the BRAC model (studied here) seem to be standing the test of time while providing results that are larger than the sum of their parts in response to the highly complex challenge that multidimensional poverty traps pose (Sulaiman et al., 2016; Banerjee et al., 2015; Bandiera et al., 2017; Banerjee et al., 2021; Balboni et al., 2022; Chowdhury et al., 2017; Sedlmayr et al., 2020; Bedoya et al., 2019). If the first Sustainable Development Goal of ending extreme poverty by 2030 is to be achieved, the development community needs to better understand how UPGs scale. My study highlights that a key barrier to scale will be the access to a broader set of sectors through the provision of diverse hard skills – a challenging barrier considering that providing a diverse set of hard skills within the same location will erode economies of scale.

Spillovers are a common concern in poverty alleviation programming. While Egger et al. (2022) find almost no inflationary pressure from cash transfers in Western Kenya, Filmer et al. (2021) argue that this may be linked to relatively low saturation in the study area while in their own study they find important price increases when cash transfer saturation is high and markets remote and Cunha et al. (2019) find differential price effects by village wealth. In the context of business training interventions, McKenzie and Puerto (2021) find no evidence of negative side effects, while Drexler et al. (2014) show that in their context, at least part of the benefits to the treated come from business stealing from the control group. Such spillovers may also extend to programs that

support new entrepreneurial activity, the focus of this study. I do not find impacts of program saturation on the price of the key commodity, maize. Spillovers from the treatment group onto the control group are likewise modest. But I show that when taking an intervention to scale, it is important to consider spillovers from treated onto other treated people.

This paper proceeds as follows. Section 2 describes the setting, research design, and intervention. Section 3 discusses the analysis, while results can be found in Section 4.

2 Experimental design

Setting and census This study takes place in Mangochi district in the South of Malawi, one of the poorest districts in the country. To determine program eligibility and convey an accurate picture of the local economy, I conducted a household census of all 29,000 households in the area, covering a population of almost 140,000 individuals (Appendix Table A1). The majority of the population are subsistence or pre-commercial farmers (Appendix Figure A1). Outside the harvest season, few have any income sources other than ad hoc daily labor ("ganyu"). About 40% of aggregate production in the study area is consumed without being traded in a market. Three quarters of households report at least one member having skipped a meal in the past week (Appendix Table A3). This is in line with the fact that agricultural yields are low. The 2018 harvest lasted fewer than 20% of farmers longer than six months and four months after the 2019 harvest only half of registered households had any food reserves at all. It is six months after harvest that investments in the next planting, both in terms of seeds and fertilizers as well as in terms of labor need to be made, but when a majority of the rural poor are running precariously low on food and money, which they compensate for with casual labor, drawing them away from their own fields. Existing businesses are overwhelmingly one-person enterprises and a sizable share depend on demand within their own village.

Sample selection and treatment assignment The sampling process is described in Appendix Figure A2. Using the Malawi-validated Simple Poverty Score Card (Schreiner, 2019) and self-reported income, I identified about 40% of households as ultra-poor and eligible.[1] Using k-means geographic clustering, these roughly 11,000 households were assigned to 271 clusters. Due to how households tend to live nearby one another, clusters often roughly correspond to villages when these are small and remote. In larger towns, villages are usually comprised of several clusters or a cluster can combine households from more than one village. To determine the sample, I deployed a custom algorithm to select 72 clusters into the study in such a way as to maximize the distance between clusters and minimize the potential for spillovers with buffer clusters in between that are

[1] The implementing partner excluded households from participation that have no female; have no children under the age of 5 as of August 1, 2023; did not appear ultra-poor based on either their stated income or the Simple Poverty Scorecard; in which the female was unable to carry out physical labor; whose data appeared implausible to the research team or enumerator; or who were not willing to be registered or could not be found during several (mop up) visits. The latter two criteria were very rare.

not part of the study.[2] The median household cluster is about 30 hectares in size and is home to about 70 households. The minimum distance between study clusters corresponds to the distance beyond which Egger et al. (2022) no longer find spillover effects of cash transfers.[3] In my analysis, I will therefore allow for spillovers within but not across clusters. Though this assumption may not exactly hold in practice (for example information is likely to "travel further" than cash), the assumption is that spillovers are substantially greater between neighbors within a cluster than they are across clusters. Clusters were grouped into sets of three that were assigned to a program implementer as well as being randomly assigned to one of three cohorts to be launched a few months apart from each other (early 2024, mid 2024, mid 2025, see Appendix Figure A3).

Within each implementer's area (stratum), the three clusters were assigned at random to either high treatment saturation (everybody treated), low treatment saturation (half treated, randomly selected at the individual level) or pure control group (nobody treated). The saturation levels were chosen following Baird et al. (2018).[4] Assigning each implementer one of each type of cluster is tantamount to geographic and implementer stratification (Bruhn and McKenzie, 2009; Bai, 2022).

In order to balance the sample on a number of key outcome and mechanism variables (see Table A4), I first run the treatment assignment algorithm 100,000 times (seeding each with the counter of its iteration). I then estimate baseline balance under each theoretical treatment assignment and discard assignments that lead to a statistically significant imbalance in any of the key mechanism variables at the 10% significance threshold (Morgan and Rubin, 2012). In other words, I enforce a "visually clean" balance table. I then randomly pick one of the permissible treatment assignments. As pointed out in both Bruhn and McKenzie (2009) and Morgan and Rubin (2012), classical inference that does not take this rerandomization into account would now on average (though not in every single case) be too conservative, leading to confidence intervals on the average treatment effect with overly wide coverage and too few rejections of null hypotheses. A Fisher (1935) exact test has the correct coverage, though it tests the sharp "randomization hypothesis" that the treatment had no effect on anyone rather than on average (Ritzwoller et al., 2025). Traditionally, researchers using rerandomization methods have done so somewhat *ad hoc*, rerandomizing until the balance table was satisfactory. My approach is more transparent. Given R simulated assignments and k baseline balance table variables, the likelihood of discarding an assignment at the 10% significance level is $q = 1 - (1 - 0.1)^k$ and I can therefore use the remaining $R \times (1 - q)$ simulated assignments for the randomization inference (Young, 2019).

[2] After cluster selection and baseline survey, but before the start of the intervention, the local government designated one of our clusters as a flood-prone area and ordered it to be vacated. As our study may have encouraged people to stay in an unsafe area, we shifted to a neighboring cluster in this one instance.

[3] If anything, my study area's road network is substantially less developed and personal ownership of bicycles or even motorbikes is lower in Southern Malawi than in Western Kenya and the terrain is sliced up by seasonal rivers that make effective distances larger than straight-line distance.

[4] I am designing my study to be optimally powered for spillovers among treated households. Modifying equation (6) in Baird et al. (2018) to omit the standard error of the spillover effect on the control group, I restate their Theorem 2 to conclude that the optimal high saturation is 100% and the optimal low saturation would have been below 30% and dependent on intra-cluster correlation of outcomes. However, in conversation with the implementing partner, we decided to increase the lower saturation to 50% which would represent a more policy-relevant comparison. Less saturated clusters would also make program components like savings groups difficult to implement.

I leverage the exogenous variation between high and low saturation to answer whether increased saturation leads to excess competition and overcrowding of small village economies or to critical mass and spending multipliers through which the budding entrepreneurs make each other viable.

Intervention The intervention is a standard BRAC-style UPG (Banerjee et al., 2015; Bandiera et al., 2017) with two key modifications. First, instead of livestock and related training, households receive a one-off labeled[5] cash grant of $300 intended for starting a business. Flexibility is given to participants for leveraging their own informational advantage over ability and preferences and they receive four months of weekly general-purpose business training covering financial literacy, market analysis, budgeting, record keeping, funding sources, and the development of business ideas. Second, program participants (mothers of children under five) receive an additional four-month training course on early childhood development. Otherwise, the program is comparable to BRAC's and other UPGs in the literature: Treated households receive six monthly cash transfers of $20 (labeled for investments in child health and wellbeing as well as business experimentation prior to the large grant). The structure of cash transfers is closely aligned with evidence on how poor households themselves prefer to receive such funds, combining a regular buffer with a large lump-sum that arrives after sufficient time for deliberation (Kansikas et al., 2023).

All households in the study receive fortnightly visits from program implementers which in the original program design are intended for tracking training progress and answering questions the participant may have. During the program evaluation, we additionally use these meetings for brief surveys on consumption and investment choices, livelihood activities, as well as shocks and coping behavior. Both the treatment and control group receive a basic phone for scheduling of surveys and for receipt of cash transfers or survey incentives through mobile money. At the community level, the formation of savings and loan groups is encouraged. Local childcare centers are supported financially and with the same ECD training. Both treatment and control group are expected to benefit from this last aspect.

In half of the treated clusters, implementers do not guide which kinds of businesses participants start start, while the other clusters receive a business selection guidance treatment. In these clusters, participants first listen to a story of three women in a similar position as them, having participated in an entrepreneurship training program and ready to receive their cash grant when they discover that they all three had the same business idea. The women in the story come to the realization that there would probably not be enough demand for all three of them to pursue the same business and they instead choose different businesses. After reading the story, the implementers ask participants to share with each other what kinds of businesses they would like to start (as a ranked preference list). Participants are then asked to consider what would happen to their profitability if they all started similar businesses and encouraged (but not forced) to coordinate so that they do not all

[5]The grant is largely unconditional, but the NGO encourages that it be used for starting a business or supporting an existing one. Receipt requires attending the training and being able to adequately verbalize a concrete business plan, but participants can miss sessions without penalty and are coached until their plan is satisfactory and the actual implementation of said plan is not afterwards enforced.

start the same businesses as one another. The goal is to first widen participants' pool of potential business ideas and second to prime them to think about the general equilibrium consequences of their business choice. The coordination intervention is motivated by the fact that when asked why they would like to pursue a particular research idea, the majority of participants mentioned that they had seen other people earn good money this way. For an individual prospective entrepreneur, imitation is of course a potentially sensible strategy. But the fact that this is by far the predominant motive points to a potential neglect of general equilibrium effects that may be detrimental at scale (Figure 1).

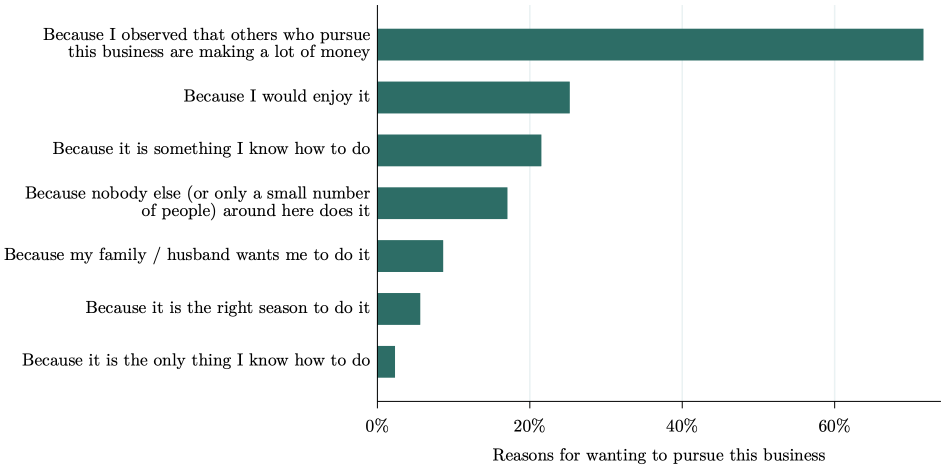


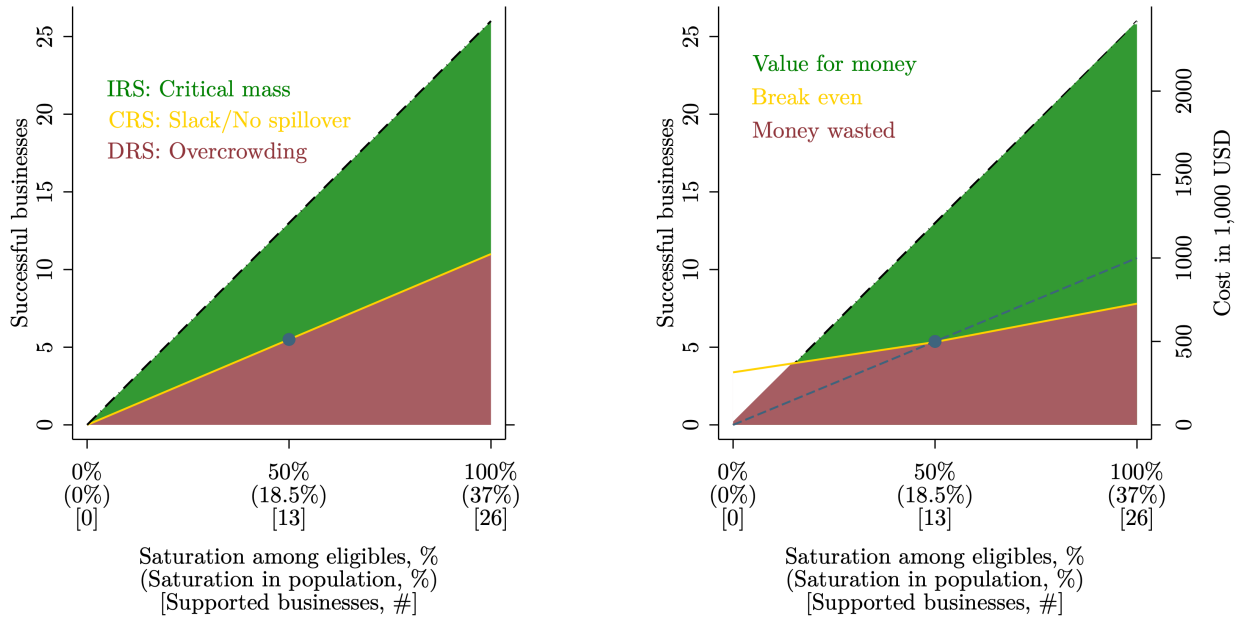
Figure 1: Evidence of general equilibrium neglect

Notes: This figure summarizes common motivations that participants state for why they would like to pursue a given business idea. The overwhelming majority expresses what I term general equilibrium neglect: That they would like to emulate the profits of others, perhaps without internalizing the effects of their own entry. I hypothesize that this will have detrimental consequences in a high-saturation implementation and it motivates devising a business coordination intervention.

3 Empirical strategy

There are two natural benchmarks for program performance. One is to ask, whether more or fewer businesses become successful relative to how many were supported (and how well they perform) as more people are treated (Figure 2(a)). So for example, if five out of thirteen supported businesses are successful in a low saturation cluster, then if the program scales with constant returns to saturation, we would expect ten to succeed when 26 are supported. Under increasing returns, it would be more than ten and vice versa. But in practice, some losses in program effectiveness are permissible considering the important economies scale in implementation cost (Figure 2(b)). High-saturation clusters, where all 26 households (100% of eligible households) are treated, correspond to a saturation level among the broader population surveyed during census of about 37% in the median cluster. In low-saturation clusters, 13 randomly chosen households (50% of the eligibles) are treated,

corresponding to 18.5% of all households in the median cluster. Given my intimate familiarity with the cost structure of the program I evaluate, I am able to investigate not just whether there are *any* adverse saturation effects (which will be of interest to academic economists), but also how they relate to the cost savings (which will be of interest to policy makers and practitioners) when the goal is to maximize the number of beneficiaries with viable businesses *given* a fixed budget. The yellow line in the right panel, shows the program cost increasing less than one-for-one between the low and high-saturation scenario. The gap between the blue dashed and the yellow solid line represents the case where the program has decreasing returns to saturation, but they are outweighed by the economies of scale.



(a) Suppose at low saturation, the treatment produces about five successful businesses (blue dot). Relative to that, scaling with constant returns to scale (no spillovers) is indicated by the yellow line while increasing and decreasing returns to scale, are scenarios indicated by green (critical mass) and red (overcrowding) respectively.

(b) Implementation costs, however, increase less than proportionally with saturation and the area between the constant returns to scale line (blue, dashed) and the decreasing cost of scale curve (yellow) may still be attractive if the goal is the most successful businesses per dollar spent.

Figure 2: Conceptual framework

As pre-specified, I estimate the following linear regressions:[6]

$$y_i = \beta_0 + \beta_T \text{Treatment}_i + \beta_C \text{Coordination Intervention}_i + \beta_{\text{BL}} y_i^{\text{BL}} + \beta_{\text{M}} \mathbb{1}[y_i^{\text{BL}} \text{ missing}] + \beta_{\text{FE}} S_i + \varepsilon_i \quad (1)$$

$$y_i = \beta_0 + \beta_T \text{Treatment} + \beta_H \text{High saturation} + \beta_{\text{BL}} y_i^{\text{BL}} + \beta_{\text{M}} \mathbb{1}[y_i^{\text{BL}} \text{ missing}] + \beta_{\text{FE}} S_i + \varepsilon_i \quad (2)$$

Standard errors are clustered at the strata (cluster triplet) level (De Chaisemartin and Ramirez-Cuellar, 2024) which is the level of randomization and I am controlling for baseline levels of all outcome variables (ANCOVA). I investigate three key outcome variables, sector choice, business survival and profits.[7] Equation 1 investigates whether the addition of the coordination information intervention alters sector choices among treated households.

The sign of β_H in equation (2) is the main parameter of interest in this study, theoretically ambiguous, and merits discussion. If it is relatively close to zero for both outcomes, there would be no evidence of spillovers among the treated for business performance. This would indicate room in the local economy for more firms. If the coefficient is positive for either outcome, this would speak in favor of the spending multipliers and critical mass hypothesis. If either one or both of the coefficients are negative, this would support overcrowding. If one were positive and the other negative (e.g. more surviving firms but with lower profits), then the local economy may be at capacity, but the constraint operates only along one margin.

4 Results and mechanisms

4.1 Regression results

In Appendix Table A4, I show sample balance. As discussed above, the clean balance table follows mechanically from the randomization approach. Attrition at the one-year follow-up (midline) survey is 2.7% at midline and 9.5% at endline. Attrition is balanced between treatment and control group.

The coordination lesson did not meaningfully increase the number of distinct or unique businesses in a cluster. It appears to have marginally compelled program participants to *propose* non-retail service businesses ahead of receiving the cash grant (Table 1), but it did not meaningfully affect the *actual* sectors they are operating in after receiving the cash grant (Table 2). Since the coordination intervention did not affect business choices meaningfully, I henceforth pool their results to compare high and low saturation clusters.

Turning to the performance of the businesses, I compare high-saturation clusters to two benchmarks, as described in section 3. First, I compare them to low-saturation clusters to investigate

[6] The pre-analysis plan additionally contained an interaction of the saturation level and coordination intervention. Since the latter turned out to be ineffective, I concentrate the statistical power on the simple comparisons instead, which should therefore be interpreted as weighted averages of the treatment effect of the uninteracted treatment effects (Muralidharan et al., 2023).

[7] In what follows, I will use the term "profit" somewhat loosely. To be precise, I will focus on the gross margin, defined as revenue minus all ongoing expenses and abstract away from any amortization of a potential asset that households may already own or buy from intervention cash grant.

whether program impacts scale linearly or display positive or negative spillovers that intensify with saturation. Second, I compare performance to a cost benchmark, since implementation at higher saturation saves the implementing partner about 30% of their per-household costs. I use two main metrics: One view of these entrepreneurship support programs is that not everybody is a good entrepreneur and it is very hard to know ex ante who will be talented (or lucky) enough to perform well (Hussam et al., 2022; Haushofer et al., 2022; Banerjee et al., 2025). I therefore count as successful entrepreneurs those who earn at least threefold the average control group earnings between after the cash grant and for the remainder of the study duration. I estimate that a little over half of households in both low and high-saturation clusters are successful by this metric, an increase by 25 percentage points over the control group (Figure 3, left). The estimate is somewhat but not significantly lower in the high-saturation group where it sits in between the low-saturation and the cost benchmark.

When it comes to earnings, the point estimates group are about 21% lower in high-saturation clusters than in the low-saturation clusters, close to the cost savings benchmark, though not statistically significantly different from either benchmark (Figure 3, right). I can confidently reject large positive spillovers as found in Egger et al. (2022). Figure 4 shows that earnings diverge from time to time between the two cluster saturation levels (Table 3). Earnings co-move within clusters, making standard error clustering very taxing.

With more firms and lower earnings per firm, it seems reasonable to ask whether the business environment in the treated areas has become more competitive, entailing spillovers outside the study sample. I track maize prices from August 2024 onwards in the entire study area. As Malawi's foremost staple crop, maize has a pivotal role as a barometer for price levels more broadly and many of the studied businesses engage in retail of food crops and may exert downward pressure on maize prices. Figure 5 suggests that maize prices track each other through heavy seasonal fluctuations in all cluster types for the first few months of the study, though during the lean season price peak, the treated clusters seem to have had slightly lower prices. This divergence, however, is driven if anything by the low saturation cluster, suggesting that higher saturation is not a driver of more competitive retail markets for maize.

Table 1: Business choices (planned)

	(1) Treatment mean (SD)	(2) Coordination Intervention	(3) N (BL)
Agriculture, Fishing, Forestry	.0733	-.009082	914
	(.2608)	(.01564)	✓
Manufacturing	.0339	-.0008599	914
	(.1811)	(.01664)	✓
Non-retail Services	.0481	.0351*	914
	(.2142)	(.01824)	✓
Retail	.8239	-.01487	914
	(.3812)	(.02813)	✓
Distinct businesses	7.604	.9583	48
	(2.06)	(.9639)	×
Unique businesses	3.75	.6667	48
	(1.952)	(.8102)	×
Cluster HHI	.066	-.0005026	48
	(.0238)	(.01337)	×

Notes: Analysis in this table is among treated households (who wrote a business plan) and the treatment coefficient is therefore omitted. The last column indicates the number of observations. All regressions are implemented as an ANCOVA specification, controlling for strata fixed effects. A ✓ indicates that the outcome variable is available for all observations at baseline. If it was not collected at baseline, a × is shown, and if it is partially available, missing observations are set to zero and dummied out (✓). Where available, regressions control for baseline levels of the outcome. Standard errors are clustered at the household cluster level. Significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Business sector choices (actual)

	(1) Control mean (SD)	(2) Treatment Effect	(3) Coordination Intervention	(4) N (BL)
Casual labor, charity	.6015	-.3605***	.05629	997
	(.4902)	(.0568)	(.04526)	✓
		[0]	[.2398]	
Farming, fishing, forestry	.3239	.04006	-.009991	997
	(.4686)	(.05583)	(.02932)	✓
		[.4755]	[.7153]	
Manufacturing	.0411	.06793**	-.03696	997
	(.1988)	(.02477)	(.0291)	✓
		[.02797]	[.2607]	
Non-retail services	.0848	.04641**	.04792*	997
	(.279)	(.02039)	(.02251)	✓
		[.02697]	[.04396]	
Retail	.0823	.3169***	-.04279	997
	(.2751)	(.03162)	(.03735)	✓
		[0]	[.2787]	
Unique income source	.1732	.06441*	.08439	970
	(.3789)	(.03188)	(.05303)	×
		[.07293]	[.1159]	
New income source (village)	.0577	.08364*	.03768	970
	(.2336)	(.04063)	(.04401)	×
		[.05594]	[.4066]	

Notes: The last column indicates the number of observations. All regressions are implemented as an ANCOVA specification, controlling for strata fixed effects. A ✓ indicates that the outcome variable is available for all observations at baseline. If it was not collected at baseline, a × is shown, and if it is partially available, missing observations are set to zero and dummied out (✓). Where available, regressions control for baseline levels of the outcome. Standard errors are clustered at the strata level. Fisher (1935) exact test p-values from 1000 simulated alternative treatment assignments are shown in brackets. Significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Business performance

	(1) Control mean (SD)	(2) Treatment Effect	(3) High Saturation	(4) N (BL)
Does business	.4722	.3796***	.01082	995
	(.3505)	(.04556)	(.01443)	✓
		[0]	[.4625]	
Investment in inventory	216,500	1,537,971***	-256,487	995
	(506,077)	(296,993)	(210,084)	×
		[0]	[.2428]	
∼ productive assets	11,874	27,742***	28,202*	995
	(105,511)	(8,522)	(14,667)	×
		[.000999]	[.06793]	
Hours worked	257	29.53	-21.94***	995
	(163.4)	(18.97)	(6.804)	×
		[.1568]	[.000999]	
Earnings	723,816	1,297,310***	-278,848	995
	(707,048)	(390,094)	(224,934)	×
		[.001998]	[.2677]	

Notes: The last column indicates the number of observations. All regressions are implemented as an ANCOVA specification, controlling for strata fixed effects. A ✓ indicates that the outcome variable is available for all observations at baseline. If it was not collected at baseline, a × is shown, and if it is partially available, missing observations are set to zero and dummied out (✓). Where available, regressions control for baseline levels of the outcome. Standard errors are clustered at the strata level. Fisher (1935) exact test p-values from 1000 simulated alternative treatment assignments are shown in brackets. Significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

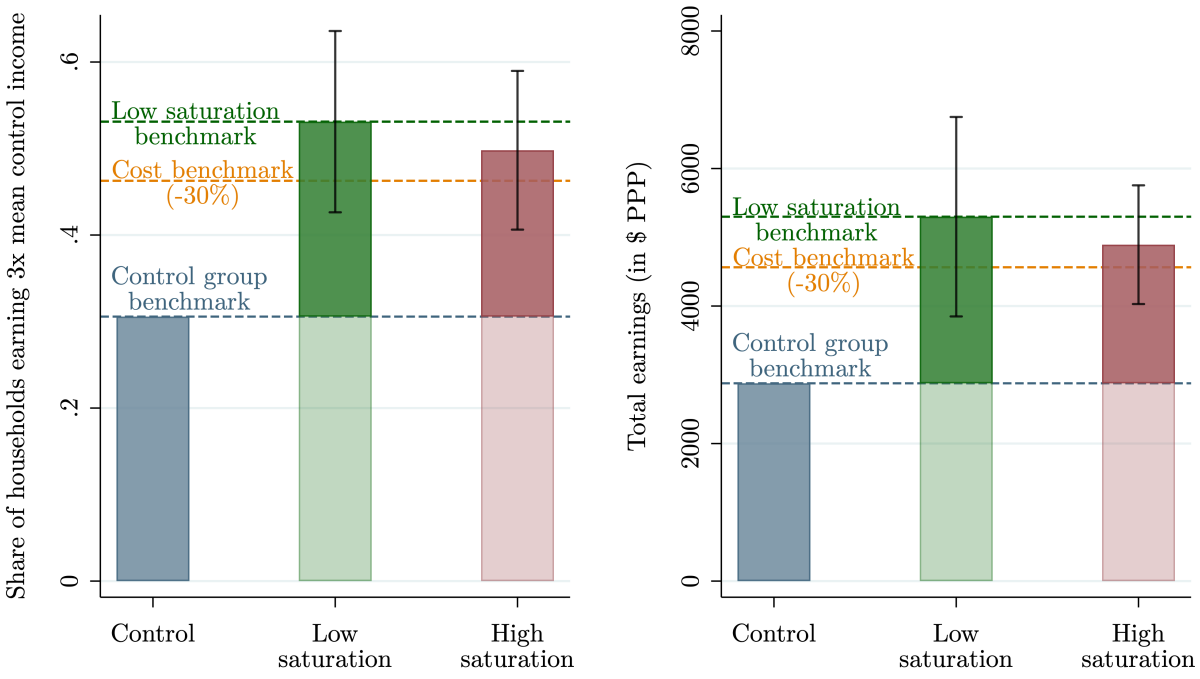


Figure 3: Earnings by saturation

Note: This graph compares regression results for indicators of being a successful entrepreneur (left) as well as average earnings (right) across cluster types.

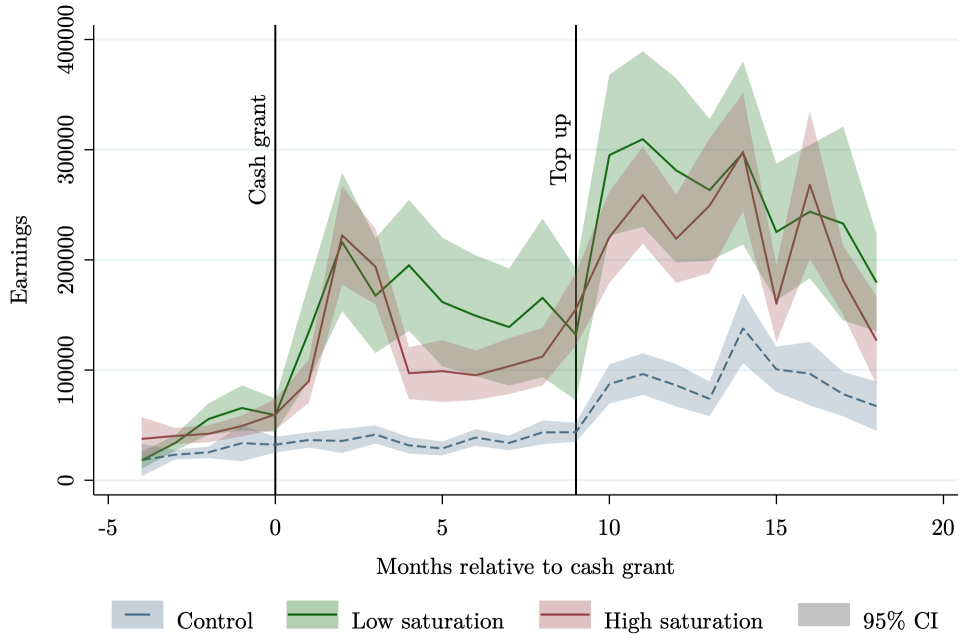


Figure 4: Earnings by saturation

Note: This graph plots monthly earnings over time relative to the cash grant. After an initial spike in the treatment group, earnings diverge and stay higher for the low-saturation clusters.

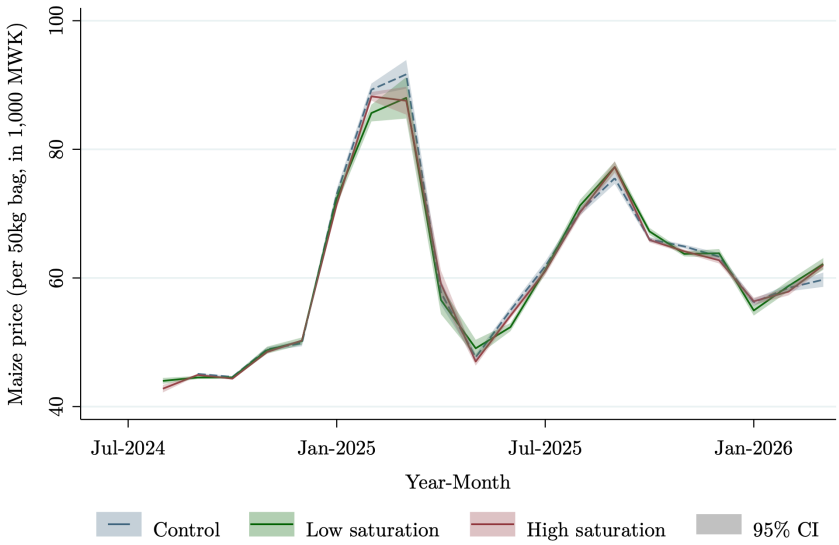


Figure 5: Maize price by saturation

Note: The price of maize, the dominant staple crop in the region, follows strong seasonal fluctuations, but does not appear affected by the high- or low-saturation treatment.

4.2 Mechanisms and robustness

The section discusses potential channels through which overcrowding might operate or manifest itself. Lower profits are not a mechanical byproduct of adverse selection into business as the 13 households selected into treatment among the 26 households that make up a low saturation cluster are chosen at random and Table A4 shows balance on ability and prior experience. I do not find any evidence that entrepreneurs in the different types of clusters spend more or less time working (Figure 6) or invest differentially in assets (Table 4).

All treated participants, regardless of saturation status, are more likely than the control group to *ever* sell goods at another town's market (Table 5) but the coefficient for the high-saturation clusters is sizable and negative though just shy of statistical significance. This may be a mediating response to increased competition in high saturation clusters and would explain why the spillover effects are not stronger. I find no effects on further away trips, longer travel in response to the treatment. The strongest indicator of competitive pressure is that households in high-saturation clusters are more likely to send an adult member permanently away or move with the household as a whole. The most common place for businesses to sell on a day-to-day basis, however, remains nearby (Figure 7).

A key concern for the validity of my results are spillovers across clusters. As discussed in section 2, household clusters are buffered in space in order to avoid spillovers across them. I now look for evidence of such spillovers. The presence of non-treated households in low-saturation clusters allows me to test for another type of spillovers from treated onto control households, by comparing them to pure control clusters. This is not a direct test of across-cluster spillovers. If one is willing to posit, however, that spillovers should generally decay with distance, then finding few spillovers onto immediate untreated neighbors from the treatment should be reassuring that there are no spillovers across clusters either. I test this using the pre-specified regression equation

$$y_i = \beta_0 + \beta_T \text{Treatment}_i + \beta_S \text{Spillover}_i + \beta_{BL} y_i^{BL} + \beta_M \mathbb{1}[y_i^{BL} \text{ missing}] + \varepsilon_i.$$

Here, Spillover is an indicator for being a control household in a low-saturation cluster and coefficient β_S compares these households directly to control group households in pure control clusters. Results in Table 6 show relatively little evidence of such spillovers onto non-treated households who are much less likely to run competitor firms to the treated.

It does not appear as though the larger cluster size deteriorated training quality significantly. Table 7 shows that the high-saturation clusters score consistently (but not statistically significantly) lower than their low-saturation counterparts, though only by less than 1 point out of 100. I also verify whether the coordination intervention is reflected in the test scores by investigating whether coordination-related themes are mentioned by program participants. Those who received the coordination intervention were indeed more likely to mention being aware of one's competitor's weaknesses as factors that contribute to business success, though they are no more likely to say that one should find one's own niche and there is no difference in whether they highlight the presence

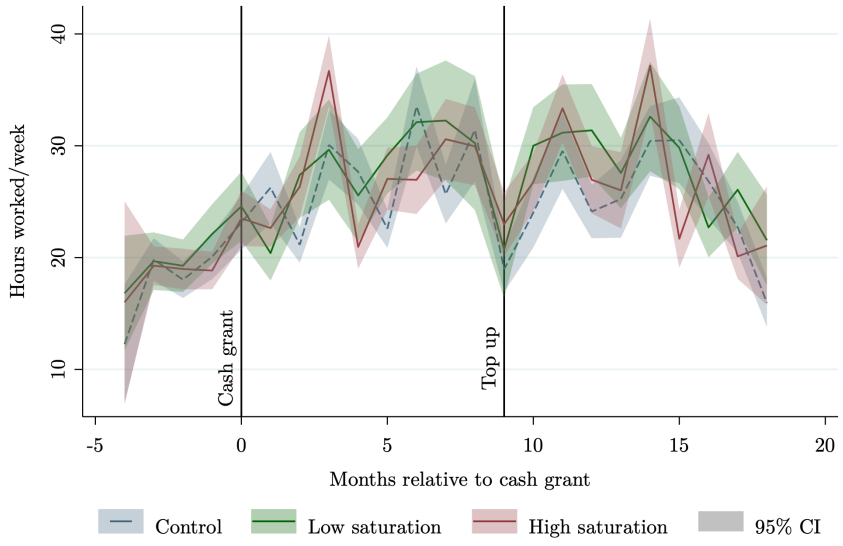


Figure 6: Hours worked by saturation

Note: Hours worked per week are not affected by treatment.

Table 4: Assets owned

	(1) Control mean (SD)	(2) Treatment Effect	(3) High Saturation	(4) N (BL)
Total assets (MWK)	391,459	170,593***	22,104	1214
	(473,494)	(26,692)	(37,693)	✓
		[0]	[.5984]	
Business assets (MWK)	25,543	49,496**	38,951*	1214
	(114,443)	(17,258)	(21,047)	✓
		[.02098]	[.09091]	
Agricultural Assets (MWK)	14,960	8,907***	-5,802*	1214
	(19,353)	(2,539)	(2,909)	✓
		[.004995]	[.06793]	
Personal assets (MWK)	341,534	117,109***	-12,298	1214
	(387,610)	(25,032)	(26,639)	✓
		[.000999]	[.6683]	

Notes: The last column indicates the number of observations. All regressions are implemented as an ANCOVA specification, controlling for strata fixed effects. A ✓ indicates that the outcome variable is available for all observations at baseline. If it was not collected at baseline, a × is shown, and if it is partially available, missing observations are set to zero and dummied out (✓). Where available, regressions control for baseline levels of the outcome. Standard errors are clustered at the strata level. Fisher (1935) exact test p-values from 1000 simulated alternative treatment assignments are shown in brackets. Significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Mobility

	(1) Control mean (SD)	(2) Treatment Effect	(3) High Saturation	(4) N (BL)
Sell at local market	.3183	.1853***	-.09576	1214
	(.4662)	(.05113)	(.06305)	✓
		[.001998]	[.1708]	
Sell elsewhere	.2663	.261***	-.0371	1214
	(.4424)	(.04984)	(.04665)	✓
		[0]	[.4396]	
Minutes to market	56.72	2.54	5.38	1213
	(52.91)	(4.875)	(7.104)	✓
		[.6374]	[.4885]	
Adults left HH	.2613	-.03173	.1113**	1214
	(.612)	(.03974)	(.04364)	✓
		[.4356]	[.02597]	
HH left study area	.0609	-.04647**	.02163**	1248
	(.2393)	(.01703)	(.008637)	×
		[.01598]	[.02897]	
Times left district	.4774	1.044	-.7362	1214
	(2.098)	(.6224)	(.6285)	✓
		[.1009]	[.2587]	
Traveled to Lilongwe	.1022	.06059	-.01138	1214
	(.9561)	(.09984)	(.1059)	✓
		[.6384]	[.9421]	
∼ Blantyre	.1139	-.07932	.09273	1214
	(1.276)	(.05907)	(.05423)	✓
		[.2388]	[.05594]	
∼ Mangochi town	.3903	.5313***	.3821	1214
	(1.381)	(.1576)	(.3918)	✓
		[.003996]	[.3916]	

Notes: The last column indicates the number of observations. All regressions are implemented as an ANCOVA specification, controlling for strata fixed effects. A ✓ indicates that the outcome variable is available for all observations at baseline. If it was not collected at baseline, a × is shown, and if it is partially available, missing observations are set to zero and dummied out (✓). Where available, regressions control for baseline levels of the outcome. Standard errors are clustered at the strata level. Fisher (1935) exact test p-values from 1000 simulated alternative treatment assignments are shown in brackets. Significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Spillovers onto untreated households

	(1) Control mean (SD)	(2) Treatment Effect	(3) Spillover Control	(4) N (BL)
Does business	.6851	.2096***	-.03621	1214
	(.4649)	(.03662)	(.04995)	✓
Casual labor, charity	.804	-.167***	.03153	1214
	(.3973)	(.03578)	(.03996)	✓
Farming, fishing, forestry	.6131	-.08689**	-.006033	1214
	(.4875)	(.04058)	(.05075)	✓
Manufacturing	.1424	-.01298	-.02965	1214
	(.3497)	(.02128)	(.02125)	✓
Non-retail services	.1457	.1022***	-.02338	1214
	(.3531)	(.02086)	(.02668)	✓
Retail	.1977	.4538***	.0491	1214
	(.3986)	(.03969)	(.03716)	✓
HH income (12 months, MWK, winsorized)	664,083	384,265***	-80,732	1213
	(647,484)	(62,354)	(57,500)	✓
Total assets (MWK)	391,459	169,153***	-48,596*	1214
	(473,494)	(33,447)	(22,897)	✓
Business assets (MWK)	25,543	75,036***	-836.2	1214
	(114,443)	(18,611)	(8,842)	✓
Agricultural Assets (MWK)	14,960	4,160**	-2,721	1214
	(19,353)	(1,691)	(2,055)	✓
Personal assets (MWK)	341,534	95,330***	-41,254*	1214
	(387,610)	(23,894)	(23,101)	✓
Any livestock	.4941	.4076***	-.04251	1214
	(.5004)	(.0466)	(.03304)	✓
Livestock value (MWK)	41,333	70,479***	-4,984	1214
	(91,642)	(10,588)	(5,338)	✓
Any savings	.3166	.4685***	-.0777*	1214
	(.4655)	(.0396)	(.03802)	✓
Total savings (MWK)	12,791	59,421***	-131.8	1214
	(32,717)	(8,584)	(2,936)	✓

Notes: The last column indicates the number of observations. All regressions are implemented as an ANCOVA specification, controlling for strata fixed effects. A ✓ indicates that the outcome variable is available for all observations at baseline. If it was not collected at baseline, a × is shown, and if it is partially available, missing observations are set to zero and dummied out (✓). Where available, regressions control for baseline levels of the outcome. Standard errors are clustered at the strata level. Significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

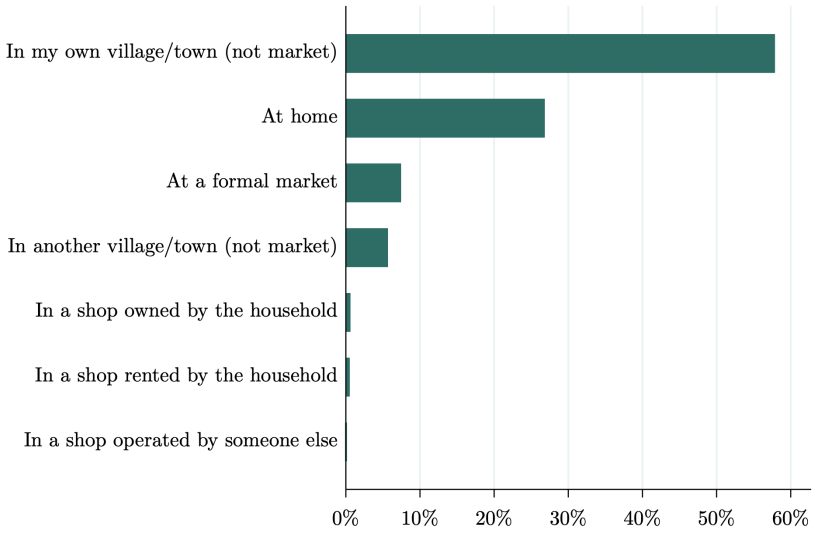


Figure 7: Location of most business sales

Note: The fact that most business sales happen within the same village or right in front of an entrepreneur's house reduces across-cluster spillover concerns through markets.

of competitors (Table 8).

Table 7: Post-training evaluation results

	(1) Treatment mean (SD)	(2) High Saturation	(3) N (BL)
Business knowledge	.4562	-.005078	923
	(.16)	(.01237)	✓
ECD knowledge	.5212	-.008697	923
	(.1347)	(.00958)	✓
Preventive health knowledge	.6042	-.009652	923
	(.2078)	(.01519)	✓

Notes: Analysis in this table is among treated households (who wrote a business plan) and the treatment coefficient is therefore omitted. The last column indicates the number of observations. All regressions are implemented as an ANCOVA specification, controlling for strata fixed effects. A ✓ indicates that the outcome variable is available for all observations at baseline. If it was not collected at baseline, a × is shown, and if it is partially available, missing observations are set to zero and dummied out (✓). Where available, regressions control for baseline levels of the outcome. Standard errors are clustered at the household cluster level. Significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9 splits the sample by whether households were engaged in retail businesses at baseline (column (3)) and those who were not (column (2)). It reveals that prior experience correlates with improvements in every reported metric of business performance. While not all of the differences

Table 8: "Name factors that contribute to business success."

	(1) Treatment mean (SD)	(2) Coordination Intervention	(3) N (BL)
Competitor weaknesses	.3424	.1081***	923
	(.4748)	(.03739)	✓
Business niche	.623	.0213	923
	(.4849)	(.03017)	✓
Aware of competitors	.4875	-.01049	923
	(.5001)	(.03557)	✓

Notes: Analysis in this table is among treated households (who wrote a business plan) and the treatment coefficient is therefore omitted. The last column indicates the number of observations. All regressions are implemented as an ANCOVA specification, controlling for strata fixed effects. A ✓ indicates that the outcome variable is available for all observations at baseline. If it was not collected at baseline, a × is shown, and if it is partially available, missing observations are set to zero and dummied out (✓). Where available, regressions control for baseline levels of the outcome. Standard errors are clustered at the strata level. Significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

are statistically significant (at the current sample size), they all point in the same direction.

4.3 Discussion

A broad but perhaps under-appreciated pattern in structural transformation is that the retail sector is the first livelihood outside of agriculture for many extremely poor people. My paper shows that while recipients express general equilibrium neglect in how they chose their businesses, this cannot have been the only barrier to broader sectoral diversification. The lack of hard skills emerges as a likely next bottleneck.

There are first signs of a divergence in earnings a few months after the cash grant, which come into sharper relief when focusing on business survival and exit. The point estimates on earnings reductions and exit are roughly in line with the cost savings for the implementer from scale economies. A particularly stark sign of overcrowding is that the higher saturation arm induces significant out-migration.

Prices do not appear to be systematically affected by saturation, suggesting that increased demand and supply either cancel each other out or the lower earnings come about by spreading a limited number of sales at fixed prices among more entrepreneurs who are already at a competitive benchmark.

On the whole, these findings cast doubt on the ability to scale an entrepreneurship-focused graduation model through saturation in contexts where a large share of the population is eligible, and without removing barriers to entry into a broader range of sectors. This, however, would require innovation in hard skills training delivery that can provide different skills to different people and

Table 9: Business performance by prior experience

	(1) Control mean (SD)	(2) Treatment Effect (0)	(3) Treatment Effect (1)	(4) Difference p-value	(5) N (BL)
Does business	.4722	.3756***	.4377***	[.0466]**	995
	(.3505)	(.04547)	(.05059)		✓
		[0]	[0]		
Investment in inventory	216,500	1,234,756***	2,022,983***	[.006641]***	995
	(506,077)	(191,062)	(241,990)		×
		[0]	[0]		
∼ productive assets	11,874	44,695***	54,868***	[.2495]	995
	(105,511)	(12,042)	(15,828)		×
		[0]	[.002997]		
Hours worked	257	10.65	36.39	[.05719]*	995
	(163.4)	(19.96)	(23.92)		×
		[.6424]	[.3257]		
Earnings	723,816	1,001,482***	1,658,562***	[.02288]**	995
	(707,048)	(263,836)	(365,164)		×
		[0]	[0]		
Gross margin	506,909	-240,344	-377,880	[.4223]	995
	(482,403)	(163,984)	(244,955)		×
		[.1568]	[.1608]		
Profitable	.9588	-.3198***	-.3813***	[.09847]*	995
	(.1991)	(.07384)	(.07585)		×
		[0]	[0]		

Notes: The last column indicates the number of observations. All regressions are implemented as an ANCOVA specification, controlling for strata fixed effects. A ✓ indicates that the outcome variable is available for all observations at baseline. If it was not collected at baseline, a × is shown, and if it is partially available, missing observations are set to zero and dummied out (✓). Where available, regressions control for baseline levels of the outcome. Standard errors are clustered at the strata level. Fisher (1935) exact test p-values from 1000 simulated alternative treatment assignments are shown in brackets. Significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

thus may not attain scale economies. Apprenticeship programs are an avenue worth exploring in future research.

References

Aggarwal, Shilpa, Jenny Aker, Dahyeon Jeong, Naresh Kumar, David Sungho Park, Jonathan Robinson, and Alan Spearot, "The dynamic effects of cash transfers: Evidence from Rural Liberia and Malawi," 2022.

Aizer, Anna, Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney, "The long-run impact of cash transfers to poor families," *American Economic Review*, 2016, *106* (4), 935–71.

Altındağ, Onur and Stephen D O'Connell, "The short-lived effects of unconditional cash transfers to refugees," *Journal of Development Economics*, 2023, *160*, 102942.

Araujo, Maria Caridad, Mariano Bosch, and Norbert Schady, "Can Cash Transfers Help Households Escape an Inter-Generational Poverty Trap?," *The Economics of Poverty Traps*, 2019, pp. 357–382.

Bai, Yuehao, "Optimality of matched-pair designs in randomized controlled trials," *American Economic Review*, 2022, *112* (12), 3911–3940.

Baird, Sarah, Craig McIntosh, and Berk Özler, "Cash or condition? Evidence from a cash transfer experiment," *The Quarterly Journal of Economics*, 2011, *126* (4), 1709–1753.

_ , J Aislinn Bohren, Craig McIntosh, and Berk Özler, "Optimal design of experiments in the presence of interference," *Review of Economics and Statistics*, 2018, *100* (5), 844–860.

Balboni, Clare, Oriana Bandiera, Robin Burgess, Maitreesh Ghatak, and Anton Heil, "Why do people stay poor?," *The Quarterly Journal of Economics*, 2022, *137* (2), 785–844.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman, "Labor markets and poverty in village economies," *The Quarterly Journal of Economics*, 2017, *132* (2), 811–870.

Banerjee, Abhijit, Emily Breza, Esther Duflo, and Cynthia Kinnan, "Can microfinance unlock a poverty trap for some entrepreneurs?," Technical Report, National Bureau of Economic Research 2025.

_ , Esther Duflo, and Garima Sharma, "Long-term effects of the targeting the ultra poor program," *American Economic Review: Insights*, 2021, *3* (4), 471–86.

_ , _ , Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry, "A multifaceted program causes lasting progress for the very poor: Evidence from six countries," *Science*, 2015, *348* (6236), 1260799.

Bauchet, Jonathan, Jonathan Morduch, and Shamika Ravi, "Failure vs. displacement: Why an innovative anti-poverty program showed no net impact in South India," *Journal of Development Economics*, 2015, *116*, 1–16.

Bedoya, Guadalupe, Aidan Coville, Johannes Haushofer, Mohammad Isaqzadeh, and Jeremy P Shapiro, "No household left behind: Afghanistan targeting the ultra poor impact evaluation," Technical Report, National Bureau of Economic Research 2019.

Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, and Bertil Tungodden, "Human and financial capital for microenterprise development: Evidence from a field and lab experiment," *Management Science*, 2015, *61* (4), 707–722.

Blattman, Christopher and Laura Ralston, "Generating employment in poor and fragile states: Evidence from labor market and entrepreneurship programs," *Available at SSRN 2622220*, 2015.

_ , Nathan Fiala, and Sebastian Martinez, "Generating skilled self-employment in developing countries: Experimental evidence from Uganda," *The Quarterly Journal of Economics*, 2014, *129* (2), 697–752.

Brooks, Wyatt, Kevin Donovan, and Terence R Johnson, "Mentors or teachers? Microenterprise training in Kenya," *American Economic Journal: Applied Economics*, 2018, *10* (4), 196–221.

_ , _ , _ , and Jackline Oluoch-Aridi, "Cash transfers as a response to COVID-19: Experimental evidence from Kenya," *Journal of Development Economics*, 2022, *158*, 102929.

Bruhn, Miriam and David McKenzie, "In pursuit of balance: Randomization in practice in development field experiments," *American Economic Journal: Applied Economics*, 2009, *1* (4), 200–232.

Bryan, Gharad, Shyamal Chowdhury, Ahmed Mushfiq Mobarak, Melanie Morten, and Joeri Smits, "Encouragement and distortionary effects of conditional cash transfers," 2021.

Cahyadi, Nur, Rema Hanna, Benjamin A Olken, Rizal Adi Prima, Elan Satriawan, and Ekki Syamsulhakim, "Cumulative impacts of conditional cash transfer programs: Experimental evidence from Indonesia," *American Economic Journal: Economic Policy*, 2020, *12* (4), 88–110.

Chaisemartin, Clément De and Jaime Ramirez-Cuellar, "At what level should one cluster standard errors in paired and small-strata experiments?," *American Economic Journal: Applied Economics*, 2024, *16* (1), 193–212.

Chowdhury, Reajul, Elliott Collins, Ethan Ligon, and Munshi Sulaiman, "Valuing Assets Provided to Low-Income Households in South Sudan," 2017.

Cunha, Jesse M, Giacomo De Giorgi, and Seema Jayachandran, "The price effects of cash versus in-kind transfers," *The Review of Economic Studies*, 2019, *86* (1), 240–281.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar, "Keeping it simple: Financial literacy and rules of thumb," *American Economic Journal: Applied Economics*, 2014, *6* (2), 1–31.

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael Walker, "General equilibrium effects of cash transfers: Experimental evidence from Kenya," *Econometrica*, 2022, *90* (6), 2603–2643.

Fiala, Nathan, Julian Rose, Filder Aryemo, and Jörg Peters, "The (very) long-run impacts of cash grants during a crisis," 2022, (961).

Filmer, Deon, Jed Friedman, Eeshani Kandpal, and Junko Onishi, "Cash transfers, food prices, and nutrition impacts on ineligible children," *The Review of Economics and Statistics*, 2021, pp. 1–45.

Fisher, Ronald Aylmer, *The design of experiments*, Edinburgh: Oliver and Boyd, 1935.

Gertler, Paul J, Sebastian W Martinez, and Marta Rubio-Codina, "Investing cash transfers to raise long-term living standards," *American Economic Journal: Applied Economics*, 2012, *4* (1), 164–92.

Giné, Xavier and Ghazala Mansuri, "Money or management? A field experiment on constraints to entrepreneurship in rural Pakistan," *Economic Development and Cultural Change*, 2021, *70* (1), 41–86.

Goldberg, Pinelopi Koujianou and Tristan Reed, "Presidential Address: Demand-Side Constraints in Development. The Role of Market Size, Trade, and (In)Equality," *Econometrica*, 2023, *91* (6), 1915–1950.

Hardy, Morgan and Gisella Kagy, "It's getting crowded in here: experimental evidence of demand constraints in the gender profit gap," *The Economic Journal*, 2020, *130* (631), 2272–2290.

Haushofer, Johannes and Jeremy Shapiro, "The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya," *The Quarterly Journal of Economics*, 2016, *131* (4), 1973–2042.

_ and _ , "The long-term impact of unconditional cash transfers: Experimental evidence from Kenya," *Busara Center for Behavioral Economics, Nairobi, Kenya*, 2018.

_ , Paul Niehaus, Carlos Paramo, Edward Miguel, and Michael W Walker, "Targeting impact versus deprivation," Technical Report, National Bureau of Economic Research 2022.

Hussam, Reshmaan, Natalia Rigol, and Benjamin N Roth, "Targeting high ability entrepreneurs using community information: Mechanism design in the field," *American Economic Review*, 2022, *112* (3), 861–898.

Ikegami, Munenobu, Michael R Carter, Christopher B Barrett, and Sarah Janzen, "Poverty traps and the social protection paradox," *The economics of poverty traps*, 2017, pp. 223–256.

Kansikas, Carolina, Anandi Mani, and Paul Niehaus, "Customized cash transfers: Financial lives and cash-flow preferences in rural Kenya," 2023.

Leight, Jessica, Daniel O Gilligan, Michael Mulford, Alemayehu Seyoum Taffesse, and Heleene Tambet, "Aspiring to more? New evidence on the effect of a light-touch aspirations intervention in rural Ethiopia," Technical Report 2021.

_ , Josué Awonon, Abdoulaye Pedehombga, Rasmané Ganaba, and Aulo Gelli, "How light is too light touch: The effect of a short training-based intervention on household poultry production in Burkina Faso," *Journal of Development Economics*, 2022, *155*, 102776.

McKenzie, David and Susana Puerto, "Growing markets through business training for female entrepreneurs: A market-level randomized experiment in Kenya," *American Economic Journal: Applied Economics*, 2021, *13* (2), 297–332.

Mel, Suresh De, David McKenzie, and Christopher Woodruff, "One-time transfers of cash or capital have long-lasting effects on microenterprises in Sri Lanka," *Science*, 2012, *335* (6071), 962–966.

Millán, Teresa Molina, Karen Macours, John A Maluccio, and Luis Tejerina, "Experimental long-term effects of early-childhood and school-age exposure to a conditional cash transfer program," *Journal of Development Economics*, 2020, *143*, 102385.

Misha, Farzana A, Wameq A Raza, Jinnat Ara, and Ellen Van de Poel, "How far does a big push really push? Long-term effects of an asset transfer program on employment trajectories," *Economic Development and Cultural Change*, 2019, *68* (1), 41–62.

Morgan, Kari Lock and Donald B Rubin, "Rerandomization to improve covariate balance in experiments," *The Annals of Statistics*, 2012, *40* (2), 1263–1282.

Muralidharan, Karthik, Mauricio Romero, and Kaspar Wüthrich, "Factorial designs, model selection, and (incorrect) inference in randomized experiments," *Review of Economics and Statistics*, 2023, pp. 1–44.

Ritzwoller, David M, Joseph P Romano, and Azeem M Shaikh, "Randomization inference: Theory and applications," *Journal of Political Economy: Microeconomics*, 2025.

Rosenstein-Rodan, Paul N., "Notes on the theory of the 'big push'," in "Economic Development for Latin America: Proceedings of a conference held by the International Economic Association" Springer 1957, pp. 57–81.

Schreiner, Mark, "Simple Poverty Scorecard® Tool Malawi," 2019.

Sedlmayr, Richard, Anuj Shah, and Munshi Sulaiman, "Cash-plus: Poverty impacts of alternative transfer-based approaches," *Journal of Development Economics*, 2020, *144*, 102418.

Sulaiman, Munshi, Nathanael Goldberg, Dean Karlan, and Aude de Montesquiou, "Eliminating extreme poverty: Comparing the cost-effectiveness of livelihood, cash transfer, and graduation approaches," 2016.

Valdivia, Martin, "Business training plus for female entrepreneurship? Short and medium-term experimental evidence from Peru," *Journal of Development Economics*, 2015, *113*, 33–51.

Wydick, Bruce, "When Are Cash Transfers Transformative?," 2020.

Young, Alwyn, "Channeling Fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results," *The Quarterly Journal of Economics*, 2019, *134* (2), 557–598.

Appendix

A Additional figures and tables

Table A1: Study population

	(1) Count
HOUSEHOLDS	
Number of surveys conducted	28,853
Female primary adult present	27,684
U5 present	16,882
No children present	3,254
INDIVIDUALS	
Population	138,548
Adults	65,162
Children	73,386
U5s	20,998
Number of villages	162
Number of CBOs	25
Number of ADCs	5

Census of all households in Mangochi district's subdistrict ("Traditional Authority") Nankumba. The urban center Monkey Bay and the tourist destination Cape Maclear were omitted from this census as well as an area in the South-West of Nankumba where another NGO is implementing a livelihood program. Area Development Committees (ADCs) are administrative subdivisions of a subdistrict, while Community-Based Organizations (CBOs) are associations of several villages organizing development initiatives or providing local public goods together.

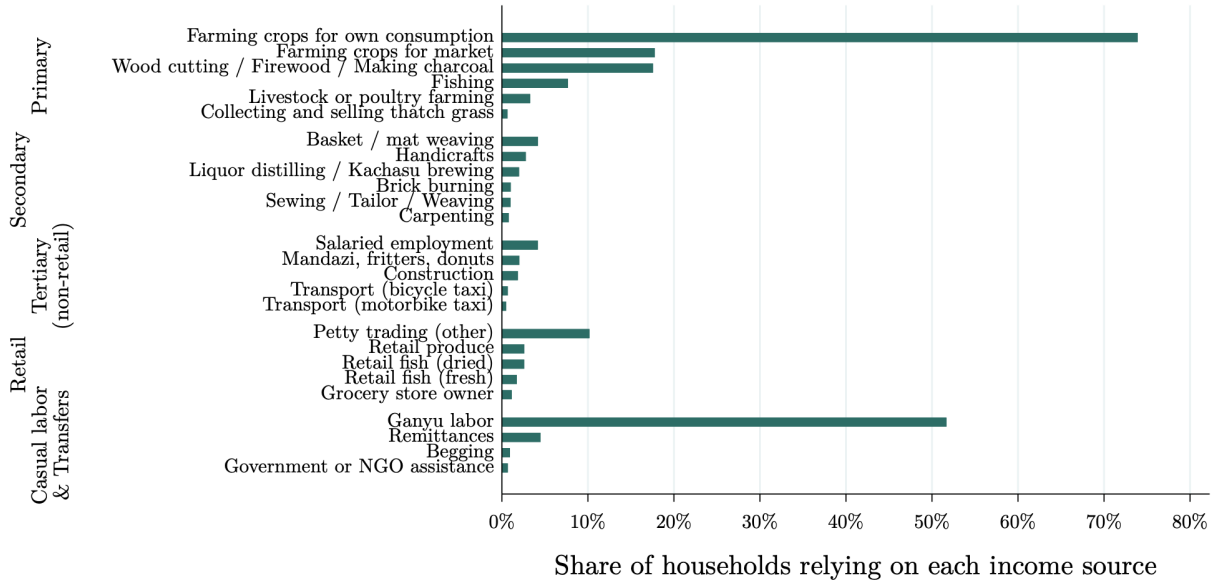


Figure A1: Most common income sources in the study area

Note: This figure describes the income sources of the population as stated during the area-wide census. People are predominantly farmers or casual "'ganyu" laborers. Outside of the primary sector, retail is another common income source.

Table A2: Livelihoods

	(1) Percent
Agriculture	
Crop farming	76.5
Subsistence crop farming	73.5
Marketed crop farming	17.8
Share of yield marketed	29.2
Share of yield marketed (conditional)	69.8
Livestock	3.3
Piece work	
Ganyu	51.7
Exclusively subsistence + ganyu	27.6
Business	
Business	58.0
Employees > 0	4.1
Local non-ag employment potential	3.4
Observations	29084

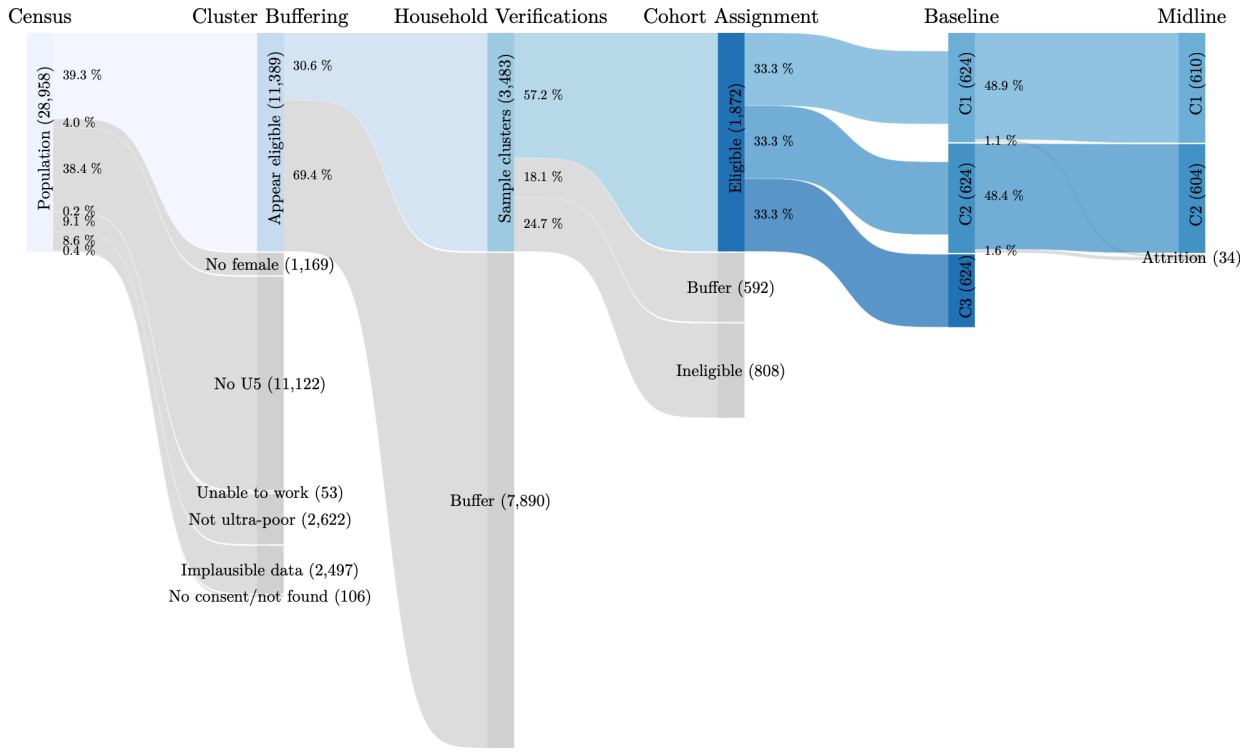


Figure A2: **Sampling funnel**

Note: This figure describes how the sample is selected from the census population. Starting from the census population of about 29,000 households, about 60% were deemed ineligible, mostly for lack of a child under age five. Among apparently eligible households, 271 clusters of 40-50 households were formed by proximity using k-means clustering. A second custom algorithm then selected 72 of these clusters in such a way as to maximize the distance between the nearest two clusters, in order to minimize spillover concerns (cf. Appendix Figure A3). The other clusters serve as buffer zones. Among the selected clusters, household verifications were conducted that resulted in some households being deemed ineligible under the same criteria as in the census, while selecting the 26 eligible households closest to the cluster centroid and excluding the remainder as additional buffer households on the periphery of the clusters. The 72 clusters were grouped into triplets by proximity and assigned an implementing staff. They were also assigned at random to three cohorts to undergo the intervention at different times in the agricultural season. The full sample was surveyed at baseline and the graph shows attrition in the subsequent survey rounds.

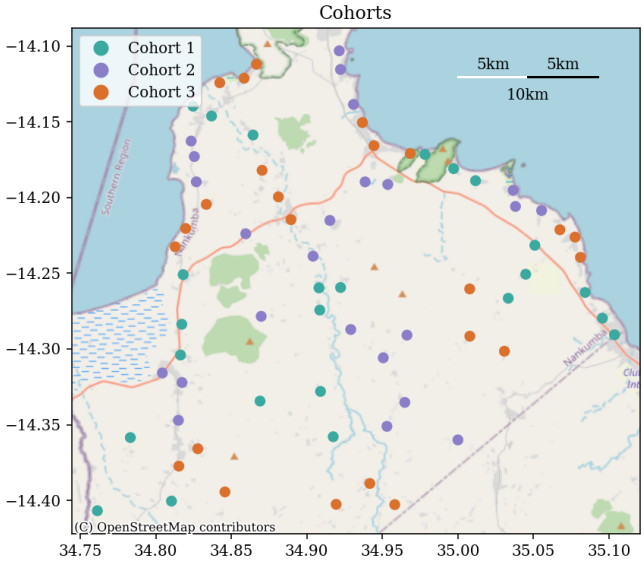


Figure A3: **Clusters** by cohort

Note: This map shows the 72 study clusters of 26 households each that make up the study sample, color-coded by the cohort they are assigned to.

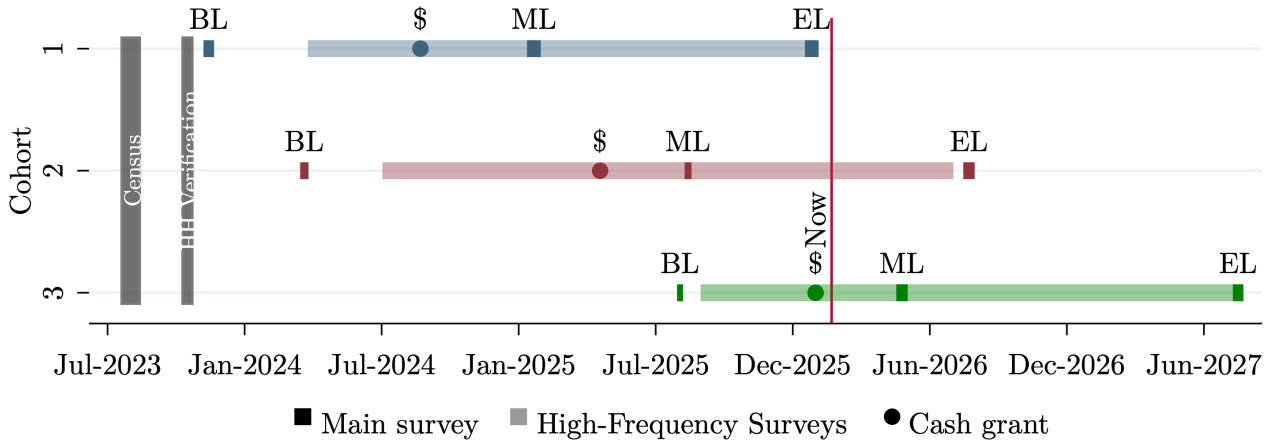


Figure A4: Study timeline

Note: This figure describes the study timeline. A census of the study area was conducted in August 2023. Household eligibility verifications followed in October 2023. The baseline survey for cohort 1 was conducted in December 2023 and high-frequency surveys in this cohort started in April 2024. The large one-off cash grant was paid out in August 2024. The midline survey was collected in January 2025, followed a year later by the endline survey.

Table A3: Poverty

	(1) Percent
Skip meal (past two weeks)	76.8
Female illiterate	37.7
Thatch roof	64.5
No furniture	80.3
No phone	51.9
No bedding	15.7
<2.15 PPP 2017 USD (self report)	92.7
<2.15 PPP 2017 USD (PMT)	85.1
# children / adult (if > 0)	1.5
# income sources	1.5

Table A4: Balance

	(1) Control Mean (SD)	(2) Treatment (SE)
HOUSEHOLD AND PARTICIPANT CHARACTERISTICS		
Household size	5.4	-.1175
	(1.8)	(.09726)
Age	32	-.5641
	(8.4)	(.4244)
Years of education	5.7	-.1218
	(2.9)	(.1906)
Literate	.69	.005342
	(.46)	(.02665)
LIVELIHOOD AND ASSETS		
Hours worked (7 days)	22	-.3643
	(18)	(.7828)
HH income (12 months, MWK)	475,629	11,315
	(480,571)	(47,515)
Total savings, assets, livestock, minus debt (MWK)	256,874	-17,471
	(365,607)	(24,353)
Largest purchase (12 months, MWK)	38,607	-2,816
	(121,528)	(6,344)
POVERTY AND FOOD SECURITY		
Thatch / tarp / no roof	.75	.02137
	(.44)	(.02314)
Food insecurity score (0-8)	7.5	.04915
	(1.3)	(.0641)
HEALTH AND NON-HEALTH SHOCKS		
Health episodes (12 months)	16	-.578
	(13)	(.9197)
Non-health shocks (12 months)	21	-.3771
	(18)	(.8531)
NOT PRE-SPECIFIED		
Entrepreneurship knowledge score (0-100)	19	.1714
	(7)	(.4259)
Engaged in retail at baseline	.18	-.007479
	(.38)	(.0236)
Observations		1,872
F		0.58
$p(F = 0)$		.85

This table describes the baseline balance. The main panel reports the control group means and coefficients of the treatment indicator in a regression predicting each baseline outcome variable. At the bottom of the table, an F statistic is reported for a regression in which all outcomes jointly predict the treatment. Only the pre-specified variables are included in the F-test. Standard errors are clustered at the household cluster level and reported in parentheses. Significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.