Why Do People Stay Poor?

Clare Balboni, Oriana Bandiera, Robin Burgess, Maitreesh Ghatak and Anton Heil*

March 10, 2020

Abstract

There are two views as to why people stay poor. The equal opportunity view emphasizes that differences in individual traits like talent or motivation make the poor choose low productivity jobs. The poverty traps view emphasizes that access to opportunities depends on initial wealth and hence poor people have no choice but to work in low productivity jobs. We test the two views using the random allocation of an asset transfer program that gave some of the poorest women in Bangladesh access to the same job opportunities as their wealthier counterparts in the same villages. The data rejects the null of equal opportunities. Exploiting small variation in initial endowments, we estimate the transition equation and find that, if the program pushes individuals above a threshold level of initial assets, then they escape poverty, but, if it does not, they slide back into poverty. Structural estimation of an occupational choice model reveals that almost all beneficiaries are misallocated at baseline and that the gains arising from eliminating misallocation would far exceed the costs. Our findings imply that large one-off transfers that enable people to take on more productive occupations can help alleviate persistent poverty.

*We thank Abhijit Banerjee, Supreet Kaur, Michael Kremer, Ben Moll, Víctor Quintas-Martínez, Alejandro Sabal-Bermúdez, Alwyn Young and several seminar participants for comments. Ghatak would like to thank the ESRC-DFID Grant ES/L012103/1 for financial support. All errors are our own.

1 Introduction

Why do people stay poor? This is one of the key questions within economics. Finding solutions to the mass poverty problem is what motivated early contributors to development economics (Lewis, 1954; Myrdal, 1968; Schultz, 1980) and what continues to motivate the current generations. It is also the central goal of development policy — the Sustainable Development Goal signed up to by the majority of the worlds governments is to "eradicate extreme poverty for all people everywhere by 2030". Given that in 2015, when these goals were set, 10% of the World's population (734.5 million people) was classified as living in extreme poverty, this is an ambitious objective.[1] Finding answers ultimately requires us to understand whether and why people stay poor.

Most of the world's poor are employed but have low earnings, so to understand why they stay poor we must understand why they work in low-earning jobs. One view is that the poor have the same opportunities as everyone else, so if they work in low-earning jobs they must have traits that make them unsuitable for other occupations. The alternative view is that the poor face different opportunities and hence do low-earning jobs because they are born poor. That is, the poor are stuck in a poverty trap. Scaled up at the macro level, the two views underpin growth models with convergence (Solow (1956)) or with multiple steady states (Rosenstein-Rodan, 1943; Nurkse, 1961; Myrdal, 1957; Myrdal, 1968; Rostow, 1960).[2]

Distinguishing between these two views is as important as it is difficult. It is important because they have dramatically different policy implications - if people are in a poverty trap then big push polices which move them into more productive forms of employment might offer a permanent solution to the global poverty problem (Murphy, Shleifer, and Vishny (1989) and Hirschman (1958)). It is difficult because both explanations produce outcomes that are observationally equivalent and indeed it has been remarkably hard to empirically identify poverty traps.[3]

[1] Atamanov et al. (2019).

[2] See, for example, Dasgupta and Ray (1986), Dasgupta (1997), Ray and Streufert (1993), Banerjee and Newman (1993), and Galor and Zeira (1993) for the modern literature on poverty traps in the context of development economics. One can also have aggregate poverty traps where there is interaction among individuals through prices (such as wages) and one can have multiple *stationary* steady states. For example, the presence of a lot of individuals who are poor can depress the wage rate and dampen mobility (see Banerjee and Newman, 1993; Galor and Zeira, 1993). Aggregate poverty traps can arise even in absence of individual poverty traps, so that no single individual is trapped in poverty and can move up and down, but the economy can have different stationary distributions of capital or income that differ in terms of the fraction of population that is poor.

[3] Existing evidence fits poverty traps models to country-level data (see Easterly (2006)) or tests individual assumptions of poverty trap models Kraay and McKenzie (2014).

The main contribution of this paper is to provide an empirical test for the existence of poverty traps using individual-level data. Our setting are village economies situated in the poorest districts of Bangladesh. The occupational structure of these villages is very simple and correlated with asset ownership. Those who own land or livestock combine it with their labor and hire those who do not on a casual basis. Land cultivation and livestock rearing yield higher earnings than casual labor. The distribution of productive assets therefore is bimodal.

The question is whether the bimodality is symptomatic of a poverty trap, namely whether poor people do casual jobs and hold nearly no productive assets because they do not have the talent to do anything else or whether the fact that they are poor prevents them from acquiring the assets needed to climb the occupational ladder. The main problem in identifying poverty traps is that, by definition, the threshold is an unstable equilibrium so we normally do not observe anyone near it. Fortuitously our setting is an exception, as BRAC's Targeting the Ultrapoor program (Bandiera et al. (2017)) transfers large assets (cows) to the poorest women in these villages and the value of the transfer is such that it moves over 3,000 households from the low mode to the lowest density point of the asset distribution.

Tracing how their assets evolve after the transfer allows us to test for poverty traps. This is because the equal opportunity and the poverty traps views of poverty produce different transition equations. In the equal opportunity view the transition equation is continuous and concave, while in the poverty traps view it has a convex segment or a discontinuity. Since BRAC targeted the program at households without significant productive assets, there are small initial differences in asset ownership before the transfer. As the asset transfer moves beneficiaries out of their steady state, we can exploit these marginally different levels of productive assets to estimate the transition equation between capital after the transfer and capital four years later. We can then test for the null of concavity and, upon rejection, identify the capital threshold past which people can escape the poverty trap. We then test whether the change in capital around the threshold is consistent with the existence of a poverty trap and examine the mechanisms that underpin it.

Our main results are as follows. First, we reject the null of concavity. Rather, we find that the transition equation is S-shaped with a threshold level of capital at 2.333 log points. At this threshold, assets are worth 9,309 BDT (504 USD PPP).[4] For comparison, the median value of a cow in our sample is around 9,000 BDT (488 USD PPP). People whose baseline assets were so low that the transfer was not enough to bring them past the threshold slide

[4]Throughout, we use the 2007 PPP adjusted exchange rate of 18.46 BDT to one dollar.

back into poverty. These are about one-third of the sample and on average they lose 16% of their asset value (inclusive of the transfer) by year 4. In contrast, those who do go past the threshold keep accumulating assets year after year and have 14% more by year 4.

Second, we probe the validity of the two assumptions underlying our test: that variation in baseline capital is orthogonal with future changes in capital and that it is orthogonal to the response to the program. For the former we use data from control villages to compute the correlation between baseline capital and subsequent changes. We find that this negative throughout as is to be expected because of mean reversion. For the latter we exploit variation in savings rate and earnings from livestock rearing across villages to estimate group specific transition equations and thresholds for the same level of baseline capital. The fact that the coefficients remain stable reassures us that different patterns of accumulation above and below the threshold are not due to unobservables correlated with baseline capital or heterogeneous responses to the training component of the program.

The theoretical literature points to several alternative mechanisms behind poverty traps (see, for example, Dasgupta and Ray, 1986; Dasgupta, 1997; Galor and Zeira, 1993; Banerjee and Newman, 1993).[5] We rule out poverty traps based on the link between nutrition and productivity or on savings and borrowing. We then provide indicative evidence that the mechanism at play is the combination of technological indivisibilities and credit market imperfections that prevent poor people from buying the assets, specifically cows, that would allow them to escape the trap. We show that technological indivisibilities take the form of ownership of complementary assets (mostly vehicles needed to procure fodder or to sell livestock products).

Informed by this analysis, we go on to estimate a structural model of occupational choice that allows us to calibrate individual-specific values of productivity, disutility of livestock rearing, and disutility of wage work. Being able to do this is unusual, as typically people are only observed in the job they can do best, making it impossible to estimate their returns in alternative ones. Two features of our setting allow us to circumvent this problem. The first is that at baseline only 6% of beneficiaries owned cows, so nearly all were engaged in wage labor. The second is that, as part of the program, BRAC required participants to hold on to the asset for two years, and so we observe *everybody* tending livestock, at least for some time, in our data.[6] To validate the model we use the estimated parameters to predict hours worked in the different occupations after four years as a function of capital.

[5]For reviews of the literature, see Azariadis, 1996, Carter and Barrett, 2006, and Ghatak, 2015.

[6]After two years, they were allowed to liquidate the asset if they wished.

With these values in hand, we can estimate individual-specific misallocation of resources (occupational choice, capital, and labor) that we can use to evaluate different policy counterfactuals. Three findings are of note. First, the total value of misallocation per year in steady state, that is the sum of the gains that accrue to each beneficiary by changing their labor allocation across jobs, is 15 times larger than the one-off cost of taking them across the threshold. Second, general equilibrium effects might reduce the benefit-to-cost ratio, but returns to livestock rearing would have to fall by 89% to equalise the cost of eliminating the trap and the value of misallocation. Third, since ability is not correlated with initial wealth, small transfers have small effects.

To conclude, we use our estimates to compute the share of households that would be propelled past the threshold with different levels of transfers. Seen through the lens of poverty traps, several common policies such as microfinance and workfare programs are too small relative to the distance to the threshold for most households. This is in line with the finding that microfinance generally fails unless the borrowers already had a business as these are probably closer to their thresholds (see Banerjee et al., 2019, Banerjee et al., 2015a, Meager, 2019).

The rest of the paper proceeds as follows. Section 2 details the context of the intervention we study. Section 3 describes the reduced form tests we use to test for poverty traps and hence discern between the two views of why people stay poor. In Section 4 outline and estimate our structural model of occupational choice that allows us to quantify the extent of misallocation in the work that people do. Section 4. In Section 5 we draw out the key policy implications from our findings. Finally, Section 6 concludes.

2 Context

We test for the existence of a poverty trap using data collected to evaluate BRAC's Targeting the Ultra-poor Program in Northern Bangladesh (Bandiera et al., 2017). The data covers 23,000 households living in 1,309 villages in the 13 poorest districts in the country. Of these households, over 6,000 are considered extremely poor. The program offers a one-off transfer of productive assets and training with the aim of simultaneously relaxing credit and skill constraints to create a source of regular earnings for poor women who are mostly engaged in irregular and insecure casual labor.[7] Beneficiaries are offered to choose from several asset

[7]The program also includes consumption support in the first 40 weeks after the asset transfer, as well as health support and training on legal, social, and political rights in the two years following the program.

bundles, all of which are valued at around 560 USD in PPP and can be used for income-generating activities. Out of all eligible women, 91% chose an asset bundle containing a cow. BRAC encourages respondents to retain the asset for at least two years, after which they can liquidate it. To identify beneficiaries, BRAC runs a participatory wealth assessment exercise in every village. This yields a classification of households into three wealth classes (poor, middle and upper class) which forms our sampling frame. We survey all of the poor and 10% of the other classes in each village — a total sample size of just over 23,000 households. The group of poor households is further split into program eligibles (ultra-poor) and non-eligibles (other poor) according to BRACs eligibility criteria. The baseline survey was conducted before the intervention in 2007, and two follow-up surveys with the full sample in 2009 and 2011. Two further survey rounds in 2014 and 2018 collected data only on the households eligible for the program.

Table 1 describes the economic lives of the women in these villages by wealth class before the program was implemented in 2007. Panel A shows that labor force participation is nearly universal with rates above 80% in all wealth classes. However, poor women work more hours in fewer, longer days and earn much less, both in total and per hour worked. Panel B illustrates how differences in labor outcomes are correlated to differences in human and physical capital. Human capital is very low in these villages, and, while there are differences across classes, even the richest women have only 3.7 years of education on average and 49% of them are illiterate. By contrast, differences in physical capital (land, livestock, vehicles, etc.) are huge: the average upper class household owns 94 times more productive assets than the average poor household.[8]

Ownership of physical capital is what sets apart rich and poor in these villages. We measure physical capital as productive assets, which include poultry, livestock, tools, machines, vehicles, and land. We argue in this paper that ownership of physical capital is a crucial determinant of occupation and social class. As a first indication of this, consider the distribution of productive assets depicted in Figure 1a. The figure shows a kernel density estimate of the distribution of productive assets.[9] The distribution is bimodal, with a mass of households around 0.25 and 6.5, and hardly anyone in between. [10] Households in these village economies either own a lot of productive assets or almost none. Differences in asset ownership relate directly to differences in consumption. For example, households at the low

[8]Assets belong to the household rather than to the individual.

[9]Sampling weights are used to account for the different sampling probabilities of households across social classes.

[10]Henceforth, we report the value of productive assets in logs of 1,000 Bangladesh Taka (BDT).

mode with assets of less than 0.5 have an average annual per capita expenditure of 637 USD. For those at the high mode with assets between 6 and 7 this number is 1110 USD.

Richer households do not just own more assets, they also own more expensive assets. Figure 2a shows that the program beneficiaries, 85% of whom own assets valued less that 2 log points (7,390 BDT), own mostly poultry and goats, whilst their richer counterparts own cows and land. This ordering corresponds to the unit value of these assets. The median unit price of chickens and goats is 100 BDT and 1,000 BDT, respectively, while a typical cows costs around 9,000 BDT. The fact that people with more assets own more expensive assets rather than more of the same assets suggests that indivisibilities might be important: with imperfect rental markets it is not possible to obtain livestock or complementary inputs for a share of the time and the price. Furthermore, differences in asset composition give rise to differences in occupational choice. Figure 2b, shows how hours allocated to different occupations vary with the value of a household's productive assets. Casual employment in agriculture or domestic services prevail at low levels of productive assets while self-employment in livestock rearing and land cultivation gradually take over as the ownership of productive assets increases.

By transferring livestock the program thus gives the poorest women in these villages the opportunity to access the same jobs as their rich counterparts. It is key to note that this opportunity would not have arisen without the program. Indeed, in control villages, only 3% of the households that are poor at baseline reach the assets stock of a median middle class household within four years. The probability of catching up with the upper classes is therefore close to zero. This is thus a setting where the poor stay poor. The key question is whether this reflects differences in immutable characteristics such as talent for different occupations, or different access to capital. The next section illustrates how we can use responses to the program to test between the two views.

3 Reduced Form Tests

3.1 Framework

We present a simple framework to illustrate two ways in which the observed differences in asset holdings can be explained: differences in individual characteristics and asset dynamics that create a poverty trap. We then use this framework to test between the two views. We will see that the policy implications are radically different in the two cases.

As mentioned earlier, the notion of an individual poverty trap that we focus on is closely related to the dynamics of capital accumulation. To formalize this notion in a general way, define the transition equation as the function that relates individual i's capital stock across two time periods:

$$K_{i,t+1} = \Phi_i(K_{i,t})$$

where $K_{i,t}$ denotes i's capital, or productive assets, at time t. To fix ideas, assume that individual i in village v generates earnings according to $Y_i = A_{iv} f(K_i)$, where $f(\cdot)$ is the production function[11] and A_{iv} captures all immutable traits—either of individuals or the village—that determine productivity. Let s_i denote the individual's savings rate and δ a common rate of depreciation. In this special case, the transition equation can be expressed as [12]

$$\Phi_i(K_{i,t}) = s_i A_{iv} f\left(K_{i,t}\right) + \left(1 - \delta\right) K_{i,t} \tag{1}$$

To capture the idea of persistence, define a steady state as a fixed point of $\Phi_i(\cdot)$, that is a level of capital, K_i^*, such that $K_i^* = \Phi_i(K_i^*)$. In the above example, this is a point where the amount of savings exactly offsets the amount of depreciation.

This framework allows us to precisely define an asset based poverty trap. For illustration consider the transition equations depicted in the top panels of Figures 3a and 3b. In each graph, the diagonal 45° line represents the set of points such that $K_{i,t+1} = K_{i,t}$. The transition equation in Figure 3a is globally concave and has a unique steady state, K_i^*. This transition equation can arise in the above example under the assumptions of constant s_i, A_{iv} and δ, and a production function, $f(\cdot)$, that satisfies the Inada conditions. In our context, a transition equation like this, implies that each household eventually converges to a household specific steady state K_i^*, determined by the household's productivity A_{iv} and savings rate s_i. An explanation for poverty, in this view, is that poor households have low productivity, which yields a low steady-state level of productive assets, and hence, low income.

Another example of a transition equation is given in the top panel of Figure 3b. In this case, there are three steady states: two stable steady states, K_{iP}^* and K_{iR}^*, and an unstable steady state, $\widehat{K_i}$, between them. If this is an accurate description of household's capital accumulation dynamics, then poverty can arise because of a low initial endowment.

[11]The production function here should be interpreted as the results of household's optimization across the choice of all available occupations or production technologies. This can be fleshed out by endogenizing occupational choice, as we do in Section 4.

[12]Note that we are here also assuming that there are no credit or rental markets. If there is a frictionless credit market, individuals will immediately borrow the amount needed to produce at the optimal level of capital input. For details see Ghatak (2015).

Households with initial capital below $\widehat{K_i}$ lose capital over time and converge towards the low steady state, K^*_{iP}. The same household (or a household with identical productivity and savings rate) could be at a higher steady state capital level, and hence higher income, had it had access to an initial endowment above $\widehat{K_i}$. This is a case of an asset based poverty trap.

Note that the S shape of the transition equation can be due to different mechanisms. If the true relationship between $K_{i,t+1}$ and $K_{i,t}$ is given by Equation (1) above, such a shape could for example arise due to increasing returns to scale in $f(\cdot)$ or if s_i is an increasing function of $K_{i,t}$.[13] Under the (strong) assumptions that Equation (1) holds and that s_i, A_{iv}, and δ are constant in $K_{i,t}$, there is a direct mapping between the transition equation and the production function, allowing us to draw conclusions about the latter based on estimates of the former.

The S-shaped transition equation is not the only way in which there can be an asset based poverty trap. Figure 3c shows a transition equation with a discontinuity. There are again two stable steady states, K^*_{iP} and K^*_{iR}, but now there is no steady state between them. Instead, households at and above the discontinuity point $\widehat{K_i}$ accumulate capital whereas those just below $\widehat{K_i}$ decumulate. Such a transition equation can describe a situation where households choose between two different production technologies and where switching to the 'high capital' technology requires an investment in a large indivisible asset.

The bottom panels of Figures 3a, 3b, and 3c show the change in capital over one period, $\Delta K_{i,t+1} = K_{i,t+1} - K_{i,t}$, against the initial level of capital implied by each of the transition equations. We will use these to interpret the empirical results, where we measure $\Delta K_{i,t+1}$ as the change in productive assets in the four years following the asset transfer.

Returning to Figure 1a, this framework illustrates different interpretations of the baseline distribution of productive assets. Let's assume that most households are close to a steady-state level of productive assets when we observe them at baseline. If asset dynamics are governed by a concave transition equation with a unique steady state as in Figure 3a, then the bimodal distribution of assets suggests that there are two groups of inherently different households: those whose steady state is close to zero and those who have a high steady state. By contrast, if asset dynamics are better described by Figures 3b or 3c, then a bimodal distribution of assets might naturally arise as some households conglomerate at a low steady state K^*_{iP}, and others at the high steady state, K^*_{iR}. This could happen even if households are identical with respect to their immutable characteristics captured in A_{iv}. In which of

[13]For a review of different micro foundations see Ghatak (2015). We test several such mechanisms below in Section 3.5.

the two steady states any individual household ends up only depends on their initial asset endowment.

3.2 Method

The above discussion illustrates two general conclusions. First, if the transition equation, $\Phi_i(k_{it})$ is globally concave there cannot be multiple stable equilibria of the capital accumulation process, and hence no asset based poverty trap as we have defined it above. The first step of the empirical analysis formally tests the concavity of $\Phi_i(k_{it})$ using the non-parametric shape test developed by Hidalgo and Komarova (2019).[14]

The second insight from the previous section is that, we can speak of a poverty trap if and only if there is a threshold level of capital, which we call $\widehat{K_i}$, such that those below $\widehat{K_i}$ converge to a low stable steady-state level of capital and those above converge to a high stable steady-state level of capital. In the vicinity of $\widehat{K_i}$, this implies formally that for households with $K_{i,t} < \widehat{K_i}$ we expect $K_{i,t+1} < K_{i,t}$, whereas for households with $K_{i,t} > \widehat{K_i}$ we expect $K_{i,t+1} > K_{i,t}$. The next step of the analysis is, therefore, to construct several estimates of the transition equation and identify a candidate threshold level, $\widehat{K}$.

The sample for this exercise consists of the group of ultra poor households in treatment villages for a period of four years after receiving the transfer. Households with initial post-transfer assets above 3 are dropped, since under perfect targeting, these should not have been included as beneficiaries of the program. This leaves us with a total of 3,276 households in the treatment sample.

We use the following notation. Let $k_{i,0} = \ln K_{i,0}$ denote log productive assets (in thousands of BDT) of household i without the transfer at baseline (in 2007), $k_{i,1} = \ln(K_{i,0} + T_i)$ log productive assets including the value of the transfer T_i at baseline (in 2007),[15] and $k_{i,3} = \ln K_{i,3}$ log productive assets at survey wave 3 four years after the transfer (in 2011). The evolution of households' asset stock after the transfer allows us to estimate an empirical

[14]The test makes use of the fact that concavity restrictions can be written as a set of linear inequality constraints when using an approximation by B-splines. Imposing those restrictions yields a constrained sieve estimator taking a B-splines base. The constrained residuals, adjusted for heteroscedasticity, are used to calculate Kolmogorov-Smirnov, Cramer-Von Mises and Anderson-Darling test statistics after applying a Khmaldaze transformation to eliminate the dependence induced by the use of the nonparametric estimator. Critical values for these tests are obtained by bootstrap using the unconstrained residuals. See Hidalgo and Komarova (2019) for further details.

[15]BRAC distributes the same asset bundles in all villages, hence their value depends on local prices. Since most households chose a cow bundle, we value this using the median cow prices within the catchment area of their BRAC branch.

transition equation:[16]

$$k_{i,3} = \phi(k_{i,1}) + \varepsilon_i, \quad (2)$$

where we should think of $\phi(k_{i,1}) = \mathbb{E}[k_{i,3} \mid k_{i,1}]$ as a transition equation in logs averaged across households.

A key challenge in estimating the transition equation is that, if there is indeed a threshold level at which asset dynamics bifurcate, with those above and below moving in different directions, then in the absence of large shocks there would be no observations close to that threshold. The advantage of our setting is that the program moves over 3,000 households to the hollow part of the distribution of assets. Figure 1b illustrates this. It shows kernel density estimates of the asset distribution in treatment and control villages after the the asset transfer. This puts us in a unique position to estimate the shape of the transition equation for a range of asset values that is typically not observed.

The ideal experiment to causally identify the transition equation would allocate asset transfers of different values to randomly selected households. In our case, variation in $k_{i,1}$ is induced by initial differences in $k_{i,0}$. Since the transfer program was targeted at households without significant productive assets, all eligible households own close to zero assets at baseline and initial differences in assets are therefore small. Nevertheless, as figure 1b illustrates, there is some variation which we can exploit. This means that variation in $k_{i,1}$ is potentially endogenous to the household's capital allocation decision. When estimating the transition equation (Equation (2)), we impose the identifying assumption that the variation in $k_{i,0}$ at baseline is orthogonal to unobservable determinants of changes in capital assets after the program. This assumption might fail for two reasons. First $k_{i,0}$ might be systematically correlated with shocks that affect capital accumulation independently of the program. Second, $k_{i,0}$ might be correlated with unobservables that shape the response to the program. For instance, baseline capital might be correlated to latent talent for livestock rearing or to the effect of the training component of the program. In this case, post-transfer asset dynamics might be driven by individuals' transitions to the new, steady state rather than by poverty trap dynamics.

[16]Since there are additional survey rounds in 2009, 2014, and 2018 it would have been possible to use different time horizons for this analysis. However, the 2009 survey round might be more strongly confounded with other components of the TUP program. For example, BRAC discouraged recipients from selling the transferred assets until two years after the transfer. By the time of the 2014 survey, BRAC had started delivering the program in control villages. Since we use the control group in our analysis below, we decided to focus on the 2011 survey round.

We therefore present additional evidence to support the identifying assumption in Section 3.4. Our strategy has two prongs. First, we use the random allocation of the program and estimate a difference-in-differences model using potential beneficiaries in control villages as a counterfactual for actual beneficiaries in treatment villages. Randomization ensures that, in expectation, these two groups are identical in every respect, including unobservable determinants of capital accumulation correlated with $k_{i,0}$. Second, guided by the theory we use differences in savings rates s_i and productivity A_{iv} to estimate different thresholds conditional on $k_{i,0}$. This allows us to test whether the process of accumulation is consistent with poverty traps without using the (potentially) endogenous variation in $k_{i,0}$.

3.3 Results

Figure 4 shows our main estimate of Equation (2), using a kernel-weighted local polynomial regression.[17] Alternative specifications are presented in Appendix Figure 18. Panel (a) of figure 18 reports the fitted values of a third order polynomial[18] and panel (b) reports the B-spline estimator.[19]

All three specifications show that the transition equation is S-shaped. The Hidalgo-Komarova shape test indeed rejects the null of global concavity with $p < 0.01$ and, in line with that, we also reject the null that the cubic term of the polynomial shown in Panel (a) of Figure 18 is zero.

All three estimation methods impose continuity of the transition equation. This implies that any poverty threshold will appear as an unstable steady state, with $\phi(\widehat{k}) = \widehat{k}$ and $\phi'(\widehat{k}) > 1$, such as shown in Figure 3b. Working for now under the assumption of continuity, we find this threshold level of $\widehat{k}$ by numerically approximating the intersection of $\widehat{\phi}(\cdot)$ with the 45° line. For example, for the local polynomial regression (Figure 4) this is done by finding a point in the smoothed graph just above and just below the 45° line and averaging their coordinates. Adjusting the number of smoothing points allows us to approximate this

[17] Local polynomial regression estimates the conditional expectation $\mathbb{E}[k_3 \mid k_1 = k]$ at each smoothing point k of a pre-specified grid as the constant term of a kernel weighted regression of $k_{i,3}$ on polynomial terms $(k_{i,1} - k), (k_{i,1} - k)^2, \ldots, (k_{i,1} - k)^p$. For more details, see Fan and Gijbels (1996)

[18] This specification is similar to those in Antman and McKenzie (2007), Jalan and Ravallion (2004), and Lokshin and Ravallion (2004). However, these authors analy the dynamics of household income instead of productive assets.

[19] A regression spline is a nonparametric smoothing method that uses spline functions as a basis. In general, an M^{th} order spline is a piecewise $M-$ degree polynomial with $M-2$ continuous derivatives at a set of pre-selected points (called the knots). B-splines are a particular type of splines. For more details, see Wasserman (2006).

point with arbitrary precision. Using this method, we find $\widehat{k} = 2.333$ with a bootstrapped standard error of 0.015.[20] At this threshold, assets are worth 9,308.82 BDT (504 USD). For comparison, the median value of a cow in our sample is around 9,000 BDT. Alternatively we can use the parametric estimates to compute the crossing point analytically, this yields a value of $\widehat{k} = 2.339$ (bootstrapped standard error 0.194), which corresponds to 9,379.14 BDT (508 USD).[21]

Note that the estimation of this poverty threshold only used data on ultra-poor households in treatment villages. The exercise is hence independent from the baseline distribution of production assets in the whole (control) population. It is, therefore, remarkable that the estimated threshold falls exactly in the low density range of the distribution of assets (Figure 1a). As would be expected if there is a poverty threshold separating a low and high equilibrium, we find a low density of households in the vicinity of the estimated threshold at baseline. The multiple equilibrium model thus explains the bimodal distribution of assets. By contrast, a bimodal asset distribution does not arise naturally under a concave production technology. While possible in theory, it requires a bimodal distribution of the savings rate or individual productivity, neither of which we observe in the data (see Appendix Figure 19).

Having identified a potential poverty threshold, we can further analyze asset dynamics in a regression model. Denote by Δ_i asset accumulation in the four years after the transfer over and above the value transferred by BRAC, that is $\Delta_i = k_{i,3} - k_{i,1}$. The bottom panels of Figures 3a–3c illustrate the close relation between Δ_i and the transition equation. As is evident from these figures, if $\widehat{k}$ has indeed the characteristics of a poverty threshold, one would expect $\Delta_i > 0$ for individuals whose baseline level of capital is large enough that, in combination with the transfer, it exceeds the threshold ($k_{i,1} > \widehat{k}$), whereas $\Delta_i < 0$ for those whose baseline level of capital is not large enough ($k_{i,1} < \widehat{k}$). The first column of Table 2 reports the estimates of:

$$\Delta_i = \alpha + \beta I(k_{i,1} > \widehat{k}) + \varepsilon_i \tag{3}$$

where $I(k_{i,1} > \widehat{k}) = 1$ if $k_{i,1} > \widehat{k}$. The estimates suggest that beneficiaries who stay below the threshold despite the transfer lose 14% of the assets over four years whilst those who are pushed past the threshold grow their assets by 16%.

[20] Due to the bootstrap sampling variation, there are cases where the poverty threshold is not unique, i.e there is more than one point at which the transition equation crosses the 45° line from below. In these cases we record the lowest of the estimated thresholds. However, across all 1000 bootstrapped samples, we always find at least one unstable crossing point.

[21] We compute this threshold as the second root of the polynomial $76.9 - (96.9 + 1)k + 41k^2 - 5.7k^3$, which is shown in figure 18.

3.4 Identification

3.4.1 Control group as counterfactual

The evidence provided so far relies on the identifying assumption that the small variation in k_0 at baseline is orthogonal to unobservable determinants of changes in capital assets after the program. However, alternative mechanisms might give rise to the same S-shaped pattern so we now provide evidence to rule these out. First, consider the case of shocks to households' capital stock that are correlated with their baseline capital. Concretely, this can take various forms. For example, households with more baseline assets might be better connected and, hence, more likely to receive windfall inheritances or gifts, or may be able to take greater advantage of some other economic opportunity that may arise independently of the asset-transfer program. Similarly, households with less baseline assets may suffer more from weather or health shocks (Burgess et al., 2017).[22]

Given that the allocation of the program is randomised, we can use potential beneficiaries in control villages, who have the same range of k_0 but do not receive the transfer to control for unobservables correlated with baseline capital and its change over four years. To do so, we define a placebo threshold indicator $I(\widetilde{k}_{i,1} > \widehat{k})$ which is equal to 1 if and only if the household would have been above the threshold had they received a transfer of the same value as the treatment households.[23] If households whose baseline capital is such that the transfer would place them above the threshold systematically receive different shocks than those whose baseline capital is such that the transfer would place them below, this difference will be captured by β in Equation (3) above. In contrast, column 2 of Table 2 shows that the estimated β is close to zero, which suggests that households above and below the threshold are not systematically different in absence of the transfer.

Column 3 shows the results of the following difference-in-differences model:

$$\Delta_{iv} = \alpha + \eta P_v + \beta I(k_{i,1} > \widehat{k}) + \gamma I(k_{i,1} > \widehat{k}) \times P_v + \varepsilon_i$$

[22] The results of this section also cover another plausible scenario in which households with a concave production technology receive random productivity shocks prior to our study but haven't converged to their new steady states when we observe them at baseline. Those with a high productivity draw have started to converge to a high steady state and will be measured with a high k_0. Over the study period, they will then continue to accumulate assets. If this could explain our results, we should see the same pattern in the treatment and control group. In particular, $I(\widetilde{k}_{i,1} > \widehat{k})$ should be a strong positive predictor of Δ_i also in the control group.

[23] We impute the hypothetical transfer $\widetilde{T}$ for the control group in the same way as we did for treatment by assigning to each household the median value of a cow within its BRAC branch. Then $\widetilde{k}_{i,1} = \ln(K_{i,0} + \widetilde{T})$.

where $P_v = 1$ if the village is treated. Under the assumption that, had it not been for the program, ultra-poor households in treatment villages would have experienced the same pattern of capital accumulation as their counterparts in control, γ measures how much treatment households gain from being to the right of the poverty threshold. The estimate of γ is similar to that in column 1, reflecting the fact that the pattern of capital accumulation is not significantly different around the (placebo) threshold for control households.[24]

3.4.2 Heterogeneous thresholds

The difference-in-differences estimates relative to individuals in randomly allocated control villages control for common shocks endogenous to baseline capital but, by definition, cannot take care of endogenous responses to a program which is not offered in control villages. For instance, the fact that the program offers training together with assets might increase households' productivity, A_i, and shift the steady state(s). If the effect of the training component is larger for individuals with a higher level of baseline capital, we can build a scenario where there is a level of k_0 that looks like a poverty threshold even if the production technology is globally concave. This is illustrated in Figure 5.

In Figure 5, we model the program as a combination of an asset transfer T, and an upward shift in each household's steady state resulting from the training component in a world of globally concave transition equations. As before, let $\Delta_i = K_i^* - \widetilde{K_i^*} - T$ denote the post-transfer change in assets, where now K_i^* and $\widetilde{K_i^*}$ refer to the steady state for person i before and after the program, respectively. The new steady state is a function of the old steady state, that is $\widetilde{K_i^*} = g(K_i^*)$, embedding the view that the training effect can vary by initial assets endowment. The change in capital is $\Delta_i = g(K_{i,0}^*) - K_{i,0}^* - T$ which can be positive or negative depending on whether the transfer brings individuals above or below the new steady state. In the example of Figure 5 the effect of the training is increasing in K_0. In this particular case, there exists level of capital, K_2^* such that $K_2^* + T = \widetilde{K_2^*}$. Individual $i = 1$ with $K_1^* < K_2^*$ gains less from the training which means that their new steady state is below their initial steady state plus the transfer, that is $K_1^* < \widetilde{K_1^*} < K_1^* + T$, which implies $\Delta_1 < 0$. Conversely, individual $i = 3$ with $K_3^* > K_2^*$ gains more from the training, raising their new

[24] The estimates for the control group in columns 2 and 3 of Table 2 might at first seem confusing: control households below the threshold see large asset gains on average while treatment households lose assets. This is explained by the fact almost all control households own close to zero assets (see Table 1) — the indicator $I(\widetilde{k}_{i,1} > \widehat{k})$ is hypothetical for controls. In the absence of rental markets, assets cannot fall below zero. Hence, the amount of asset loss is bounded for households with close to zero assets and their average Δ_i likely to be positive.

steady state above their post-transfer asset value, that is $K_3^* < K_3^* + T < \widetilde{K_3^*}$, which implies $\Delta_3 > 0$.[25]

To address this concern, we use variation in the parameters that shift the transition equation to estimate different thresholds for different groups. This will allow us to test for poverty traps exploiting the differences in thresholds conditional on baseline capital. Consider the transition equation,

$$K_{i,t+1} = s_i A_{iv} f(K_{i,t}) + (1 - \delta) K_{i,t}.$$

There are two factors that determine the rate at which capital is accumulated. The first is the saving rate s_i: for a given level of capital and earnings, individuals who can save more will have more capital the next period. The second is the productivity parameter A_{iv}, which depends both on individual traits such as entrepreneurship and village level characteristics such as access to markets and the quality of infrastructure. Individuals who are able to save a large fraction of income, or to generate more income for the same level of capital will be able to accumulate more assets at a given point in time, other things equal. Under the assumption of a poverty trap, this then implies that their threshold will be lower, that is, a smaller transfer will be sufficient to push them out of the trap. This means that two households with the same endowment but different savings rate or productivity, might experience different asset dynamics, allowing us to hold k_0 fixed and thus rule out a mechanism as described above (figure 5)

To test whether individuals with higher saving rate face a lower threshold we use the dependency ratio as an instrument for savings. The rationale for this is that a larger share of earnings can be saved when there are fewer household members who consume but do not earn.[26] To test for differences due to earning potential we use a village measure of excess livestock earnings for non-beneficiaries at baseline. To do so, we regress livestock earnings on the number of cows, both linear and squared, and take the mean residuals at the village level. Intuitively, villages where individuals earn more than predicted from their livestock holdings must have the right infrastructure for livestock businesses.

Figure 7, panel (a) reports non-parametric estimates of the transition equation for households above and below the median saving rate, instrumented by the dependency ratio, while

[25]Note that while this mechanism would give rise to a level of capital that looks like a poverty threshold, it cannot explain the exact shape of the empirical transition equation as shown in Figure 4. One would have to assume a more elaborate relationship between the treatment effect of the training and K_0 in order to construct such an empirical transition equation under the concave production technology.

[26]The fact the median age of respondents is 40 at baseline implies that we can assume that fertility is exogenous to asset accumulation.

panel (b) splits households in those above and below the median earnings potential. Both panels show that the transition equation for households above the median is vertically above that for households below the median. The threshold for households with larger savings (earnings) potential is 2.29 (2.24), while that for households below the median is 2.36 (2.39). For inference we randomly split the sample into two equal sized sub-samples, either using the individual or the village as unit, estimate the thresholds in each and take the difference between the two. The probability of a random sample split yielding the observed differences is less than 0.01 for both.

The fact that differences in savings and earnings potential imply different thresholds provides us with an alternative identification strategy that uses the differences in thresholds for the same level of baseline capital. Table 3 estimates three regressions for each of the two dimensions. Columns 1 and 4 estimates the change in capital stock above and below the individual threshold, i.e. the high threshold if the household is below the median and vice versa. In line with the earlier findings we see that individuals for whom the transfer is not large enough to bring them past the earnings-specific threshold lose 16% of asset value in four years, whilst those who pass the threshold accumulate 14%. Similarly, individuals for whom the transfer is not large enough to bring them past the savings-specific threshold lose 15% of asset value in four years, whilst those who pass the threshold accumulate 17%. Columns 2 and 5 control for the level of baseline capital. Strikingly the coefficients remain stable, which is consistent with the fact that neither savings nor potential earnings are correlated with baseline capital. More importantly, and in line with the analysis in the previous section, these results reassure us that different patterns of accumulation above and below the threshold are not due to unobservables correlated with baseline capital.

Finally, columns 3 and 6 test whether it is the relevant individual thresholds that bind. To implement this test we restrict the sample to individuals with *high* thresholds and estimate:

$$\Delta_i = \alpha + \beta_L I(k_{i,1} > k^u_L) + \beta_H I(k_{i,1} > k^u_H) + \varepsilon_i,$$

here β_L measures the effect of being past the low threshold while β_H measures the effect of being past the high threshold. The results show that these individuals lose capital regardless of whether they are above the low threshold but they start accumulating once they go above the high threshold, which further allays the concern that results are driven by unobservables related to baseline capital.

3.5 Mechanisms

To provide evidence on the mechanisms underlying the poverty trap we begin by testing two broad classes of explanations.

In the first, the non-convexity is generated by a standard and continuous form of (local) increasing returns to scale, e.g., due to learning by doing. In the second, it is generated by a discontinuity which results in local increasing returns as well, for instance, due to lumpiness or indivisible investment opportunities. For instance, larger livestock herd size increases returns for pastoralists in southern Ethiopia, as livestock in larger herds is more likely to survive weather shocks and profitable migrations. (Lybbert et al. (2004), Santos and Barrett (2016)). Figure 3b shows the case of a continuous transition equation. For example, this results from increasing returns to scale at a low level of production. By contrast, Figure 3c shows the case of a discontinuous transition equation. Assuming constant s_i, A_i and δ, such a transition equation would suggest a production technology with indivisible investments. The transition equation maps directly to $\Delta(k)$ which is shown in the bottom panel of both figures. We can, therefore provide some evidence on the mechanism by testing whether $\Delta(k_0)$ is continuous.[27]

We estimate a linear model of $\Delta(k)$ above and below the threshold, allowing for a discontinuity at the threshold:

$$\Delta_i = \alpha + \beta_0 I(k_{i,1} > \widehat{k}) + \beta_1 k_{i,1} + \beta_2 I(k_{i,1} > \widehat{k}) \times k_{i,1} + \varepsilon_i.$$

Our null hypothesis is $H_0 : \beta_0 = 0$, i.e., there is no discontinuity. The results of this regression are shown in columns 4 to 6 of Table 2. Column 4 rejects the null and shows a discontinuity at $\widehat{k}$, where the change in capital goes from -0.29 to 0.19. To account for shocks that would have occurred in the absence of the program, Column 5 estimates the above regression in the control group. As before, we set $I(\widetilde{k}_{i,1} > \widehat{k}) = 1$ if the baseline level of assets is such that control households would be past the threshold had they received the transfer. Here, there is no discontinuity at the (placebo) threshold as β_0 is close to zero and precisely estimated. Column 6 pools treatment and control villages together and shows that the difference in the

[27] Note that the observed Δ depends on the time horizon over which households are observed after being placed close to the threshold and the speed of convergence. If the four years are long enough that household fully converge to the steady state, there might be an apparent discontinuity as those who started closest to the threshold changed their assets the most. We are here assuming that households have not yet fully converged, which is consistent with asset density estimates two and four years after the transfer. $\Delta(k_0)$ is therefore indicative of the speed of convergence.

discontinuity at $\widehat{k}$ is precisely estimated.

Column 4 of Table 2 also shows that the pattern of change is consistent with the production function in Figure 3c throughout the range of baseline capital: Δ is decreasing in k until $\widehat{k}$, where it "jumps" above zero and then remains constant. This is further illustrated in Figure 6 where we relax the assumption of linearity and estimate two local polynomial regressions of Δ_i on $k_{i,1}$ allowing for a discontinuity at $k_1 = \widehat{k}$. The left panel showing the estimates of the treatment group confirm the linear regression results. The same estimate for control villages shows only mean reversion.[28]

We also note that the shape of Δ is inconsistent with heterogeneous responses to training endogenous to baseline capital as in Figure 5 because that implies $\beta_1 > 0$ throughout. Taken together, the evidence in Table 2 and Figure 3c indicates that while individual heterogeneity endogenous to baseline capital can explain why individuals with low k_0 decumulate assets after the transfer on average, only a convexity in the transition equation can explain the response to the transfer that we observe in the data.

Next, we investigate the mechanisms that underpin the discontinuity, focusing on the three broad classes suggested in the literature: nutrition, credit and savings constraints, and indivisible assets. We estimate the following panel data model:

$$y_{i,s} = \sum_{j=1}^{3} \left[\gamma^j I(k_{i,1} > \widehat{k}) \times S_j + \delta^j S_j \right] + \varepsilon_{i,s}$$

where now $s \in \{1, 2, 3\}$ denotes the survey waves at baseline, two, and four years after the transfer, respectively. $y_{i,s}$ is an outcome of interest, and $S_j = I(s = j)$ indicates the survey wave. Our sample still consists of ultra-poor households in treatment villages, but since we are now interested in discontinuous changes around the threshold, we further restrict the sample to those in a narrow window of baseline assets $k_{i,1} + T \in [2.24; 2.44]$.

This allows us to test whether individuals just above and below the threshold are similar at baseline ($H0 : \gamma^1 = 0$) and whether there is a discontinuity in any of these variables after the program is implemented. For instance, if the trap were driven by low caloric intake,

[28] As with Table 2 one has to be careful with the interpretation of the control group effects. Again, direct comparison between treatment and control is not possible, since control households do not actually receive the transfer. We plot $k_0 + \tilde{T}$ on the horizontal axis (where $\tilde{T}$ is an imputed, placebo transfer) and allow for a break at the same point as in the treatment group at $\widehat{k} = 2.333$. We plot this against the actual Δ_i, which is positive for households with low asset values which have nothing to lose. We interpret the negative delta at high asset levels as reversals from positive shocks prior to the study (these shocks are for the most part not large enough move a household above the true poverty threshold.)

we would expect this to increase discontinuously once the transfer puts the constrained individuals above the threshold.

Table 4 shows the resulting estimates of γ^j. Columns 1-3 in Panel A test the nutrition hypothesis, using food expenditure, calories per capita and BMI. We find that individuals just to the left and those just to the right of the threshold have almost identical nutrition at baseline, and this does not change over the years. The evidence thus indicates that nutrition is unlikely to drive the feedback mechanism that underpins poverty traps. Columns 4-5 analyse differences in financial flows. Again individuals just to the left and those just to the right have similar levels of savings and loans at baseline and no significant differences emerge over time. In particular, there doesn't seem to be a differential adjustment of the savings rate in a narrow window around the threshold.

Panel B decomposes the effect on different productive assets. Two findings are of note. The first is that within this narrow window, there is a discontinuity in the ownership of vehicles. That is, among individuals with very similar levels of capital, those just to the right of the threshold own about 4% more vehicles. Second, after the transfer the difference between individuals above and below the threshold grows rapidly overtime with the acquisition of increasingly more expensive assets: cows, sheds and goats after two years and cows after four. The latter is particularly striking: by year four, individuals above the threshold have 58% more cow stocks than those below.

In summary, the evidence in Table 4 indicates that the program does not relax a food constraint or allow better access to credit. Rather we find that indivisibilities underpin the poverty trap: indeed individuals with baseline capital high enough that the transfer will place them past the threshold own more expensive assets (mostly vehicles) and accumulate even more expensive assets (cows) after being treated.

As assets are combined with labor to generate income, the picture that emerges is one where poor people cannot afford to purchase indivisible productive assets and remain employed in low wage, insecure casual jobs that pay little relative to the price of the asset and keep them in a poverty trap. This raises several key questions for policy: do poverty traps create misallocation? That is, would the poor be more productively utilized in the occupations of the rich and if so, by how much do we lose in terms of aggregate output because of this? In the next section we develop a structural model of occupational choice to find answers and estimate policy counterfactuals.

4 Structural estimation

The results of the previous section suggest that the overwhelming concentration of the ultra-poor in wage labor at baseline is unlikely to reflect those individuals' first-best choice of occupation given their productivity and preference parameters. In this section, we use a simple model of occupational choice to estimate individual-level parameters, determine the optimal occupation for each individual in the absence of capital constraints and hence quantify the extent of misallocation at baseline. Using these results, we simulate the implied total value and distribution of transfers necessary for all households to escape the poverty trap, and consider the effects of a series of policy counterfactuals.

4.1 Simple model of occupational choice

Consider a simple environment where individuals allocate their time endowment R between self-employment in livestock rearing (l) and wage labor (h). We allow individuals to have a mixed occupational choice and allow for overall labor supply to be elastic. We also consider the possibility of hiring in external labor (h') for livestock rearing, so that the total labor input in that activity is $l + h'$. The wage rate for hired-in labor is w'.

Let the individual production function for livestock rearing be given by (we drop subscript i for simplicity):

$$q = AF(\bar{k}, l + h').$$

We assume that the capital stock $\bar{k}$ is given and there is no possibility of borrowing or depositing money in a bank and earning interest.

Since $\bar{k}$ is a constant, this is effectively a one-input production function that depends on $l + h'$. We will restrict attention to production functions that are multiplicatively separable in capital and labor:

$$F(\bar{k}, l + h') = f(\bar{k})g(l + h')$$

Notice that therefore, even if the production function may be S-shaped with respect to k when k is not given, so long as it is concave with respect to $l + h'$ we can use standard maximization techniques. Since we are mostly concerned with properties of $f(k)$ relating to convexity or non-convexity, we will assume that $g(l + h')$ is strictly concave.

For a wage laborer, the wage rate is w. We assume $w > w'$, to capture the fact that hired-in workers are usually members of the farmer's own family, and generally they are paid less than what the farmer earns by supplying wage labor herself. There is an exogenous demand constraint in the labor market, and so $h \leq \overline{H}$ where $0 < \overline{H} < \overline{R}$. Similary, there is a constraint on the maximum hours of labor a farmer can hire in, $h' \leq \overline{N}$.

We assume that the (disutility) cost of supplying labor takes the form

$$\frac{1}{2}(\sqrt{\psi_l}l + \sqrt{\psi_h}h)^2$$

where $\psi_h > 0$ and $\psi_l > 0$.

As a result, the static optimization problem becomes:

$$\max_{l\geq 0, h\geq 0, h'\geq 0} Af(\bar{k})g(l+h') + wh - w'h' - \frac{1}{2}(\sqrt{\psi_l}l + \sqrt{\psi_h}h)^2 \tag{4}$$

subject to

$$h \leq \overline{H} \tag{H}$$

$$h' \leq \overline{N} \tag{N}$$

$$h + l \leq \overline{R} \tag{R}$$

Assuming a fully interior solution, the first-order conditions for the maximization are:

$$\begin{aligned} Af(\bar{k})g'(l+h') &= \psi_l l + \sqrt{\psi_l \psi_h} h \\ w &= \sqrt{\psi_l \psi_h} l + \psi_h h \\ Af(\bar{k})g'(l+h') &= w' \end{aligned}$$

In the case of corner solutions, some of the equalities above need not hold. The full solution with all possible cases is characterized in Appendix A.

4.2 Model calibration

The first step in the estimation is to calibrate individual-level parameters for productivity in livestock rearing A and disutility of supplying wage labor and livestock rearing hours, ψ_h and

ψ_l respectively. These parameters are identified from baseline and year 2 data by assuming that, in these years, individuals choose the hours that they devote to each occupation[29] and hire in optimally given their capital endowment, production technology, prevailing wage rates and exogenous hours constraints. The assumptions used for each of these is as follows.

The production function assumed is

$$f(k_i)g(l_i + h_i') = (ak_i^2 + bk_i)(l_i + h_i')^\beta.$$

It represents the latent quadratic production function which, when combined with flat wage income that dominates at low capital levels, yields the characteristic S shape described in Section 3.1. The parameters a, b and β of this function are estimated by non-linear least squares. The prevailing market wage and wage paid for hired in labor are means at the branch level in each survey wave. We set the time endowment constraint $\overline{R}$ to be 3,650 hours per year, and drop from the estimation the three ultra-poor individuals who report total hours higher than this at baseline or year 2. The labor demand constraint $\overline{H}$ is set at the 90$^{\text{th}}$ percentile of wage labor hours worked at baseline by BRAC branch. The constraint $\overline{N}$ on how much labor can be hired in is set at the 95$^{\text{th}}$ percentile across all households and survey rounds, equal to 1,400 hours per year.

The optimization problem described in Section 4.1 yields first order conditions for several cases according to the occupation(s) in which the individual works, whether they hire in labor and whether each of the exogenous hours constraints binds. For the majority of ultra-poor individuals, these first order conditions can be combined with data on capital and occupational choice at baseline and year 2 to calibrate the values of the parameters A, ψ_h and ψ_l that are consistent with the observed hours worked in livestock rearing and wage labor, and hours of labor hired in, being chosen optimally.

In particular, 16% of ultra-poor individuals mix occupations and hire in labor at year 2 (case 1 in Appendix A), such that the three year 2 first order conditions can be solved for the three parameters of interest for these individuals. For those individuals in other cases at year 2, there are fewer first-order conditions than parameters so this method cannot be used. However, in many of these cases, first order conditions from year 2 and baseline can be combined to calibrate the parameters. In our data 24% of individuals specialize in wage labor without hiring in labor at baseline, and at year 2 either mix occupations without hiring in labor or specialize in livestock rearing with hired in labor. In these cases, the baseline

[29] These are self reported and checked for consistency.

and year 2 first order conditions again yield three equations than can be solved for the three parameters. Parameters can be calibrated for a further 23% of individuals by assigning ψ_h to be the maximum observed value for those individuals who do not work at baseline.[30]

This method yields estimated individual-level parameters for 64% of ultra-poor individuals. In all other combinations of cases at baseline and year 2, there are either very few individuals or the combination of cases does not permit calibration of all parameters (for instance, if an individual specializes in livestock rearing at baseline and year 2, it is not possible to pin down their disutility of wage labor hours). Plotting the baseline productive assets distribution for the 64% of households for whom we can conduct estimation and the 36% for whom we cannot reveals a high degree of overlap, with the latter distribution slightly rightward shifted. This suggests that those for whom we cannot conduct estimation are more likely to engage in livestock rearing and, therefore, less likely to be constrained in their choice of occupation choice (though not necessarily hours worked in each occupation).

Figure 8 plots the calibrated values of A, ψ_h and ψ_l against post-transfer baseline capital and shows that there is no systematic correlation between baseline wealth and any of these parameters, and no evidence of a discontinuity at the threshold capital level. The fact that A is not correlated with k_0 provides further support for our identification assumptions in the reduced form estimation. Moreover we find that, in line with the fact that wage labor carries social stigma, the disutility of wage labor hours ψ_h is higher than the disutility of livestock rearing hours ψ_l, as shown in Figure 9. The median value of ψ_h is more than 50% higher than the median ψ_l value. The distribution of the calibrated A parameters, shown in Figure 10, is unimodal.

4.3 Model estimation

With estimated values for each individual's productivity in livestock rearing and disutility of labor hours in hand, we can use the model structure to solve for each individual's optimal hours in wage labor and livestock rearing, their optimal hours of hired in labor, and their implied payoff at any level of capital. In a first simulation exercise of this nature, we calculate these at each individual's year 4 capital level in order to assess how well the model matches non-targeted moments in the data. In a second, we estimate the value of misallocation at baseline by comparing each individual's optimal occupational choice and payoff at the steady

[30] We abstract from the labor demand constraint and constraint on hired in labor in the parameter calibration since the choice of hours across occupations will be uninformative about underlying parameters where these constraints bind.

state capital level of the rich (i.e. in the absence of capital constraints) to those observed at their baseline capital level.

4.3.1 Testing model fit using year 4 data

We test the predictive power of the model by using the model to simulate each individual's optimal choice of hours in each occupation at their year 4 capital level, and comparing these to the choice of hours observed for that individual at year 4.

Figure 11 demonstrates that the share of individuals working in livestock rearing, wage labor or mixing in the data matches closely the model-predicted shares in each occupation. Figures 12 and 13 show local polynomial predictions of model-predicted and actual hours in livestock rearing and wage labor respectively, as a function of year 4 capital. In both cases, there is a close fit between the model-predicted and observed hours.

4.3.2 Quantifying misallocation

In order to quantify misallocation, we estimate the payoff that the model suggests would be available to each ultra-poor individual were they to have the steady state capital level of the rich, and compare this to the payoff available to them at their baseline capital level.

The steady state capital level of the middle and upper classes is estimated to be the level corresponding to the upper mode of the distribution across all wealth classes of productive assets excluding land, which occurs at 43,701 BDT.[31] This is higher than the baseline capital level of the vast majority of ultra-poor individuals, so in extrapolating to this higher capital level it is necessary to account for the income effect in the demand for leisure suggested by the observed negative correlation between income and hours worked at baseline. We achieve this by scaling up the labor disutility parameters ψ_h and ψ_l by the ratio of the median ψ_l for richer classes versus the median ψ_l for the ultra-poor.[32]

The model yields an expression for the optimal hours worked in each occupation and hired in, and respective payoffs, in each of the cases outlined in Appendix A. We use these expressions, together with the calibrated values of each individual's livestock-rearing productivity and disutility of labor hours, to calculate the occupational choice, hours worked and hours

[31] Land is excluded in choosing this level since women across wealth classes rarely cultivate land; the ultra-poor possess little land across survey rounds; and land is a very expensive asset, the purchase of which is not endogenized in our model. The distribution of productive assets excluding land is also bimodal as shown in Figure 20.

[32] For five households, this scaling up of the disutility of labor is sufficient to result in negative estimated misallocation. For these households, the estimated value of misallocation is set to zero.

hired in that would yield the highest payoff for each individual at the steady state capital level of the rich.

The results of this exercise reveal that, at the steady state capital level of the rich, 90% of ultra-poor households for whom we can conduct the structural estimation should optimally specialise in livestock-rearing, 8% should mix and just 2% should specialise in wage labor. This contrasts starkly to the observed distribution across occupations at baseline, as shown in Figure 14. At their baseline capital level, only 1% of working ultra-poor households specialize in livestock rearing, with 97% specializing in wage labor and 2% mixing occupations. As such, the model suggests that 96% of individuals for whom we can conduct the structural estimation have non-zero misallocation.

The model also yields the total value of misallocation across all households for which the estimation is conducted as the sum of the differences between the payoff available to each individual at the steady state of the rich and at their baseline capital level. The estimation suggests that the total value of misallocation thus quantified is 15 USD million.[33] The estimated total value of transfers required to bring all of these individuals to the average threshold capital level identified in Section 3.3 — from which they are able to escape the poverty trap[34] — is an order of magnitude smaller at 1 USD million.

4.4 Simulating policy counterfactuals

The structure of the model allows us to simulate the effect of counterfactual changes in the model's parameters. We use this to consider how the results are influenced by potential general equilibrium effects of the intervention and to study the effects of a series of counterfactual policy interventions.

The central simulation exercise above aims to quantify the effects of propelling large numbers of the ultra-poor to higher capital levels. A potential concern with this is that the induced large-scale increase in livestock rearing might have general equilibrium effects that influence the returns to livestock rearing, for instance due to falling produce prices. To investigate this possibility, we re-simulate the results reducing livestock income $Af(k)g(l+h')$ by a fixed factor. We find that, even in a case where this is reduced by 50%, 71% of ultra-poor

[33]This is the implied gain each period once the steady state has been reached. Here and in all simulations we top-code the top 5% of individual misallocation values at the 95^{th} percentile to reduce the effect of outliers.

[34]Beyond this point, the transition equation is concave and the individual can accumulate towards the high stable steady state. As such, the transfers required to set individuals on a stable trajectory out of poverty need only elevate them to the capital level of the unstable steady state, from where they can continue to save towards convergence.

households should specialize in livestock rearing, though the estimated value of misallocation falls by 57%. In order for the value of misallocation to fall to the estimated cost of eliminating the poverty trap, the simulations suggest that livestock income would need to be reduced by 89%. These results suggest that general equilibrium price effects may attenuate the estimated value of misallocation, but are unlikely to overturn the central finding that the value of implied misallocation far exceeds the cost of eliminating the poverty trap.

In a second set of counterfactual simulations, we consider the effects of a series of alternative policy interventions that might be considered to tackle occupational inequality in this setting. In the first of these, we simulate the effect of increasing the wage available for wage labor activities. Even with a doubling of the wage rate, the simulations suggest that the share of households optimally specializing in livestock rearing at the steady state capital level of the rich is 60%. An alternative policy counterfactual considers the effect of reducing the disutility of wage labor hours, ψ_h, for instance through increasing availability of occupations that do not bear the social status costs of agricultural or domestic service occupations. The simulations suggest that reducing all individuals' disutility of wage labor hours by 50% would reduce the share of the ultra-poor that should optimally specialize in livestock rearing to 79%. In the simulations that increase the wage rate or reduce the disutility of wage labor hours, the estimated value of misallocation falls much less than in the simulation that reduces livestock income (less than 10% in both cases). This is because the former simulations influence marginal individuals in the left tail of the misallocation distribution, while the latter shifts the entire misallocation distribution to the left.

While the share optimally specializing in livestock rearing in both policy counterfactual simulations is lower than the share in the central simulations, these are still an order of magnitude higher than the 1% observed among ultra-poor households at baseline.

5 Implications for policy

Our results point to the existence of a poverty threshold, so that households with a starting level of productive assets below that threshold are trapped in poverty, and households who are able to get past that threshold accumulate capital and approach the level of the richer classes. That allows them to switch occupations from casual laborers to the more productive business activity of livestock rearing. The existence of such a poverty threshold has important implications for policy design. Transfer programs that bring a large share of households above the threshold will see large effects on average, while transfers that fall short of this might have small effects in the long run.

As a simple illustration, we can compute the share of households in our sample that would have been moved above the threshold as a function of the transfer size. The black line in Figure 15 shows this. To construct this graph, we compute the difference between the threshold value and the initial value of productive assets for ultra-poor households. When computing this gap, it is necessary to account for the fact that some households would move above the threshold through positive shocks even without a transfer. We account for that by drawing random shocks from ultra-poor, poor, and middle class households in the control group and adding those to the initial assets. To allow comparability with alternative policies, we express the transfer value relative to annual per capita consumption. As the figure shows, almost 20% of households would reach the threshold even without a transfer. Consistent with the fact that most ultra-poor households own close to zero assets, small transfers only slightly increase the share of households that pass the threshold. At a transfer just above 80% of annual per capita consumption all households, even those with zero baseline assets, get moved past the poverty threshold.

The vertical lines in Figure 15 show the size of the actual transfer (blue) and alternative transfer policies (red). The country names refer to study sites in Banerjee et al. (2015b), who conducted randomized evaluations of graduation program similar to BRAC's program in six countries. Blattman, Fiala, and Martínez (2013) study the impact of unconditional cash grants to young adults in northern Uganda. The transfers in the different sites of the Banerjee et al. (2015b) study and in the study by Blattman, Fiala, and Martínez (2013) were similar in relative size to our setting, and in both cases the authors found similarly large, positive effects on average. For example, the cash grants in Blattman, Fiala, and Martínez (2013), roughly the size of recipients' annual income, caused large increases in business assets and earnings.[35]

We also consider the effect of alternative transfer schemes such as income support (NREGA) and microfinance. BRAC typically offers entry microloans between 100 USD and 200 USD. Our results suggest that such small transfers would only allow a small share of households to escape poverty – those that are already close to the threshold. This is consistent with evidence suggesting a negligible average impact of microfinance (Banerjee et al., 2015a; Meager, 2019) but a large effect on a small group of households that already run a successful

[35]However, the variation of effects across study sites in Banerjee et al. (2015b) is not fully in line with this figure. For example, they find much larger gains on asset ownership in India than in Honduras, highlighting the fact that contextual factors beyond the transfer size can play an important role.

business (Banerjee et al., 2019). In a similar fashion to these reduced form results, we can use the structural model to simulate the effect of different transfer policies on occupational misallocation.

The structural model can be used to estimate the value of transfers needed to reduce misallocation to zero. To achieve this, we resimulate the model under the assumption that all households are given a transfer equal to an increasing percentage of annual per capita consumption expenditure, until the point at which misallocation equals zero. This exercise suggests that the value of misallocation — measured as before against the maximum payoff available at the upper mode of the distribution of productive assets excluding land — would be zero if all ultra-poor households were given a transfer equal to 3.95 times the average level of baseline per capita consumption expenditure among ultra-poor households at baseline. The trajectory of the total value of misallocation as the transfer value is increased is shown in Figure 16. The total cost of transferring 3.95 times times the average level of baseline per capita consumption expenditure to each of the 2,283 ultra-poor households in the estimation, 5.7 USD million, remains much lower than the total value of estimated misallocation (15 USD million).

In a second set of simulations, we consider the possibility that misallocation could be measured not against the upper mode of the distribution of productive assets excluding land, but instead versus the maximum payoff available at the unstable steady state — from where the theory suggests individuals can accumulate towards the high steady state along the concave part of the transition equation. The results in this case are shown in Figure 17 and suggest that the value of misallocation would be zero if all ultra-poor households received a transfer equal to 1.05 times the average level of baseline per capital consumption expenditure.

6 Conclusions

We have shown that poverty traps exist. People are poor because of a lack of opportunity. It is not their intrinsic characteristics that trap people in poverty but rather their circumstances. Poverty is not an innate condition. This has implications for how we think of development policy and for the value of eliminating global poverty.

The first implication is that the solution to the global poverty problem will require a big push that enables the poor to take on more productive occupations. Small pushes will work to elevate consumption but will not get people out of the poverty trap. Large swathes of the poor both in Bangladesh and across the developing world are characterised by working as itinerant, casual laborers (Bandiera et al., 2017; Kaur, 2019). The type of asset transfer program studied here may be directly relevant in those contexts. However, the larger point is that a range of interventions may be effective in getting people out of poverty traps as long as they are effective in shifting people into occupations that leverage their talent. Indeed the transfer may take the form of human as opposed to physical capital. For example, Alfonsi et al. (2020) find that significant investments in vocational training and apprenticeships have large impacts on employment and earnings of disadvantaged youth in Uganda. Furthermore, as is stressed in the macroeconomics literature, investments in infrastructure or other policies which raise individual productivity can also achieve the same effect. Big push policies will typically imply that the magnitude of the transfer needed is much larger than is typical with current interventions though importantly it can be time limited.

The second implication is that poverty traps create mismatches between talent and jobs. We have shown that misallocation of labor is rife amongst the poor in rural Bangladesh. They are perfectly capable of taking on the occupations of the richer women but are constrained from doing so by a lack of resources. The value of eliminating misallocation is an order of magnitude larger than the cost of moving all the beneficiaries past the threshold. This is important as it implies that poverty traps are preventing people from making full use of their abilities. Redistributing capital is one possible way to address the mismatch. The alternative is to remove the obstacles that prevent the owners of capital from hiring individuals to work with their assets. This requires progress on our understanding of the core reasons for contractual imperfections, and evidence on policies that can eliminate them.

Ending poverty is the central focus of development economics and policy. This paper points to the importance of expanding opportunity for the poor. In effect, it sharply highlights the need to rethink our approach to tackle the problem of global poverty, and in particular, the critical importance of focusing on welfare policies that change the production activities of the poor. This is distinct from traditional consumption focused policies which have characterised welfare support both in developed and developing countries. It is only by expanding opportunities for the poor that we will be able to tap into the productive capacity of a large cross-section of humanity.

References

Alfonsi, Livia, Oriana Bandiera, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali (2020). "Tackling youth unemployment: Evidence from a labor market experiment in Uganda". *Working Paper*.

Antman, Francisca and David McKenzie (2007). "Poverty Traps and Nonlinear Income Dynamics with Measurement Error and Individual Heterogeneity". *Journal of Development Studies* 43.6, pp. 1057–1083.

Atamanov, Aziz, R Andres Castaneda Aguilar, Carolina Diaz-Bonilla, Dean Jolliffe, Christoph Lakner, Daniel Gerszon Mahler, Jose Montes, Laura Liliana Moreno Herrera, David Newhouse, Minh C Nguyen, et al. (2019). "September 2019 PovcalNet Update: What's New". Global Poverty Monitoring Technical Note no. 10.

Azariadis, Costas (1996). "The Economics of Poverty Traps Part One: Complete Markets". *Journal of Economic Growth* 1.4, pp. 449–486.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman (2017). "Labor Markets and Poverty in Village Economies". *Quarterly Journal of Economics* 132.2, pp. 811–870.

Banerjee, Abhijit, Emily Breza, Esther Duflo, and Cynthia Kinnan (2019). "Can Microfinance Unlock a Poverty Trap for Some Entrepreneurs?" *NBER Working Paper* No. w26346.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan (2015a). "The miracle of microfinance? Evidence from a randomized evaluation". *American Economic Journal: Applied Economics* 7.1, pp. 22–53.

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry (2015b). "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries". *Science* 348.6236.

Banerjee, Abhijit and Andrew Newman (1993). "Occupational Choice and the Process of Development". *Journal of Political Economy* 101.2, pp. 274–298.

Blattman, Christopher, Nathan Fiala, and Sebastián Martínez (2013). "Generating Skilled Self-employment in Developing Countries: Experimental Evidence from Uganda". *Quarterly Journal of Economics* 129.2, pp. 697–752.

Burgess, Robin, Olivier Deschenes, Dave Donaldson, and Michael Greenstone (2017). "Weather, Climate Change and Death in India". *Working Paper*.

Carter, Michael R and Christopher B Barrett (2006). "The Economics of Poverty Traps and Persistent Poverty: An Asset-based Approach". *Journal of Development Studies* 42.2, pp. 178–199.

Dasgupta, Partha (1997). "Nutritional Status, the Capacity for Work, and Poverty Traps". *Journal of Econometrics* 77.1, pp. 5–37.

Dasgupta, Partha and Debraj Ray (1986). "Inequality as a Determinant of Malnutrition and Unemployment: Theory". *The Economic Journal* 96.384, pp. 1011–1034.

Easterly, William (2006). "Reliving the 1950s: the Big Push, Poverty Traps, and Takeoffs in Economic Development". *Journal of Economic Growth* 11.4, pp. 289–318.

Fan, Jianqing and Irène Gijbels (1996). *Local Polynomial Modelling and Its Applications*. Springer US.

Galor, Oded and Joseph Zeira (1993). "Income Distribution and Macroeconomics". *The Review of Economic Studies* 60.1, pp. 35–52. ISSN: 0034-6527. DOI: `10.2307/2297811`. eprint: `https://academic.oup.com/restud/article-pdf/60/1/35/4480796/60-1-35.pdf`. URL: `https://doi.org/10.2307/2297811`.

Ghatak, Maitreesh (2015). "Theories of Poverty Traps and Anti-poverty Policies". *World Bank Economic Review* 29, S77–S105.

Hidalgo, Javier and Tatiana Komarova (2019). "Testing Nonparametric Shape Restrictions". *Manuscript* (LSE).

Hirschman, Albert O. (1958). *The Strategy of Economic Development*. New Haven: Yale University Press.

Imbert, Clement and John Papp (2015). "Labor market effects of social programs: Evidence from india's employment guarantee". *American Economic Journal: Applied Economics* 7.2, pp. 233–63.

Jalan, Jyotsna and Martin Ravallion (2004). "Household Income Dynamics in Rural China". *Insurance Against Poverty*, pp. 108–124.

Kaur, Supreet (2019). "Nominal Wage Rigidity in Village Labor Markets". *American Economic Review* 109.10, pp. 3585–3616.

Kraay, Aart and David McKenzie (2014). "Do Poverty Traps Exist? Assessing the Evidence". *Journal of Economic Perspectives* 28.3, pp. 127–48.

Lewis, Arthur (1954). "Economic development with unlimited supplies of labour". *Manchester School* 22.2, pp. 139–191.

Lokshin, Michael and Martin Ravallion (2004). "Household Income Dynamics in Two Transition Economies". *Studies in Nonlinear Dynamics & Econometrics* 8.3.

Lybbert, Travis J., Christopher Barrett, Solomon Desta, and D. Layne Coppock (2004). "Stochastic Wealth Dynamics and Risk Management among a Poor Population". *The Economic Journal* 114.498, pp. 750–777.

Meager, Rachael (2019). "Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments". *American Economic Journal: Applied Economics* 11.1, pp. 57–91.

Murphy, Kevin, Andrei Shleifer, and Robert Vishny (1989). "Industrialization and the Big Push". *Journal of Political Economy* 97.5, pp. 1003–1026.

Myrdal, Gunnar (1957). *Economic Theory and Under-Developed Regions*. Gerald Duckworth & Co., Ltd.

— (1968). *Asian Drama: an Inquiry into the Poverty of Nations*. Vol. Volumes I, II and III. Pantheon.

Nurkse, Ragnar (1961). *Problems of Capital Formation in Underdeveloped Countries*. Oxford University Press.

Ray, Debraj and Peter Streufert (1993). "Dynamic equilibria with unemployment due to undernourishment." *Econ Theory*, pp. 61–85.

Rosenstein-Rodan, Paul (1943). "Problems of Industrialisation of Eastern and South-eastern Europe". *The Economic Journal* 53.210/211, pp. 202–211.

Rostow, Walt (1960). *The Stages of Economic Growth: A Non-communist Manifesto*. Cambridge University Press.

Santos, Paulo and Christopher B Barrett (2016). *Heterogeneous Wealth Dynamics: On the Roles of Risk and Ability*. Tech. rep. NBER Working Paper.

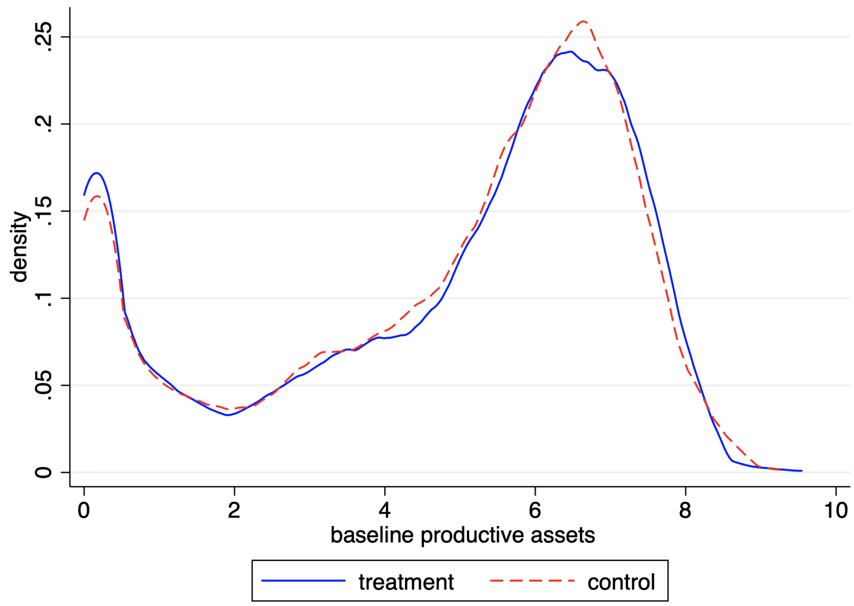
Schultz, Theodore (1980). "Nobel Lecture: The Economics of Being Poor". *Journal of Political Economy* 88.4, pp. 639–651.

Solow, Robert (1956). "A Contribution to the Theory of Economic Growth". *Quarterly Journal of Economics* 70.1, pp. 65–94.

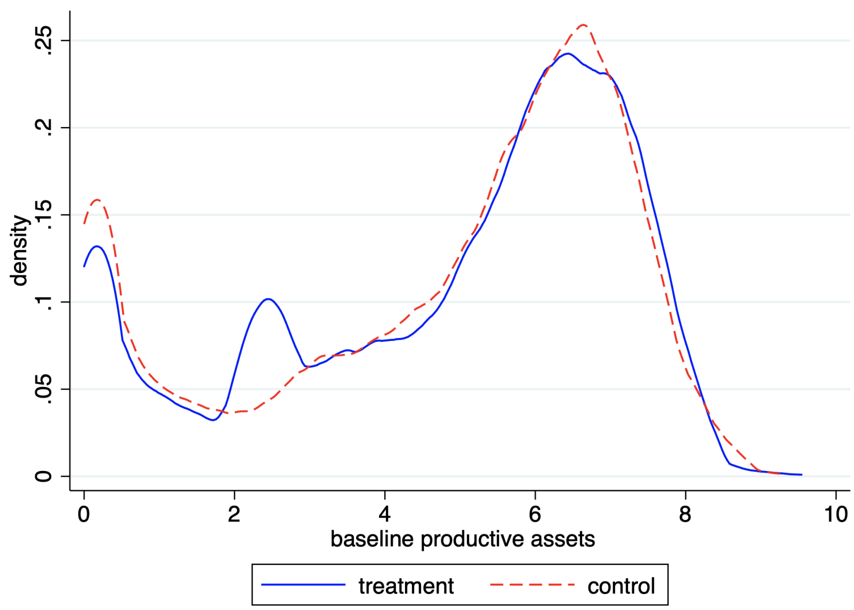
Wasserman, Larry (2006). *All of Nonparametric Statistics*. Springer Science & Business Media.

Figure 1: Distribution of Productive Assets

(a) Distribution of Productive Assets at Baseline



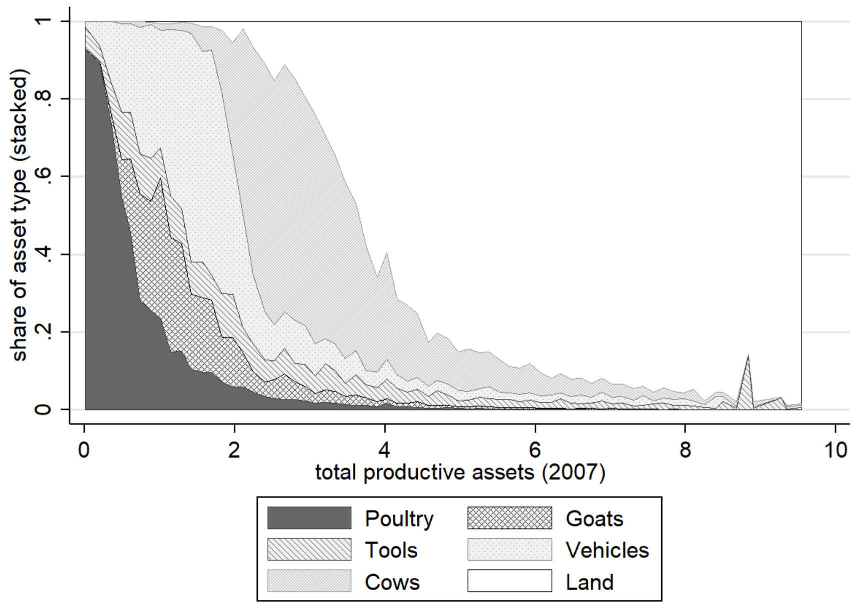
(b) Distribution of Productive Assets after Transfer



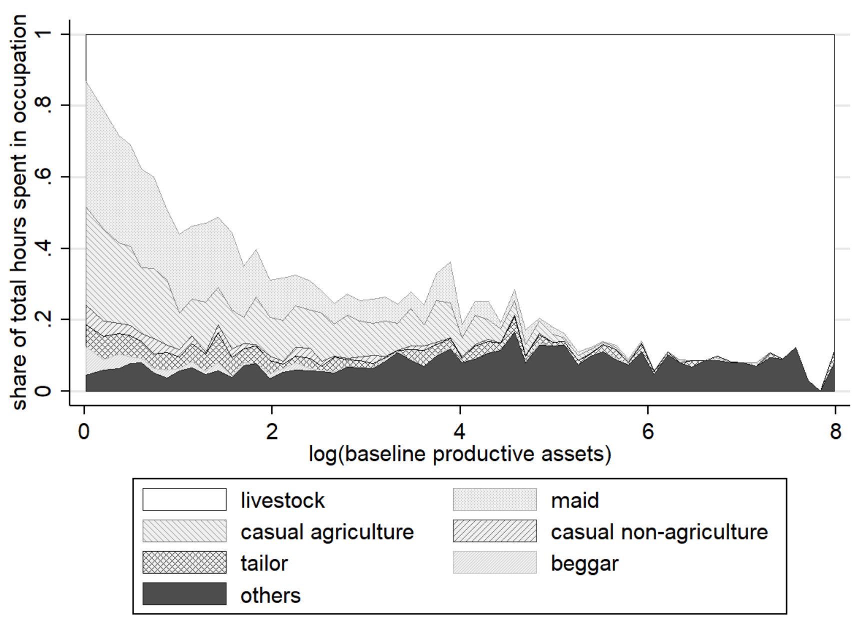
Notes: The graph shows kernel density estimates of the distribution of baseline productive assets in the full sample of 21,839 households across all wealth classes in treatment and control villages. Productive assets are measured as the natural logarithm of the total value, in 1000 Bangladeshi Taka, of all livestock, poultry, business assets, and land owned by the households. Sample weights are used to account for different sampling probabilities across wealth classes. The weights are based on a census of all households in the 1,309 study villages. Panel b) shows the post-transfer distribution. Transfers for treatment households are imputed as the median value of a cow within the catchment area of a household's BRAC branch.

Figure 2: Asset Composition and Occupation by Baseline Assets

(a) Asset Composition

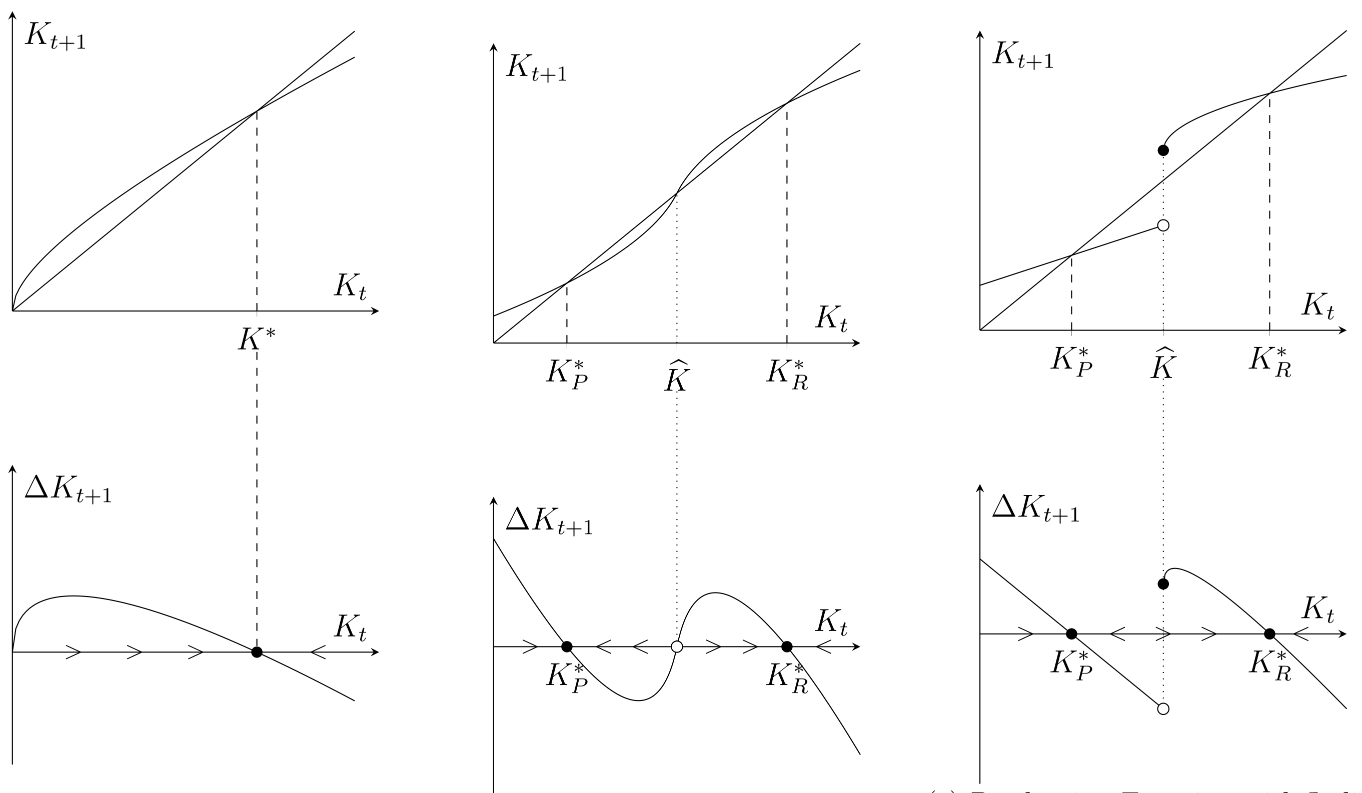


(b) Occupation and Productive Assets



Notes: The graph shows the composition of productive assets and hours spent in different occupations against baseline productive assets in the full sample of 21,839 households across all wealth classes. Productive assets are measured as the natural logarithm of the total value, in 1000 Bangladeshi Taka, of all livestock, poultry, business assets, and land owned by the households. Panel a) splits livestock into goats and cows, and business assets into tools and vehicles. In panel b) hours reportedly spent on rearing poultry are excluded. All occupations with a population average of less than 10 hours are summarised in 'others'.

Figure 3: Three Transition Equations and Implied Asset Dynamics

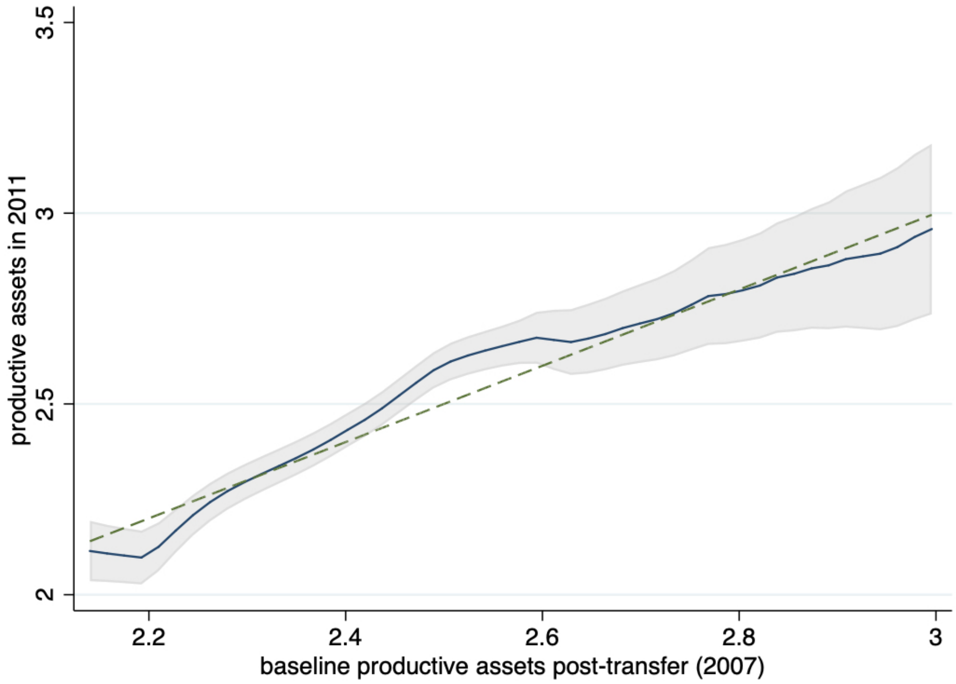


(a) Globally Concave Production Function

(b) S-shaped Production Function

(c) Production Function with Indivisibilities

Figure 4: Local Polynomial Estimate of the Transition Equation



Notes: The sample is restricted to ultra-poor households in treatment villages with log baseline productive assets below 3. Productive assets are measured as the natural logarithm of the total value, in 1000 Bangladeshi Taka, of all livestock, poultry, business assets, and land owned by the households. Post-transfer assets are imputed by adding to each household's baseline assets the median value of a cow within the catchment area of a household's BRAC branch. The blue line plots the smoothed values of a local polynomial regression with an Epanechnikov kernel of optimal bandwidth. The grey area depicts 95 percent confidence bands. The dashed line represents the 45° line at which assets in 2011 equal initial assets in 2007.

Figure 5: Endogenous response to training

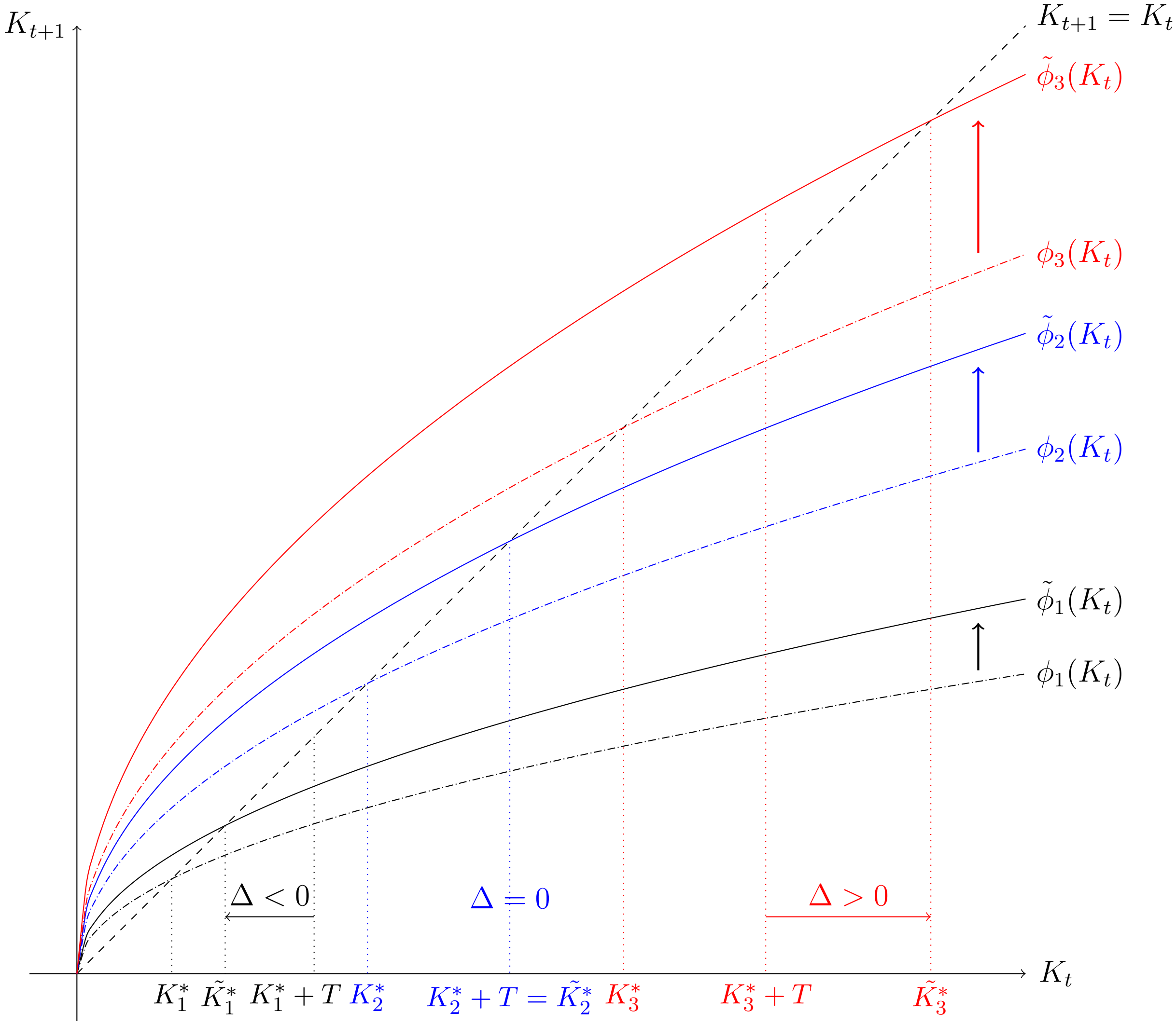
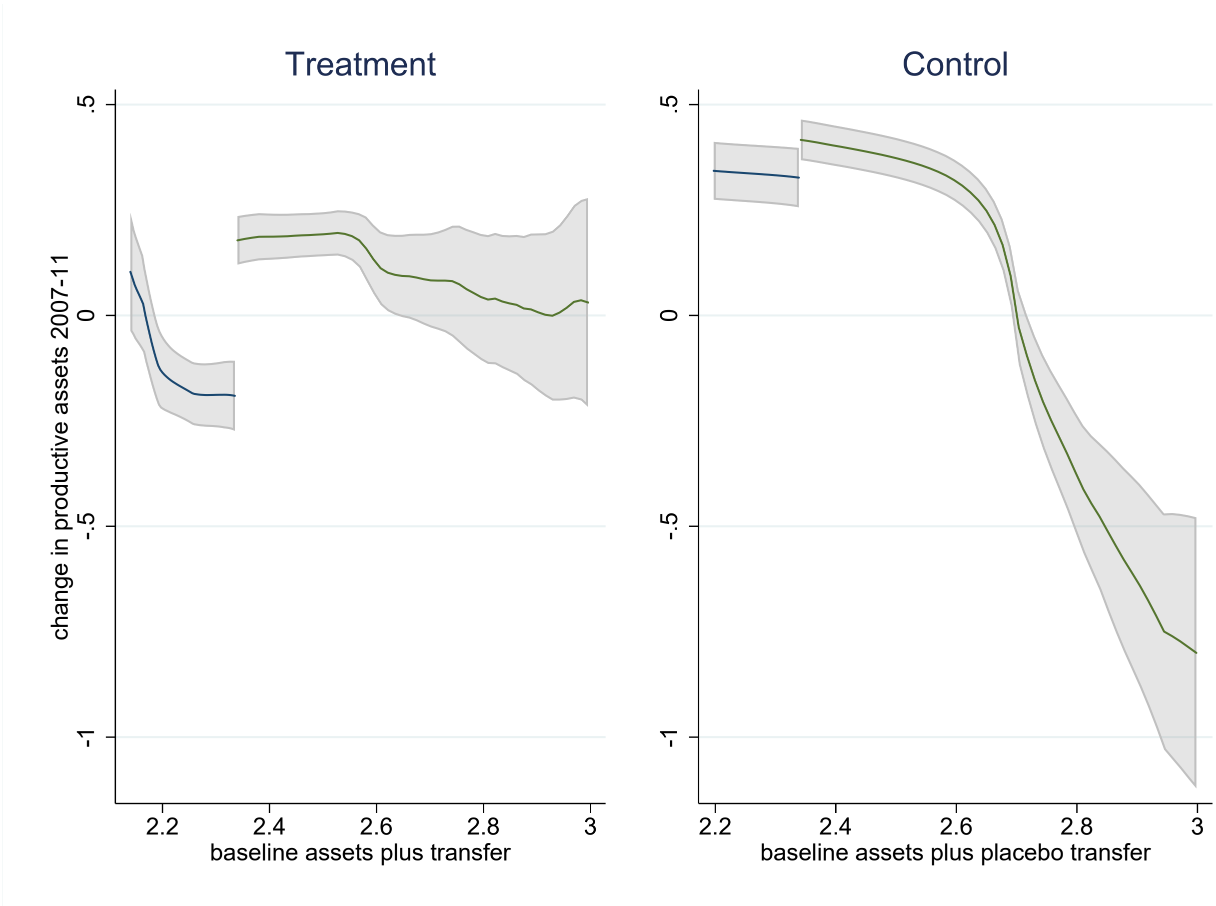
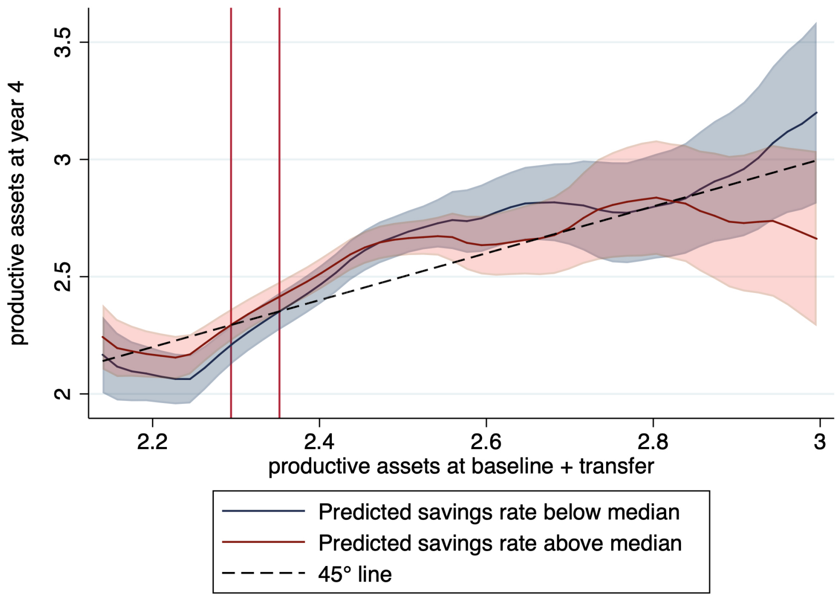


Figure 6: Change in Productive Assets

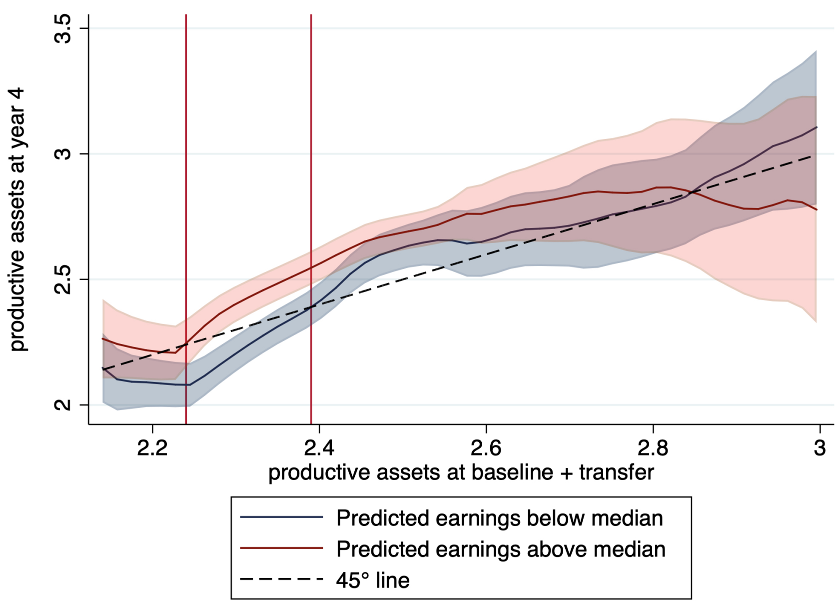


Notes: The sample is restricted to ultra-poor households in treatment and control villages with log baseline productive assets below 3. Productive assets are measured as the natural logarithm of the total value, in 1000 Bangladeshi Taka, of all livestock, poultry, business assets, and land owned by the households. Post-transfer assets for both treatment and control are imputed by adding to each household's baseline assets the median value of a cow within the catchment area of a household's BRAC branch. In the control group, this constitutes a placebo since these households didn't receive any transfers before 2011. The graphs show the smoothed values from local polynomial regressions estimated separately below and above a threshold of $\hat{k} = 2.34$. These regressions use an Epanechnikov kernel of optimal bandwidth. The grey areas depicts 95 percent confidence bands.

Figure 7: Heterogeneous Thresholds



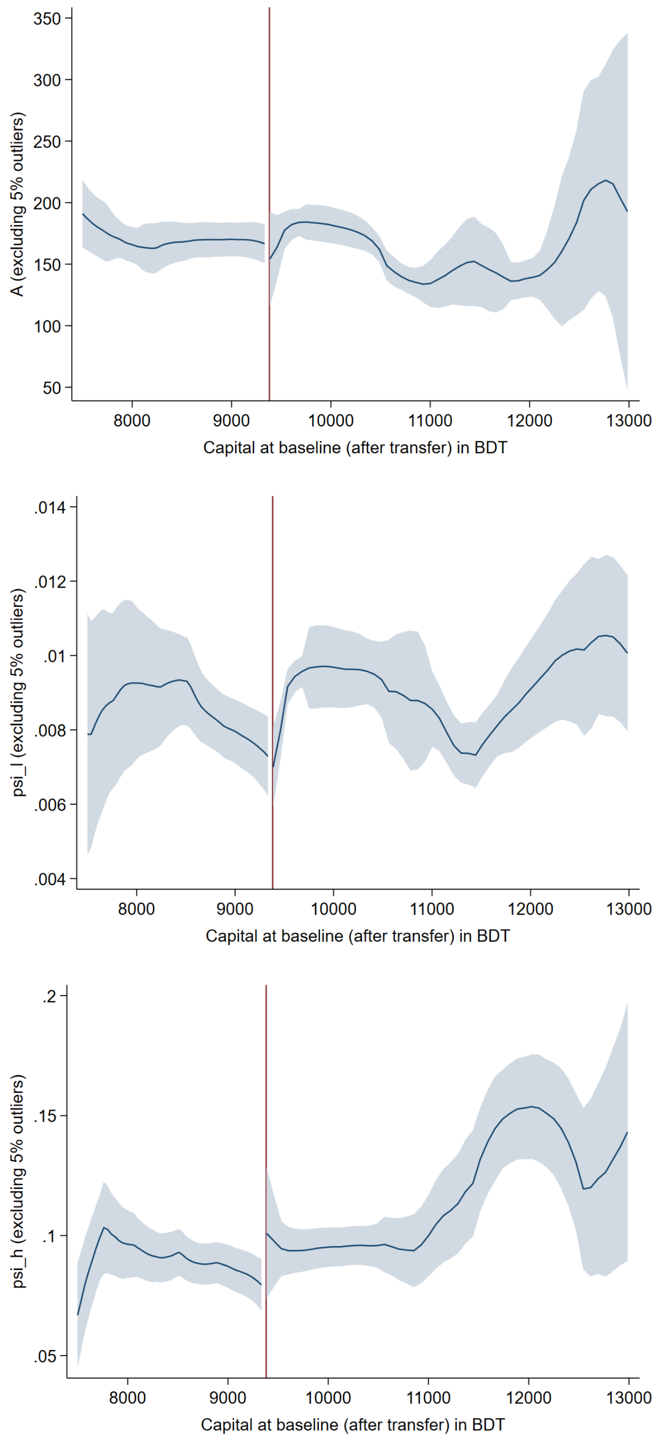
(a) Heterogeneous Thresholds: Savings Potential



(b) Heterogeneous Thresholds: Earnings Potential

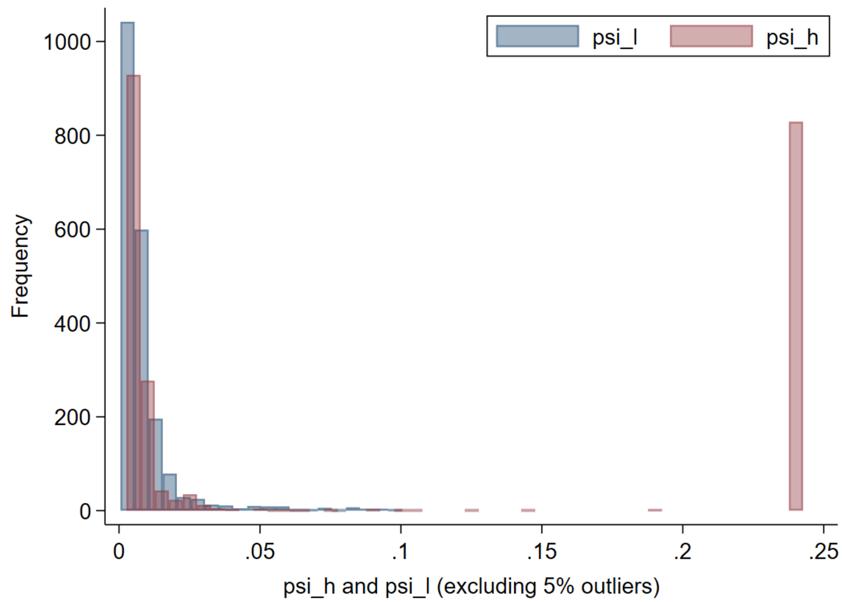
Notes: The sample and estimation method are the same as in figure 4. Panels a) and b) split the sample respectively at the median of households' predicted savings rate and earnings potential. The predicted savings rate is computed as the predicted values from regressing the observed savings rate on a constant and a fourth order polynomial of the household's dependency ratio. The latter is the ratio children (below 10), elderly (above 65), and chronically ill to total household members. Earnings potential is computed by as the residual (averaged at the branch level) from regressing livestock earnings on a constant and a second order polynomial of the number of cows owned. The vertical red lines indicating unstable steady states are at 2.29 and 2.36 in panel a), and at 2.24 and 2.39 in panel b).

Figure 8: Calibrated parameters as a function of baseline capital.



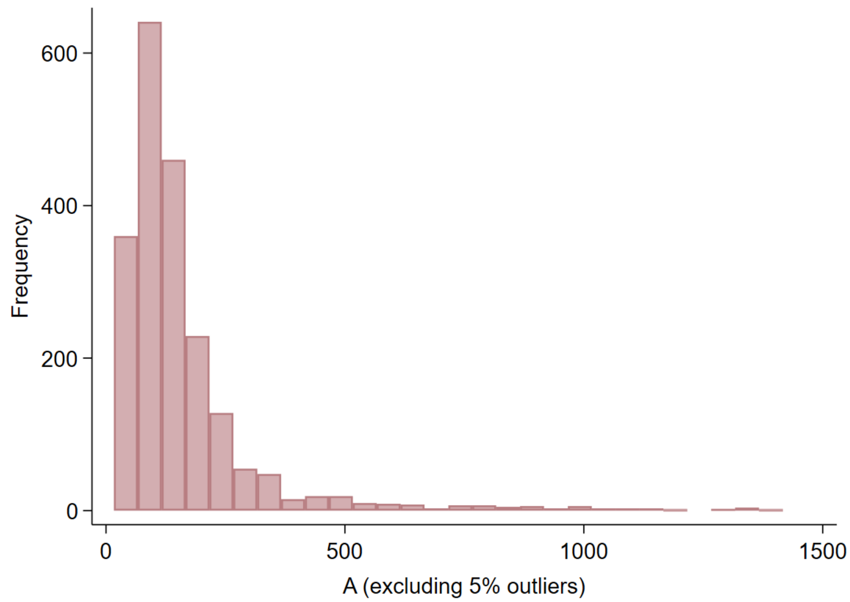
Notes: The graphs show calibrated values of individual-level parameters as a function of post-transfer baseline capital. The calibrated parameters shown are productivity in livestock rearing A (panel A), disutility of labor hours in livestock rearing (panel B), and disutility of wage labor hours (panel C). Five percent outliers are excluded. The vertical lines show the threshold level of capital. Local polynomial regressions are estimated separately on either side of the threshold. Ninety five percent asymptotic confidence intervals for the local polynomial regressions are shown.

Figure 9: Frequency distribution of calibrated disutility of labor parameters



Notes: The frequency distributions shown are of calibrated individual-level parameters for disutility of livestock rearing hours (blue) and wage labor hours (red), excluding 5% outliers, for the 64% of ultra-poor individuals for whom individual-level parameters can be calibrated using baseline and/or year 2 data (as described in the text). The upper mode in the latter frequency distribution reflects the fact that individuals who do not work at baseline are assigned the maximum calibrated value of the disutility of wage labor hours parameter.

Figure 10: Frequency distribution of calibrated productivity parameters

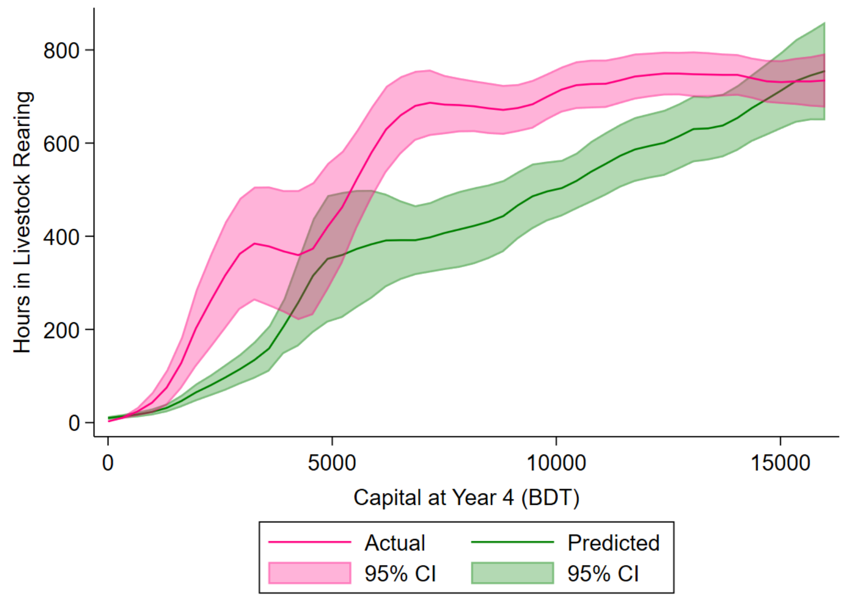


Notes: The graph shows the frequency distribution of calibrated parameters for productivity in livestock rearing, for the 64% of ultra-poor individuals for whom individual-level parameters can be calibrated using baseline and/or year 2 data (as described in the text).

Figure 11: Predicted vs. actual occupation in Year 4

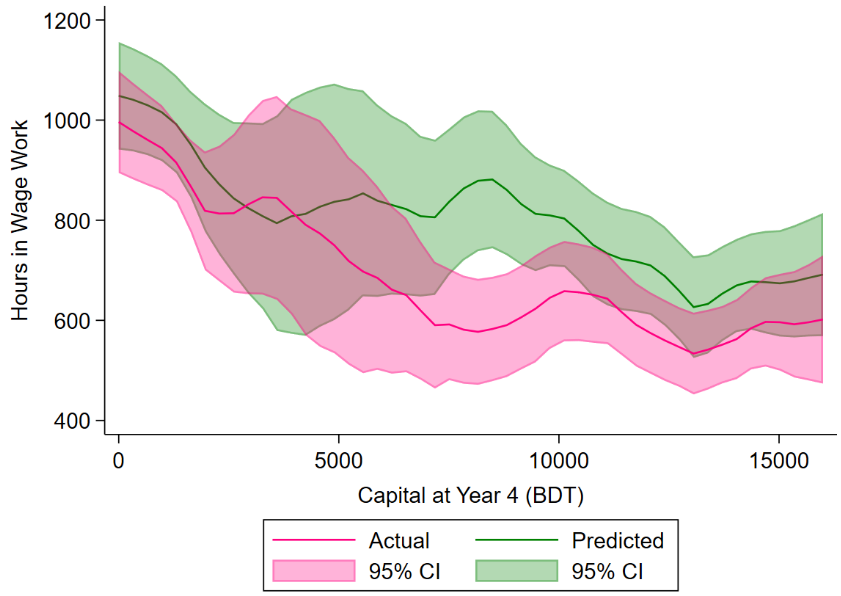
Notes: The pink bars show the observed distribution across occupations (specialization in livestock rearing, specialization in wage labor, engaging in both occupations) in year 4 for those of the 64% of ultra-poor individuals for whom individual-level parameters can be calibrated using baseline and/or year 2 data (as described in the text) who report positive labor hours at year 4. The green bars show, for the same individuals, model-implied optimal occupational choices at each individual's observed year 4 capital level.

Figure 12: Predicted vs. actual hours of livestock rearing in Year 4



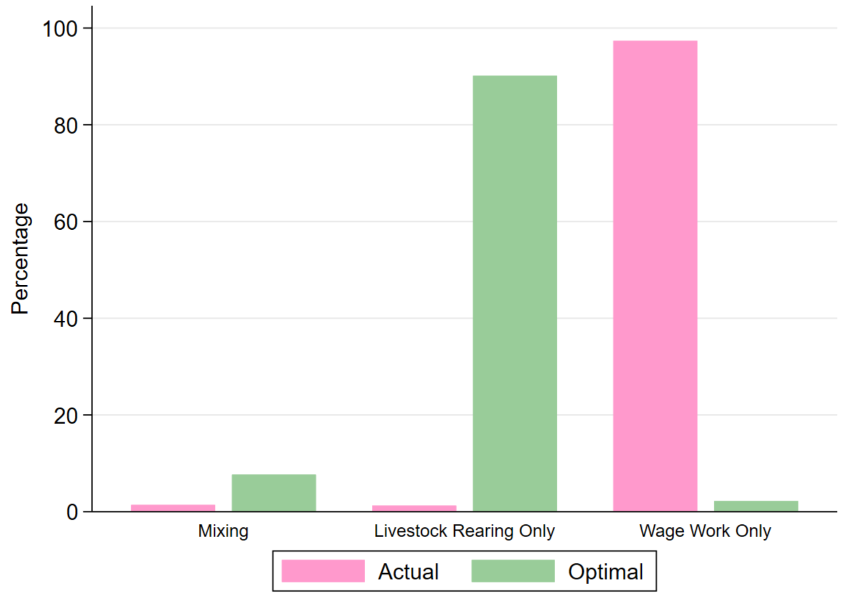
Notes: The pink graph shows local polynomial predictions of the observed hours worked in livestock rearing in year 4, as a function of year 4 capital, for those of the 64% of ultra-poor individuals for whom individual-level parameters can be calibrated using baseline and/or year 2 data (as described in the text) who report positive labor hours at year 4. The green graph shows, for the same individuals, local polynomial predictions of model-implied optimal hours worked in livestock rearing as a function of observed year 4 capital level. Ninety five percent asymptotic confidence intervals for the local polynomial regressions are shown.

Figure 13: Predicted vs. actual hours of wage labor in Year 4



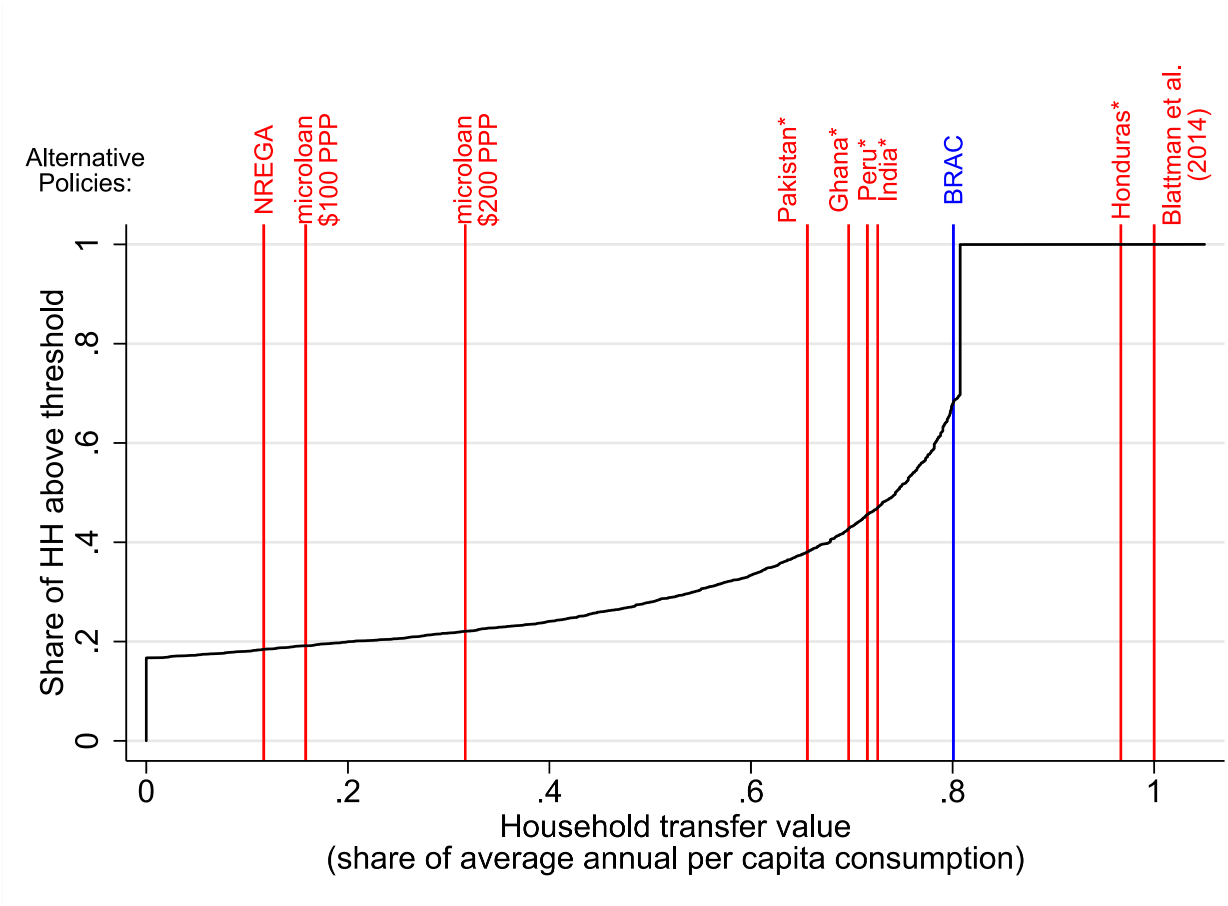
Notes: The pink graph shows local polynomial predictions of the observed wage labor hours worked in year 4, as a function of year 4 capital, for those of the 64% of ultra-poor individuals for whom individual-level parameters can be calibrated using baseline and/or year 2 data (as described in the text) who report positive labor hours at year 4. The green graph shows, for the same individuals, local polynomial predictions of model-implied optimal hours worked in wage labor as a function of observed year 4 capital level. Ninety five percent asymptotic confidence intervals for the local polynomial regressions are shown.

Figure 14: Optimal vs. actual occupation at baseline



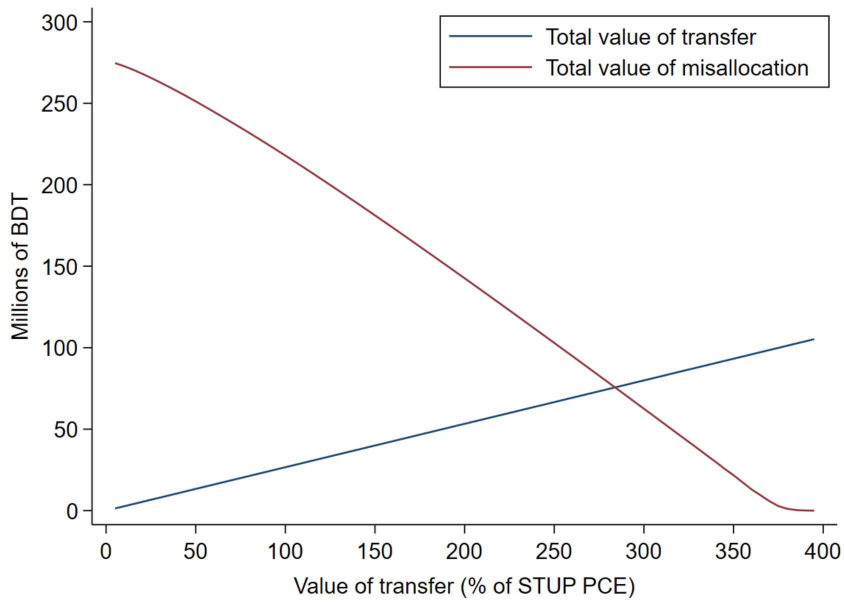
Notes: The green bars show the model-implied optimal distribution across occupations at the capital level corresponding to the upper mode of the distribution across all wealth classes of productive assets excluding land (43,701 BDT), for the 64% of ultra-poor individuals for whom individual-level parameters can be calibrated using baseline and/or year 2 data (as described in the text). The pink bars show the observed baseline distribution across occupations of those of these individuals who report positive labor hours at baseline.

Figure 15: Share of households escaping the poverty trap as a function of the transfer size



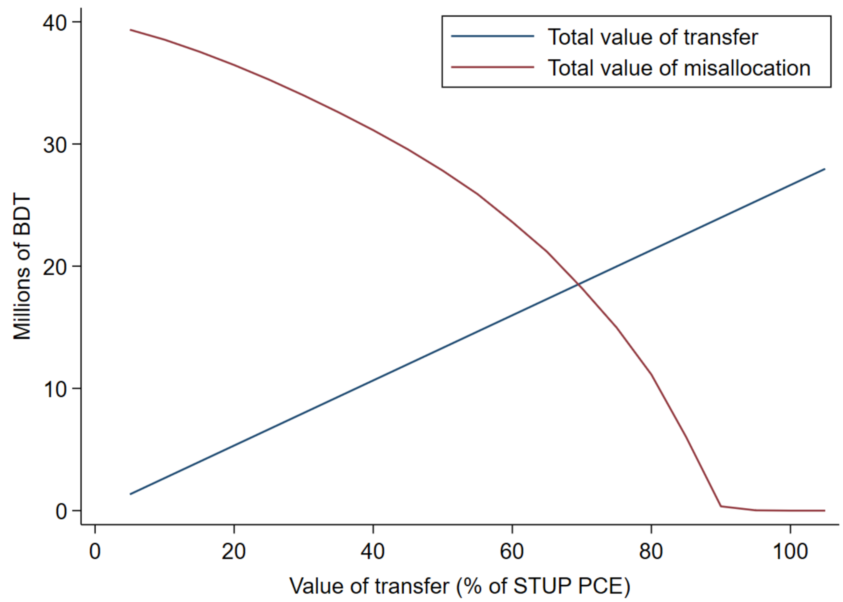
Notes: The sample includes ultra-poor households in treatment villages at baseline. The black line shows the empirical cumulative density of of the difference between the poverty threshold of $\hat{k} = 2.34$ and household's productive asset at baseline plus a shock randomly drawn from control households. Vertical lines depict different transfer sizes. The blue line shows the actual transfer, which is computed as the average of the imputed transfers we use in the main analysis. Red lines depict approximate transfer values of similar programs in the literature. The transfer size of India's National Rural Employment Guarantee Act (NREGA) is computed based on Imbert and Papp (2015) as the annual wage received when working the full 100 days to which the program is limited. (*) Country names refer to study sites of Banerjee et al. (2015b).

Figure 16: Estimated misallocation with increasing transfer value (misallocation vs upper mode of distribution of productive assets minus land)



Notes: The graph shows the model-implied total value of misallocation (red) as transfers given to all households (blue) increase in increments of percentage of annual per capita consumption expenditure. Misallocation is measured against the maximum model-implied payoff available at the capital level corresponding to the upper mode of the distribution across all wealth classes of productive assets excluding land (43,701 BDT). The top 5% of individual misallocation values are top-coded at the 95th percentile in the simulations.

Figure 17: Estimated misallocation with increasing transfer value (misallocation vs unstable steady state)



Notes: The graph shows the model-implied total value of misallocation (red) as transfers given to all households (blue) increase in increments of percentage of annual per capita consumption expenditure, where misallocation is measured against the maximum model-implied payoff available at the unstable steady state capital level. The top 5% of individual misallocation values are top-coded at the 95th percentile in the simulations.

Table 1: Descriptive Statistics

	(1) ultra-poor	(2) near poor	(3) middle class	(4) upper-class
A) Labour Outcomes				
In labour force	0.84	0.81	0.87	0.91
	(0.36)	(0.39)	(0.34)	(0.29)
Total hours worked per year	1134.31	938.53	819.82	820.79
	(888.38)	(821.22)	(639.08)	(549.77)
Total days worked per year	252.06	265.07	303.55	325.62
	(136.74)	(141.27)	(122.21)	(102.25)
Hourly income (BDT)	4.65	4.27	5.98	12.55
	(19.35)	(7.37)	(17.69)	(40.61)
B) Human and Physical Capital				
Years of formal education	0.56	1.26	1.99	3.72
	(1.63)	(2.43)	(2.99)	(3.74)
Literate	0.07	0.17	0.27	0.51
	(0.26)	(0.37)	(0.44)	(0.50)
Body mass index (BMI)	18.25	18.58	19.17	20.53
	(2.27)	(2.25)	(2.28)	(3.02)
Household Savings (1000 BDT)	0.15	0.40	1.62	8.61
	(0.83)	(1.24)	(10.62)	(29.29)
Productive assets (1000 BDT)	9.92	12.94	145.38	801.77
	(30.63)	(71.59)	(310.49)	(945.29)
Productive assets + Loans (1000 BDT)	10.53	14.83	150.23	812.83
	(31.10)	(72.47)	(312.50)	(947.65)
Observations	6732	7340	6742	2215

Notes: Standard deviations are reported in brackets. All statistics are constructed using baseline household data from both treatment and control villages. Wealth classes are based on the participatory rural assessment (PRA) exercise conducted by BRAC: the ultra-poor are ranked in the bottom wealth bins and meet the TUP program eligibility criteria, the near poor are ranked in the bottom wealth bins and do not meet the program eligibility criteria, middle-class are ranked in the middle wealth bins, and the upper classes are those ranked in the top bin.The number of households in each wealth class at baseline is reported at the bottom of the table. All outcomes, except household savings, productive assets and loans, are measured at the individual level (for the ultra-poor women in the household). The recall period is the year before the survey date. The BMI statistics trim observations with BMI above 50.

Table 2: Capital accumulation

	Dependent variable: Δ_i					
	Treatment (1)	Control (2)	Both (3)	Treatment (4)	Control (5)	Both (6)
Above $\widehat{k}$	0.296*** (0.043)	−0.010 (0.051)	−0.010 (0.057)	0.62*** (0.110)	0.35 (0.228)	0.353 (0.251)
Treatment			−0.473*** (0.059)			−0.519*** (0.268)
Above $\widehat{k}$ × Treatment			0.305*** (0.069)			0.266*** (0.271)
Baseline assets (demeaned)				−2.155*** (0.697)	−1.586 (1.346)	−1.587 (1.474)
Above $\widehat{k}$ × Baseline assets (demeaned)				1.930*** (0.728)	−0.837 (1.346)	−0.837 (1.480)
Treatment × Baseline assets (demeaned)						−0.568 (1.613)
Above $\widehat{k}$ × Treatment × Baseline assets (demeaned)						2.767* (1.630)
Constant	−0.136*** (0.033)	0.336*** (0.045)	0.336*** (0.050)	−0.446*** (0.106)	0.073 (0.227)	0.073 (0.250)
N	3,292	2,450	5,742	3,292	2,450	5,742

Notes: *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$. Standard errors in brackets. Sample: ultra-poor households in treatment and control villages with log baseline productive assets below 3. The dependent variable is the difference between log productive assets in 2011 and log of productive assets in 2007, where productive assets are defined as the total value of livestock, poultry, business assets (e.g. tools, vehicles and structures), and land. Above $\widehat{k}$ equals 1 if the baseline asset stock plus the imputed transfer is larger than 2.333, and 0 otherwise. In treatment this represents household's actual post-transfer asset stock. In control, where no transfer was received, Above $\widehat{k}$ indicates if the household would be above 2.333 if it had received a transfer. Treatment was assigned at the village level. Baseline assets (demeaned) refers to the log of household's productive assets in 2007 minus its mean among all ultra-poor households.

Table 3: Heterogeneous thresholds

	Earnings Potential			*Savings Potential*		
	Baseline (1)	FE (2)	Placebo (3)	Baseline (4)	FE (5)	Placebo (6)
Above $\widehat{k}_i$	0.301*** (0.044)	0.307*** (0.047)		0.319*** (0.045)	0.357*** (0.048)	
Above $\widehat{k}_L$			−0.268 (0.047)			0.172 (0.878)
Above $\widehat{k}_H$			0.474*** (0.072)			0.484*** (0.102)
Constant	−0.157*** (0.038)	−0.161*** (0.038)	−0.012 (0.09)	−0.154*** (0.035)	−0.177*** (0.037)	−0.262*** (0.07)
Baseline $\ln K_0$ FE	N	Y	Y	N	Y	Y
N	3,292	3,292	1,656	3,135	3,135	1,352

Notes: *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$. Standard errors in parenthesis. Sample: ultra-poor households in treatment villages with log baseline productive assets below 3. The dependent variable is the difference between log productive assets in 2011 and log of productive assets in 2007, where productive assets are defined as the total value of livestock, poultry, business assets (e.g. tools, vehicles and structures), and land. Above $\widehat{k}_i$ equals 1 if the baseline asset stock plus the imputed transfer is larger than the individual specific threshold value based on earnings potential in columns 1-3 and savings in columns 4-6. For those with below median savings (earnings potential) the individual specific threshold is at 2.36 (2.39) and for those above the median it is 2.29 (2.24) (SeeFigure 7). $\widehat{k}_{L/H}$ equals 1 if the capital stock plus the transfer is larger than the thresholds for individuals below/above the median of earnings potential in columns 1-3 and savings in columns 4-6. Columns 3 and 6 restrict the sample to households for which the high threshold applies, that is those with below median earnings potential or savings rate, respectively.

Table 4: Mechanisms

Panel A

Dependent variable:	Food PCE (1)	Calories (2)	BMI (3)	Loans (4)	Savings (5)
Above $\widehat{k}$	0.005 (0.018)	−0.004 (0.019)	−0.01 (0.007)	0.041 (0.025)	−0.005 (0.014)
Above $\widehat{k}$ × two years after transfer	−0.028 (0.026)	−0.033 (0.027)	−0.009 (0.01)	−0.028 (0.036)	−0.011 (0.02)
Above $\widehat{k}$ × four years after transfer	−0.019 (0.026)	−0.005 (0.027)	−0.004 (0.01)	−0.023 (0.036)	0.038 (0.02)
Constant	2.217*** (0.014)	7.85*** (0.024)	2.909*** (0.005)	0.127 (0.019)	−0.057*** (0.01)
N	6,023	6,021	6,064	6,402	6,402

Panel B

Dependent variable:	Poultry (1)	Goat (2)	Shed (3)	Vehicles (4)	Other Assets (5)	Cows (6)
Above $\widehat{k}$	−0.015 (0.010)	0.013 (0.019)	−0.003 (0.012)	0.037* (0.019)	0.010 (0.013)	0.0003 (0.038)
Above $\widehat{k}$ × two years after transfer	0.088*** (0.015)	0.1*** (0.027)	0.056*** (0.017)	−0.005 (0.027)	−0.01 (0.018)	0.171*** (0.054)
Above $\widehat{k}$ × four years after transfer	0.054*** (0.015)	0.1*** (0.027)	0.071*** (0.017)	−0.022 (0.027)	−0.005 (0.018)	0.575*** (0.054)
Constant	0.095*** (0.008)	0.013 (0.014)	0.008 (0.009)	0.018 (0.014)	0.004 (0.01)	0.000 (0.028)
N	6,402	6,402	6,402	6,402	6,402	6,402

Notes:*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$. Standard errors in parenthesis. The sample consists of ultra-poor households in treatment villages in 2007, 2009, and 2011. The estimates reported are based on a regression discontinuity specification within the interval of [2.24;2.44] of post-transfer assets. This regresses the outcome of interest for household i in survey wave t on an indicator for whether the the household's post-transfer assets at baseline are above 2.333 (Above $\widehat{k}$) interacted with an indicator for each survey round. All regressions also control for the non-interacted survey dummies. All dependent variables are in logs.

A Solution of the Structural Model

In this appendix we characterize the full solution of our structural model:

$$\max_{l\geq 0, h\geq 0, h'\geq 0} Af(\bar{k})g(l+h') + wh - w'h' - \frac{1}{2}(\sqrt{\psi_l}l + \sqrt{\psi_h}h)^2 \tag{1}$$

subject to

$$h \leq \overline{H} \qquad [H]$$

$$h' \leq \overline{N} \qquad [N]$$

$$h + l \leq \overline{R} \qquad [R]$$

Case 1 Mixed occupational choice with hired-in labour ($l > 0, h > 0, h' > 0$).

Case 1a All [H], [N] and [R] slack.

In this case, the optimal solution must satisfy:

$$\begin{aligned} Af(\bar{k})g'(l+h') &= \psi_l l + \sqrt{\psi_l \psi_h} h \\ w &= \sqrt{\psi_l \psi_h} l + \psi_h h \\ Af(\bar{k})g'(l+h') &= w' \end{aligned}$$

Note that this is possible under the assumption that $w > w'$ and $\psi_h > \psi_l$. We can interpret the left-hand side of the first order conditions as the marginal benefit of increasing the amount of self-employment or wage labour supplied or the amount of labour hired in (in terms of additional production or earnings), whereas the right-hand side represents the respective marginal cost. Because the agent is choosing an interior solution for these three variables, it must be that the marginal benefit is equal to the marginal cost.

Case 1b [H] binding, [N] and [R] slack.

If $h = \overline{H}$, then the optimal solution is characterised by:

$$Af(\bar{k})g'(l + h') = \psi_l l + \sqrt{\psi_l \psi_h}\,\overline{H}$$
$$h = \overline{H}$$
$$Af(\bar{k})g'(l + h') = w'$$

Moreover, because [H] is binding, we have that

$$w \geq \sqrt{\psi_l \psi_h} l + \psi_h \overline{H},$$

i.e. in the optimum the marginal benefit of wage work could be greater than the marginal cost. This might mean that the agent would like to supply more hours of paid labour, but cannot do so because of the labour demand constraint.

Case 1c [H] and [N] slack, [R] binding.

If $h < \overline{H}$ but $h + l = \overline{R}$, letting λ denote the Lagrange multiplier on the time endowment constraint, the optimal solution must satisfy

$$Af(\bar{k})g'(l + h') = \psi_l l + \sqrt{\psi_l \psi_h} h + \lambda$$
$$w = \sqrt{\psi_l \psi_h} l + \psi_h h + \lambda$$
$$Af(\bar{k})g'(l + h') = w'$$
$$h + l = \overline{R}$$

The multiplier $\lambda \geq 0$ represents the value of relaxing the binding constraint [R] at the optimum. It appears in the right-hand side of the first order conditions because, when the time endowment constraint binds, increasing the hours worked in livestock rearing implies decreasing the hours in wage labour (and vice versa). Combining the first two equations, we can characterise the solution as:

$$Af(\bar{k})g'(l + h') - \psi_l l - \sqrt{\psi_l \psi_h}(\overline{R} - l) = w - \sqrt{\psi_l \psi_h} l - \psi_h(\overline{R} - l)$$
$$h = \overline{R} - l$$
$$Af(\bar{k})g'(l + h') = w'$$

Case 1d [H] and [R] binding, [N] slack.

In this case the optimal solution is:

$$\begin{aligned} l &= \overline{R} - \overline{H} \\ h &= \overline{H} \\ Af(\bar{k})g'(\overline{R} - \overline{H} + h') &= w' \end{aligned}$$

As before, at the optimum we have

$$\begin{aligned} Af(\bar{k})g'(\overline{R} - \overline{H} + h') &\geq \psi_l(\overline{R} - \overline{H}) + \sqrt{\psi_l\psi_h}\,\overline{H} \\ w &\geq \sqrt{\psi_l\psi_h}(\overline{R} - \overline{H}) + \psi_h\overline{H} \end{aligned}$$

i.e. the marginal benefits of self-employment and wage labour could be greater than the respective marginal costs.

In all the sub-cases where [N] is binding, in the optimum we will have

$$Af(\bar{k})g'(l + \overline{N}) \geq w',$$

meaning that, because the farmer is hiring in the maximum amount of labour she can, it is possible that the marginal benefit of hiring in is still bigger than the marginal cost of doing so.

Case 1e [H] and [R] slack, [N] binding.

The optimal solution is characterised by:

$$\begin{aligned} Af(\bar{k})g'(l + \overline{N}) &= \psi_l l + \sqrt{\psi_l\psi_h}h \\ h &= \sqrt{\psi_l\psi_h}l + \psi_h h \\ h' &= \overline{N} \end{aligned}$$

Case 1f [R] slack, [H] and [N] binding.

The optimal solution is given by:

$$
\begin{aligned}
Af(\bar{k})g'(l+\overline{N}) &= \psi_l l + \sqrt{\psi_l\psi_h}\,\overline{H} \\
h &= \overline{H} \\
h' &= \overline{N}
\end{aligned}
$$

Case 1g [H] slack, [R] and [N] binding.

The optimal solution must satisfy:

$$
\begin{aligned}
Af(\bar{k})g'(l+\overline{N}) - \psi_l l - \sqrt{\psi_l\psi_h}(\overline{R}-l) &= w - \sqrt{\psi_l\psi_h}l - \psi_h(\overline{R}-l) \\
h &= \overline{R}-l \\
h' &= \overline{N}
\end{aligned}
$$

Case 1h All [H], [N] and [R] binding.

The optimal solution is:

$$
\begin{aligned}
l &= \overline{R}-\overline{H} \\
h &= \overline{H} \\
h' &= \overline{N}
\end{aligned}
$$

Case 2 Mixed occupational choice without hired-in labour ($l>0, h>0, h'=0$).

In all the sub-cases below, because $h'=0$, necessarily we have

$$
Af(\bar{k})g'(l) \leq w'
$$

This means that, as no labour is being hired in, the marginal benefit of doing so must be less than the marginal cost. Also, [N] is slack because $\overline{N}>0=h'$.

Case 2a Both [H] and [R] slack.

In this case, the optimal solution must satisfy:

$$
\begin{aligned}
Af(\bar{k})g'(l+h') &= \psi_l l + \sqrt{\psi_l\psi_h}h \\
w &= \sqrt{\psi_l\psi_h}l + \psi_h h \\
h' &= 0
\end{aligned}
$$

Case 2b [H] binding, [R] slack.

The optimal solution is characterised by:

$$\begin{aligned} Af(\bar{k})g'(l+h') &= \psi_l l + \sqrt{\psi_l \psi_h}\,\overline{H} \\ h &= \overline{H} \\ h' &= 0 \end{aligned}$$

with

$$w \geq \sqrt{\psi_l \psi_h} l + \psi_h \overline{H}$$

Case 2c [H] slack, [R] binding.

By the same argument as above, in the optimum we must have:

$$\begin{aligned} Af(\bar{k})g'(l+h') - \psi_l l - \sqrt{\psi_l \psi_h}(\overline{R} - l) &= w - \sqrt{\psi_l \psi_h} l - \psi_h(\overline{R} - l) \\ h &= \overline{R} - l \\ h' &= 0 \end{aligned}$$

Case 2d Both [H] and [R] binding. The optimal solution is:

$$\begin{aligned} l &= \overline{R} - \overline{H} \\ h &= \overline{H} \\ h' &= 0 \end{aligned}$$

with

$$\begin{aligned} Af(\bar{k})g'(\overline{R} - \overline{H}) &\geq \psi_l(\overline{R} - \overline{H}) + \sqrt{\psi_l \psi_h}\,\overline{H} \\ w &\geq \sqrt{\psi_l \psi_h}(\overline{R} - \overline{H}) + \psi_h \overline{H} \end{aligned}$$

We turn now to cases where the agent does only livestock rearing (self-employment) and no wage labour. Because $h = 0$, it must be the case that

$$w \leq \sqrt{\psi_l \psi_h} l + \lambda$$

at the optimum, where λ is again the Lagrange multiplier on the time endowment constraint (and $\lambda = 0$ if the constraint is slack). This mean that, even at $h = 0$, the marginal cost of supplying hours of paid work is higher than the marginal benefit. Also, note that the labour demand constraint [H] will always be slack, as $h = 0 < \overline{H}$.

Case 3 Livestock rearing only with hired-in labour ($l > 0, h = 0, h' > 0$).

Case 3a Both [R] and [N] slack.
The optimal solution must satisfy:

$$\begin{aligned} Af(\bar{k})g'(l + h') &= \psi_l l \\ h &= 0 \\ Af(\bar{k})g'(l + h') &= w' \end{aligned}$$

Case 3b [R] binding, [N] slack.
The optimal solution is given by:

$$\begin{aligned} l &= \overline{R} \\ h &= 0 \\ Af(\bar{k})g'(\overline{R} + h') &= w' \end{aligned}$$

with

$$Af(\bar{k})g'(\overline{R} + h') \geq \psi_l \overline{R}$$

Case 3c [R] slack, [N] binding.
At the optimum we must have:

$$\begin{aligned} Af(\bar{k})g'(l + \overline{N}) &= \psi_l l \\ h &= 0 \\ h' &= \overline{N} \end{aligned}$$

Case 3d Both [R] and [N] binding.

The optimal solution is:

$$l = \overline{R}$$
$$h = 0$$
$$h' = \overline{N}$$

Case 4 Livestock rearing only without hired-in labour ($l > 0, h = 0, h' = 0$).

Again, because $h' = 0$, we must have

$$Af(\bar{k})g'(l) \leq w'$$

at the optimum.

Case 4a [R] slack.

The optimal solution must satisfy:

$$Af(\bar{k})g'(l) = \psi_l l$$
$$h = 0$$
$$h' = 0$$

Case 4b [R] binding.

The optimal solution is:

$$l = \overline{R}$$
$$h = 0$$
$$h' = 0$$

with

$$Af(\bar{k})g'(R) \geq \psi_l \overline{R}$$

Next, we examine the cases where the agent does only wage labour but no livestock rearing herself. Because $l = 0$, we necessarily have that

$$Af(\bar{k})g'(h') \leq \sqrt{\psi_l \psi_h} h$$

Notice also that, since $h \leq \overline{H} \leq \overline{R}$, the time endowment constraint [R] is automatically slack.

Case 5 Wage work only with hired-in labour ($l = 0, h > 0, h' > 0$).

Case 5a Both [H] and [N] slack.
The optimal solution is given by:

$$l = 0$$
$$w = \psi_h h$$
$$Af(\bar{k})g'(h') = w'$$

Case 5b [H] binding, [N] slack. The optimum must satisfy:

$$l = 0$$
$$h = \overline{H}$$
$$Af(\bar{k})g'(h') = w'$$

with

$$w \geq \psi_h \overline{H}$$

Case 5c [H] slack, [N] binding. The optimum must satisfy:

$$l = 0$$
$$w = \psi_h h$$
$$h' = \overline{N}$$

Case 5d Both [H] and [N] binding. The optimum must satisfy:

$$l = 0$$
$$h = \overline{H}$$
$$h' = \overline{N}$$

In the last possible case, we have that $l + h' = 0$. Under standard regularity conditions (in particular, if we assume $g'(0) = +\infty$) this would never be an optimal choice. This is because the marginal return of starting any livestock rearing (either through self-employment or by hiring in external labour) is arbitrarily large, whereas the marginal cost is only finite.

However, to allow for this case, we consider the possibility of liquidating the physical capital stock, which would yield a profit of $\rho\bar{k}$ with $\rho \leq 1$. Hence, the problem that the agent faces is just a choice of hours of paid work:

$$\max_{h\geq 0} \rho\bar{k} + wh - \frac{1}{2}\psi_h h^2 \tag{2}$$

subject to

$$h \leq \overline{H} \tag{H}$$

Case 6 Wage work only without hired-in labour ($l = 0, h > 0, h' = 0$).

Case 6a [H] slack.
The optimality condition is

$$w = \psi_h h$$

Case 6b [H] binding. In this case we must have

$$h = \overline{H}$$

and $w \geq \psi_h \overline{H}$.

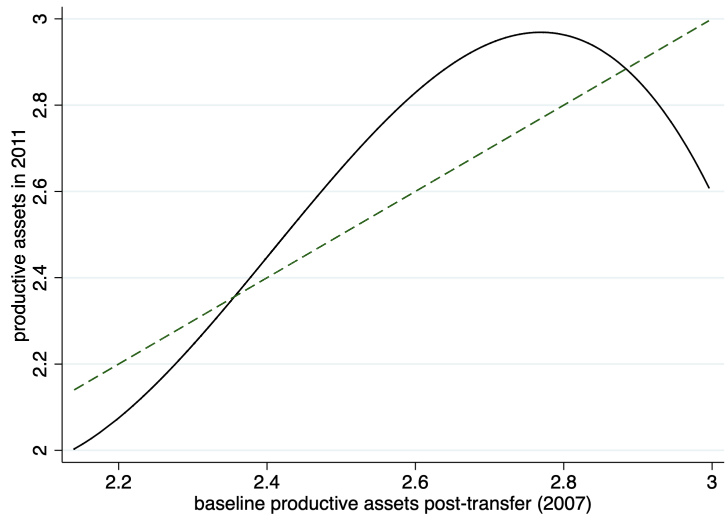
The above will be optimal when the solution to the maximization problem in (2) yields a higher payoff than the outcome of (1).

Finally, we note that with this parametrisation it is not possible to have $l = 0$ and $h = 0$ at the same time, because at those levels, the marginal cost of supplying wage labour is 0, whereas the marginal benefit is $w > 0$. However, this case seems to be empirically relevant.

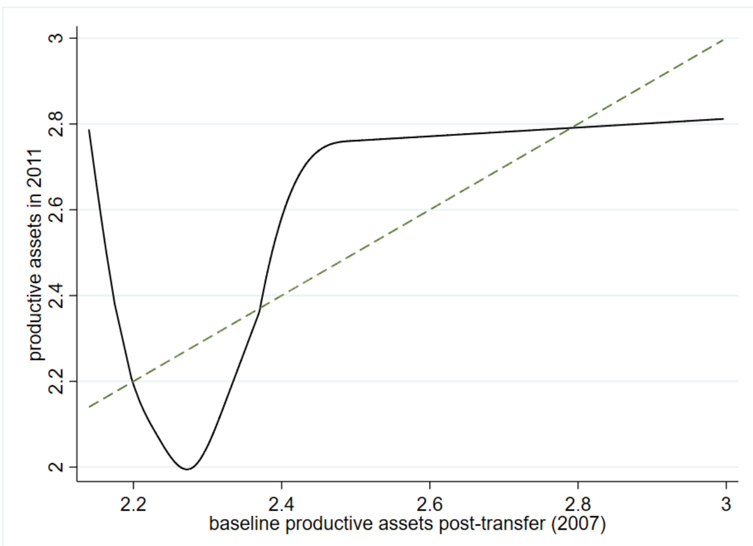
B Appendix Figures

Figure 18: Alternative Estimates of the Transition Equation

(a) 3rd Order Polynomial

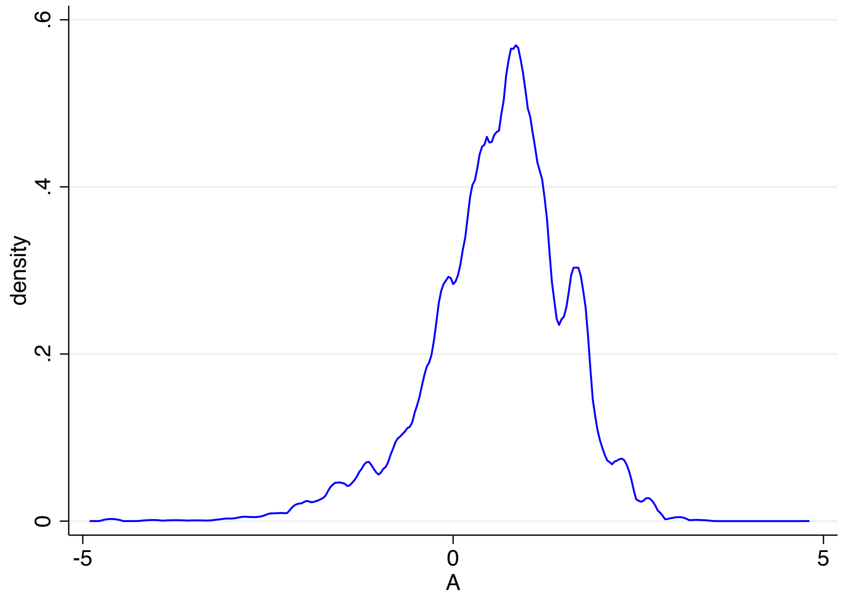


(b) B-splines

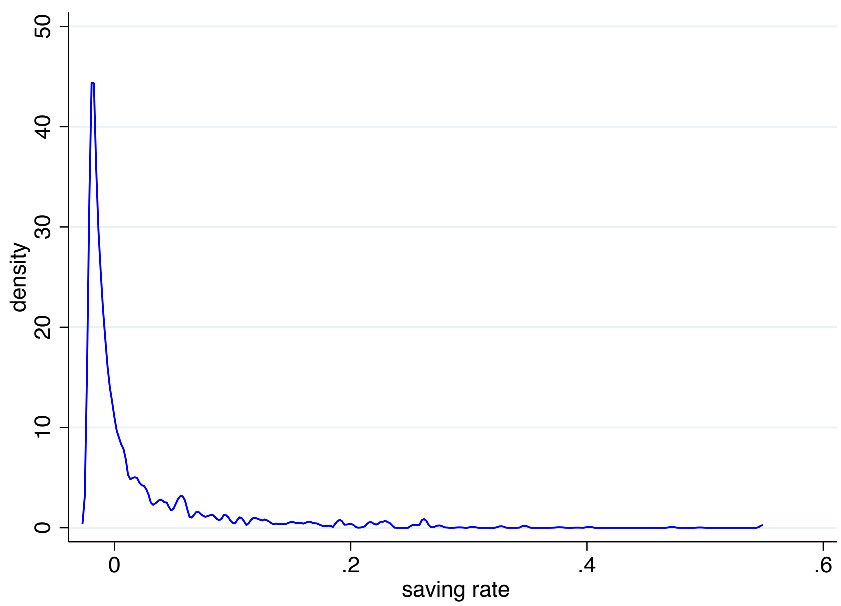


Notes: The sample is restricted to ultra-poor households in treatment villages with log baseline productive assets below 3. Productive assets are measured as the natural logarithm of the total value, in 1000 Bangladeshi Taka, of all livestock, poultry, business assets, and land owned by the households. Post-transfer assets are imputed by adding to each household's baseline assets the median value of a cow within the catchment area of a household's BRAC branch. The dashed line represents the 45° line at which assets in 2011 equal initial assets in 2007. Panel a) plots the predicted values of a regression of log productive assets in 2014 on a third order polynomial of log productive assets including the transfer in 2011. Panel b) shows a B-spline estimate of the same relationship.

Figure 19: What explains bimodal distribution of Assets?



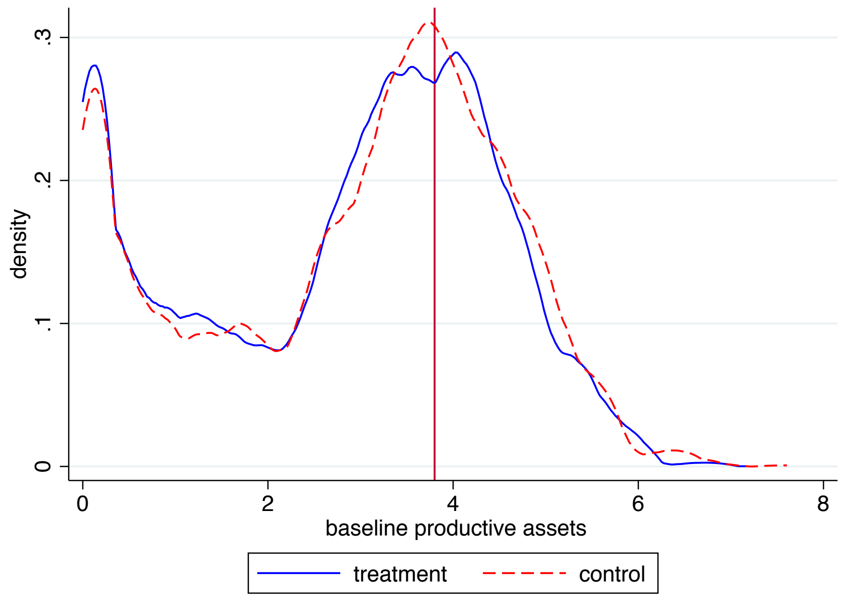
(a) Distribution of Productivity



(b) Distribution of savings rate

Notes: The graph shows kernel density estimates of the distribution of households' livestock rearing productivity (panel (a)) and savings rate (panel (b)) for all surveyed households in treatment villages. Household level productivity estimates are obtained by regressing log livestock income on log hours worked in livestock rearing, and the log of the number of cows, controlling for survey round, BRAC branch, and individual fixed effects in a panel over the survey rounds 2007, 2009, 2011, and 2014. We interpret the individual fixed effects from this regression as household productivity.

Figure 20: Distribution of productive assets excluding land



Notes: The graph shows kernel density estimates of the distribution of baseline productive assets excluding land in the full sample of 21,839 households across all wealth classes in treatment and control villages. Productive assets without land include all livestock, poultry, and business assets owned by the household. Sample weights are used to account for different sampling probabilities across wealth classes. The weights are based on a census of all households in the 1,309 study villages.

Figure 21: When $sAf(\widehat{k}) - \delta\widehat{k} > 0$, the threshold is $\widehat{k}$.

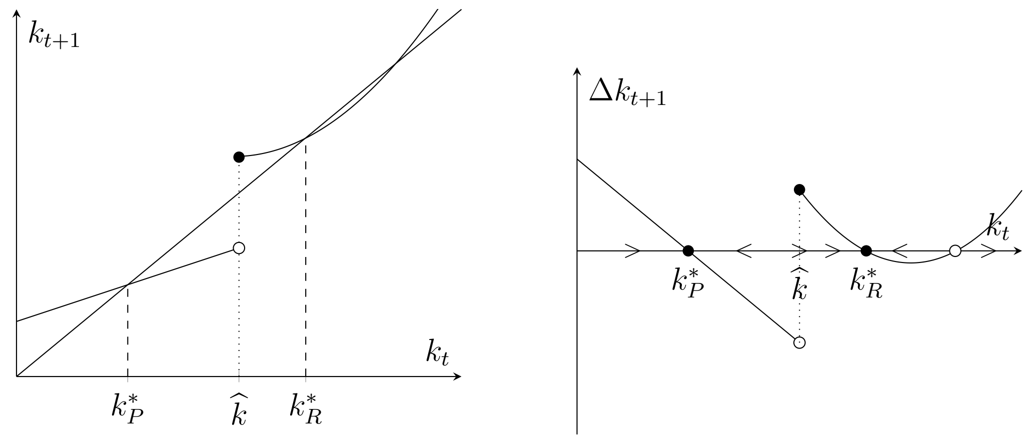


Figure 22: When $sAf(\widehat{k}) - \delta\widehat{k} < 0$, the threshold is $k^u > \widehat{k}$.