Oriana Bandiera, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, Munshi Sulaiman

Labor markets and poverty in village economies

Article (Accepted version) (Refereed)

Original citation:
Bandiera, Oriana and Burgess, Robin and Das, Narayan and Gulesci, Selim and Rasul, Imran and Sulaiman, Munshi (2017) *Labor markets and poverty in village economies.* Quarterly Journal of Economics, 132 (2). pp. 811-870. ISSN 0033-5533

DOI: 10.1093/qje/qjx003

LABOR MARKETS AND POVERTY IN VILLAGE ECONOMIES*

Oriana Bandiera, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, Munshi Sulaiman

July 2016

Abstract

We study how women's choices over labor activities in village economies correlate with poverty and whether enabling the poorest women to take on the activities of their richer counterparts can set them on a sustainable trajectory out of poverty. To do this we conduct a large-scale randomized control trial, covering over 21,000 households in 1,309 villages surveyed four times over a seven year period, to evaluate a nationwide program in Bangladesh that transfers livestock assets and skills to the poorest women. At baseline, the poorest women mostly engage in low return and seasonal casual wage labor while wealthier women solely engage in livestock rearing. The program enables poor women to start engaging in livestock rearing, increasing their aggregate labor supply and earnings. This leads to asset accumulation (livestock, land and business assets) and poverty reduction, both sustained after four and seven years. These gains do not crowd out the livestock businesses of non-eligible households while the wages these receive for casual jobs increase as the poor reduce their labor supply. Our results show that: (i) the poor are able to take on the work activities of the non-poor but face barriers to doing so, and, (ii) one-off interventions that remove these barriers lead to sustainable poverty reduction. *JEL Classification: J22, O12.*

*Earlier drafts were circulated under the titles "Can Basic Entrepreneurship Transform the Economic Lives of the Poor?" and "The Misallocation of Labor in Village Economies". We thank all BRAC staff and especially Sir Fazle Abed, Mushtaque Chowdhury, Mahabub Hossain, W.M.H. Jaim, Imran Matin, Anna Minj, Muhammad Musa and Rabeya Yasmin for their collaborative efforts in this project. We thank Wahiduddin Mahmud and Hafeez Rahman of the IGC Bangladesh office and Clare Balboni for outstanding research assistance. We thank the editor, Larry Katz, three anonymous referees as well as Arun Advani, Orazio Attanasio, Abhijit Banerjee, Timothy Besley, Gharad Bryan, Francisco Buera, Anne Case, Arun Chandrasekhar, Jonathan Colmer, Angus Deaton, Dave Donaldson, Esther Duflo, Pascaline Dupas, Greg Fischer, Doug Gollin, Chang-Tai Hsieh, Dean Karlan, Eliana La Ferrara, Costas Meghir, Ted Miguel, Mushfiq Mobarak, Benjamin Olken, Michael Peters, Steve Pischke, Mark Rosenzweig, Esteban Rossi-Hansberg, Juan Pablo Rud, Jeremy Shapiro, Chris Udry, Chris Woodruff and numerous seminar and conference participants for useful comments. This project was financed by BRAC and its CFPR-TUP donors, including DFID, AusAID, CIDA and NOVIB, OXFAM-AMERICA. This document is an output from research funding by the DFID as part of the iiG. Support was also provided by the IGC. The views expressed are not necessarily those of DFID. Rasul acknowledges funding by the ESRC Centre for the Microeconomic Analysis of Public Policy at IFS (grant number RES-544-28-5001). All errors remain our own.

I. Introduction

As of today around a billion people are deemed to be living in extreme poverty. Since labor is their primary endowment, attempts to lift them out of poverty require us to understand the link between poverty and labor markets and whether policy interventions that move them into higher return labor activities can set them on a sustainable trajectory out of poverty. To shed light on the issue we combine a detailed labor survey that tracks over 21,000 households, drawn from the entire wealth distribution in 1,309 rural Bangladeshi villages, four times over a seven year period, with the randomized evaluation of the nationwide roll-out of a program that transfers assets and skills to the poorest women in these villages.

Our survey gathers detailed data on hours worked, days worked and earnings for each labor activity of each household member. We find that, at baseline, the choice of labor activity for women is limited as they allocate over 80% of hours worked to three activities: maid services, agricultural labor and livestock rearing. These labor activities are strongly correlated with poverty: poor women engage mostly in casual wage labor as maids and agricultural laborers, while wealthier women specialize in livestock rearing. The main differences between these activities are that the returns to casual wage labor are lower and work is only available on some days of the year. Consequently, we find that poor women work two months less per year than wealthier women. These findings are consistent with evidence in other settings where the rural landless poor are employed in low-pay and insecure activities (Bardhan 1984a; Dreze 1988; Dreze and Sen 1991; Rose 1999; Kaur 2015).[1]

The key question we examine in the paper is whether enabling the poorest women to take on the same work activities as the better off women in their villages can set them on a sustainable path out of poverty. To answer this question we evaluate BRAC's Targeting the Ultra-Poor (TUP) program that provides a one-off transfer of assets and skills to the poorest women with

[1] According to the Indian National Sample Survey (NSS), 46% of the female rural workforce have agricultural wage employment as their main occupation. As is also the case in our setting for maids and agricultural laborers, 98% of agricultural wage employment is through casual employment typified by spot markets (Kaur 2015). On the fact that such agricultural wage employment is only available on some days of the year, Khandker and Mahmud (2012) and Bryan *et al.* (2014) document how lean seasons between planting and harvesting are observed throughout South Asia and Sub-Saharan Africa, and are characterized by a lack of demand for casual wage labor and higher grain prices as food becomes scarce. As a result, households face extreme poverty and food insecurity.

the aim of instigating occupational change. Intuitively, if the poor face barriers to entering high return work activities and this is what keeps them in poverty, we expect program beneficiaries to change their labor allocation and escape poverty once such barriers are removed. Because the intervention is bundled, however, we cannot measure the separate relevance of credit constraints and skills constraints, both of which could be relaxed by the program.[2] Of course, the one-off asset transfer mechanically reduces poverty in the very short run because it makes beneficiaries instantly wealthier and they can consume that wealth. The question of interest here is whether such one-off asset and skills transfers set the poorest households on a sustainable trajectory out of poverty, where their consumption and asset holdings keep increasing long after the one-off transfer, as they are able to alter their labor allocation permanently.

To evaluate the causal impacts of the program, we randomly assign forty BRAC branch offices serving 1,309 villages to either treatment or control for four years. A participatory wealth ranking is conducted before baseline in both treatment and control villages, followed by the application of TUP eligibility criteria by BRAC officers. This process classifies households into four groups in all villages: ultra-poor, near-poor, middle class and upper class. Ultra-poor households, who account for 6% of the population, are eligible to receive the program whereas other households are ineligible. We survey all the ultra-poor and near-poor households and a 10% sample of the middle and upper class households. Our design is thus a partial population experiment (Moffitt 2001) that allows us to identify indirect treatment effects on ineligible households at different points of the wealth distribution as well as distributional effects, namely the extent to which the ultra-poor close the gap with the next wealth class. This is relevant because the program aims to induce occupational change among ultra-poor women to take on the same work activities as richer women (livestock rearing). It is thus natural to trace through the economic impacts on richer women as they face increased competition in output markets for livestock produce, and in markets related to inputs into livestock rearing.

We find the program transforms the labor activity choices of ultra-poor women. Four years after

[2]Indeed, this is a bundled, multi-faceted program that also provides some consumption support in the first 40 weeks post asset transfers, as well as health support and training on legal, social and political rights across the two years of the program. As discussed throughout, we do not aim to separate out the impacts of each component.

the transfer, they devote 217% more hours to livestock rearing, 17% fewer hours to agricultural labor and 26% fewer hours to maid services relative to their counterparts in control villages. Aggregating across labor activities, there is a net positive effect on hours worked and days worked of 17% and 22%, respectively, suggesting poor women had idle work capacity and that the program enables them to put it to a productive use by taking on livestock rearing activities. Overall, the results demonstrate that the poor are able to take on the labor activities of the non-poor but face barriers to doing so, which the one-off asset and skills transfers from the program relax.

The reallocation of labor supply across work activities by the ultra-poor leads their earnings to be 21% higher than their counterparts in control villages, and the probability of being below the $1.25 extreme poverty line is 14% lower. Per capita consumption expenditure is 11% higher and the value of household durables is 57% higher, with both effects being larger after four years than after two. In line with this, earnings from livestock rearing are not entirely consumed, but used to save and invest further in productive assets. Four years post-transfer, the ultra-poor in treatment villages have more than four times the amount of savings and they are more likely to receive and give loans to other households. Moreover, the value of cows they own is over twice as large (net of the value of the asset transfer itself) and they also accumulate business assets such as livestock sheds, rickshaws, vans, pumps and trees whose value is over 159% larger than for the controls over the same period.[3] More importantly they gain access to land, which is the key productive asset in these villages. Relative to controls, treatment households are 139% more likely to rent land, 45% more likely to own land and the value of their landholdings is 82% higher.

Since individuals are likely to differ in their ability to raise livestock and manage a small business, the effect of the program is likely to be heterogeneous. The scale of our evaluation, covering more than $6,000$ ultra-poor households, allows us to estimate quantile treatment effects (QTE). These indeed reveal a large degree of heterogeneity: the effect on the 95th centile of consumption, for instance, is ten times larger than the effect on the 5th centile and differences for savings and productive assets are even larger.

[3]Land is the key asset in the densely populated rural areas of Bangladesh we study. Laboring for others is necessary, in part, because the ultra-poor do not have access to land and livestock rearing is a viable alternative, in part, because it does not require a land input (Bardhan 1984a).

The effects of the program on the labor allocation of the beneficiaries raise the possibility that ineligible households residing in treatment villages might be affected through general equilibrium effects, such as changes in livestock produce prices. Our estimates of the indirect treatment effects on ineligibles however show no evidence that the livestock rearing businesses of richer women are crowded out by the entrance of the poor into this activity: they neither reduce their labor supply nor experience a significant reduction in earnings. A likely explanation for these muted impacts is that even after four years, the ultra-poor still constitute a relatively small share of the market overall. In contrast, we do find general equilibrium impacts on the casual wage labor activities that the ultra-poor dominated at baseline: after four years, the agricultural and maid wages paid to ineligible women in treatment villages are 9% and 11% higher than in control. At the same time, the hours the ineligible devote to these work activities are lower, so their earnings are unaffected.

The partial population experiment design also allows us to estimate treatment effects of the program on the gap between wealth classes and so shed light on the distributional consequences of the intervention. This exercise reveals that the ultra-poor close the gap with the near-poor in consumption expenditures and household assets, while on other dimensions they actually overtake this group and end up with four times the level of savings and twice the value of productive assets. The program thus has powerful distributional impacts, both between wealth classes as well as within the ultra-poor, as highlighted by the quantile treatment effect estimates.

At a combined cost of USD $1,120$ in purchasing power parity (PPP) terms per household, both the asset and skill components constitute large transfers benchmarked against the baseline wealth and human capital of the ultra-poor.[4] We can use our estimates to benchmark the program's benefits against its costs. Under the assumption that the estimated consumption benefits at year four are repeated over 20 years, the program has an average benefit/cost ratio of 3.2. The estimated internal rate of return (IRR) of the program is between 16% and 22%, depending on the assumed opportunity cost of time that must be taken into account as the program causes the ultra-poor's labor supply to increase overall.

The final part of the analysis sheds light on long term impacts of the intervention. To do so

[4]Throughout the paper we stick to the convention of reporting values in USD PPP terms.

we surveyed the same households again in 2014, seven years after the intervention began. While 20% of the control group residing in 49% of the control villages had been treated by then, we are able to derive a lower bound for the effect of the program after seven years, as well as compute other bounds by using our QTE estimates to create counterfactuals for the treated controls. This comparison reveals that changes after seven years are at least as large as the four year impacts. While these results must be interpreted with caution as our counterfactuals might be imperfect, a major trend break would be needed to reverse the conclusion that the original beneficiaries are escaping poverty at a steady rate.

Overall the results show that one-off asset and skills transfers to the ultra-poor enable them to overcome barriers to accessing high return labor activities. These reallocations of labor supply across work activities lead to increases in their consumption, and a diversification of their asset base, especially through accessing land, and this process sets them on a sustained trajectory out of poverty.

By the end of our study in 2014, the program had reached $360,000$ households in Bangladesh containing 1.2 million individuals, and it has subsequently been piloted in other countries (Banerjee *et al.* 2015a). We compare our results for Bangladesh to those from six pilot studies in Ethiopia, Ghana, Honduras, India, Pakistan and Peru (Banerjee *et al.* 2015a). Across ten dimensions covering consumption, food security, assets, financial inclusion, labor supply, income, physical health, mental health, political awareness and women's empowerment, we find the three year results for these pilot studies are strikingly similar to our four year results. The fact that the program has positive effects across such a wide range of outcomes increases confidence that it has a profound effect on the lives of ultra-poor women. The comparison of our findings to those of other pilots suggests that specifically promoting occupational change is effective in different contexts. This lends support to the argument that the program may be able to be scaled-up in different contexts with different implementing partners to achieve sizable and sustainable improvements in outcomes for the poorest.

The paper is organized as follows. Section II describes key features of rural labor markets underlying our analysis. Section III describes the TUP intervention, our data and research design.

Section IV documents treatment effects on the ultra-poor. Section V looks across the wealth distribution to provide estimates of indirect treatment effects on ineligible households and the extent to which the ultra-poor close the gap with the near-poor. Section VI presents a cost-benefit analysis and estimates internal rates of return. Section VII examines the trajectories of beneficiaries after seven years. Section VIII concludes by discussing the broader implications of our study.

II. Labor Markets and Poverty at Baseline

II.A Poverty and Wealth Classes

We study labor markets in 1,309 villages in Bangladesh's 13 poorest districts. These districts were chosen by BRAC to implement the TUP program based on food security maps of the World Food Program. Our sample is drawn from two randomly selected sub-districts in each district, containing 40 BRAC branches that serve the 1,309 villages where the evaluation takes place.[5]

To construct our sample we first conducted a census of the 99,775 households in the 1,309 villages. To draw a sample for the baseline survey, we combine this data with information on household wealth, derived from a participatory wealth ranking organized by BRAC in each village. This exercise places all households into one of several wealth bins corresponding to the poor, the middle class, and the upper class. Pre-randomization, BRAC officers use inclusion and exclusion criteria to further subdivide the poorer households into the ultra-poor, who are eligible for the TUP program, and the near-poor who are not. The four wealth classes account for 6%, 22%, 59% and 14% of the village populations, respectively (Table I). We survey almost all ultra-poor and near-poor households, and a 10% random sample of households from higher wealth classes, at baseline in 2007 and then at follow-ups in 2009, 2011 and 2014. Overall the sample covers over 21,000 households in 1,309 villages, of which over 6,700 are ultra-poor. Our research design allows us to study the program's: (i) intent-to-treat effect on the ultra-poor, where the number of ultra-poor households that we track allows us to further estimate quantile treatment effects to

[5]There is a concentration of study sites in the Northern part of the country. This is because this is the poorest and most vulnerable region, often referred to as the monga or famine region (Bryan *et al.* 2014). Our evaluation is representative of the areas in which the nationwide TUP program was scaled-up in after 2007.

shed light on heterogeneous impacts of the program among the ultra-poor; (ii) general equilibrium and distributional impacts on near-poor, middle class and upper class households.

The top two panels of Table I confirm that the participatory ranking exercise is successful in identifying the poorest households: 53% of the households identified as ultra-poor are below the $1.25 a day poverty line, while the corresponding figures for the near-poor, middle and upper classes are 49%, 37% and 12%. Due to BRAC's targeting strategy, the primary woman is the sole earner in 41% of the ultra-poor households, while this only occurs in 25%, 14% and 12% of near-poor, middle and upper class households. Illiteracy is also much higher for ultra-poor women: a staggering 93% of them are illiterate compared to 83%, 74% and 49% in the other three wealth classes. These data confirm that the ultra-poor are severely disadvantaged relative to their wealthier counterparts in the same village. They also confirm that these village economies have a significant fraction of middle and upper class households lying below the extreme poverty line.

Looking across household assets, savings, livestock, land and business assets the distinguishing feature of the ultra-poor is that they are largely assetless. As we look across the columns of Table I all these variables are larger for wealthier households.

The value of cows owned by the ultra-poor is only 2.2% of the value owned by the upper classes and the corresponding figure for goats is 11.1%. This gap in the value of livestock is driven both by the ultra-poor being much less likely to own livestock (particularly cows) and then conditional on owning livestock being more likely to own goats (the average value of which is close to USD 54 in PPP terms) rather than cows (the average value of which is USD 542). As households get richer they focus on accumulating cows not goats with the former accounting for 96% of the value of livestock owned by upper class households. Therefore, as the comparison of cow and goat values in Table I shows, cows are the key livestock asset in these village economies. Table I also shows that rental markets do not equalize access to productive assets: only 7% of the poor in our sample rent in cows from other households. This is likely because of various transaction costs associated with renting out livestock to others, which have been shown to be relevant in rural labor markets (Shaban 1987, Foster and Rosenzweig 1994).[6]

[6]Even though wealthier households can in principle gain by renting livestock to the poor to take advantage of their lower labor costs, the transaction costs from doing so are high for at least three reasons: (i) the ultra poor lack

The final panel of Table I shows that the poor are much less likely to own land than wealthier households. Only 7% of ultra-poor households own land at baseline compared to 11%, 49% and 91% for near-poor, middle class and upper class households. In addition only a small fraction of the ultra-poor, 6%, rent land for cultivation. The majority of ultra-poor households are therefore landless and the value of land they own is tiny compared to middle class and upper class households. Land is the asset that most clearly differentiates rich from poor households in these villages.

What is also clear from Table I is that inequality in asset holdings across the village wealth distribution is much more marked than inequality in consumption. Average consumption expenditure per adult equivalent for ultra-poor households is 51% of that for upper class households. The corresponding figures for household assets, savings, business assets, value of cows, value of goats and value of land owned are 2.2%, 1.6%, 1.5%, 2.2%, 11% and 0.5%. The upper classes in the villages are distinguished mainly by owning more assets, particularly agricultural land. The ultra-poor, in contrast, have negligible asset holdings.

These characteristics of ultra-poor women combined with the fact that they have a median age of 40 and an average of one dependent child below the age of 10 imply that they are likely to be captive in these village labor markets. Migration to other labor markets in towns and cities is unlikely to be a possibility for the majority of ultra-poor women. In common with many ultra-poor women around the world they have to choose from the work activities on offer within the villages where they currently reside.[7]

II.B Labor Markets

Our survey collects information on all labor activities, for each household member, during the previous year. For each activity, we ask whether the individual was self-employed or hired by a third party as a wage laborer, the number of hours worked per day, the number of days

experience of livestock rearing: for centuries they have been landless and engaged in casual wage labor activities; (ii) the quality of labor inputs in livestock rearing are critical: there can be large variations in the productivity of livestock due to differences in feeding, veterinary and other practices; (iii) the economic opportunities of wealthier households means they face high opportunity costs of supervising, or training, other households when rearing livestock. More generally, Shaban (1987) and Foster and Rosenzweig (1994) provide evidence of the importance of moral hazard in labor contracts in rural India.

[7]Later we present experimental evidence that the program did not lead to differential attrition in treatment versus control villages which is consistent with this hypothesis. Cultural barriers also imply that migration, and in particular seasonal migration, is typically practised by males in Bangladesh (Bryan *et al.* 2014).

worked per year, wage rates and total earnings. We collect data related to the entire year because employment in casual wage jobs, especially those in agriculture, is irregular so a that a shorter time frame (days, weeks) is likely to severely mis-measure aggregate hours devoted to these activities. As the program targets the primary woman in ultra-poor households, defined as the head's spouse or the female head, we focus the analysis on women's labor market activities.[8]

Figure IA begins to describe the working lives of women in rural Bangladesh. It identifies the main labor activities in these villages by showing the share of women's work hours devoted to various work activities in each of the 40 BRAC branches our sample covers. The figure reveals that the set of labor activities that women engage in is extremely limited. Around 80% of women's labor hours are devoted to three activities: casual jobs in agriculture, casual jobs as domestic maids and livestock rearing. The first two are activities where unskilled labor is the only input and where women are hired daily without any guarantee of future employment.[9] For the third, women are self-employed, working with cows and goats to generate income through the sale of milk, meat, manure and young calves. The key difference between these two sets of activities is that the latter requires a capital input. It is also likely that livestock rearing requires higher levels of skills.[10] Figure IA shows that while livestock rearing is present in all labor markets, either agricultural or maid labor tends to dominate in a particular location. Hence in most villages within a given BRAC branch, women effectively choose between two labor activities – agricultural/maid labor and livestock rearing.[11]

Figure IB presents hours of work broken out by wealth class and activity to investigate whether

[8]Bardhan (1984b) and Foster and Rosenzweig (1996) document a marked differentiation in agricultural tasks by gender, which is also observed in our setting.

[9]In our data 99% (96%) of women working in agricultural wage labor (as maids) report being hired and paid daily through spot contracts. This is also what Kaur (2015) observes in India using NSS data. We do not therefore observe coexistence of temporary and permanent wage labor contracts (Eswaran and Kotwal 1985).

[10]Expertise is needed to (i) give beef cows, dairy cows and goats the right diets, (ii) be able to detect diseases and know when to contact the vet; (iii) know about vaccines and when they need to be given; (iv) be able to work with artificial insemination services (for cows); (v) be able to construct livestock sheds and keep them clean.

[11]Due to the geographical separation of casual wage labor activities described in Figure IA, agricultural work and maid work are rarely combined to make a full time job. Only 10% of women who report any wage activity are engaged in both casual agricultural labor and domestic maid work. We also note that 43% of poor women generate small amounts of income from poultry: however, the returns from such activities are far lower than even for casual wage labor. Following the earlier literature that has argued for buffer stock motivations of animal ownership (Rosenzweig and Wolpin 1993), we consider poultry holdings as a form of illiquid savings rather than representing a key choice over labor market activities.

there is a correlation between labor market activities and poverty. The figure demonstrates that there is a pronounced shift towards livestock rearing as we move up the wealth distribution. Ultra-poor and near-poor women engage predominantly in casual wage labor, although ultra-poor women are distinguished from near-poor women by relying *almost exclusively* on unskilled casual labor which requires no capital input and where they rely on others to employ them, primarily as agricultural laborers or domestic maids. In contrast, women from middle and upper class households are predominantly engaged in livestock rearing. Across all four wealth classes these three activities account for 80% of hours worked.[12]

Figure IC graphs the hourly returns for the three main work activities averaged over all individuals with non-missing earnings and positive hours in each of the three activities. We compute simple averages at the BRAC branch level. Hourly returns for casual jobs are equal to the hourly wage. To compute average hourly earnings for livestock rearing we divide yearly profits (revenues minus input costs) by total hours devoted to livestock rearing over the year. Two things are apparent from this plot. The first is that the average returns for those engaged in livestock rearing are higher than those for casual wage labor in nearly all rural labor markets in our sample. Table A.I shows that, at the village level, hourly earnings in livestock rearing are USD 0.72 per hour, more than double the hourly earnings for agricultural wage labor (USD 0.34 per hour) and maid work (USD 0.27 per hour). The choice over labor activities however depends on the *marginal* returns to labor in each. For competitive casual wage labor markets, that are governed by spot contracts without any future employment guarantee, the hourly wage closely matches the MP_L. For capital-intensive activities such as livestock rearing, measuring the MP_L requires knowing the production function for how capital and labor are combined. Assuming a Cobb-Douglas technology, MP_L is proportional to AP_L, with the constant of proportionality being labor's share of income. Given the measured returns across activities, we note that for the average branch, the MP_L in livestock rearing is larger than the MP_L in agricultural (maid) work as long as the labor

[12]The remaining 20% of hours is distributed across several other activities which typically account for less than 1% of hours each (where work on the household's own land is counted as own cultivation not agricultural labor). The activities that account for more than 1% for the ultra-poor are: begging (6%), tailoring (4%), casual day labor outside agriculture (4%), land cultivation (1%). For the near-poor they are: begging (3%), tailor (3%), casual day labor outside agriculture (3%), land cultivation (4%). For the middle classes they are: tailoring (3%), land cultivation (4%). For the upper classes they are: tailoring (1%), teacher (1%), land cultivation (5%).

share is larger than .48 (.37). Macro-wide estimates from developing countries typically lie in the range of .65-.80 (Gollin 2002).[13]

The second observation from Figure IC is that returns to casual wage labor are uniform across space whereas returns to livestock rearing vary strongly across space. The uniformity of returns to casual labor across geography reflects the fact that there is an abundant supply of low skilled women willing to work in these work activities and wages offered in village spot markets tend to fall within narrow bands (Kaur 2015). In contrast, returns to livestock rearing vary according to location-specific features such as linkages to urban markets and trade networks (Donaldson 2015).

Figure I exposes the puzzle at the heart of our study – why do the poor not allocate their labor to the activity with the highest return? One possibility is that the observed cross-sectional returns to activities might not represent the returns available to the poor if they engaged in them. The differences could be due to differences in innate ability correlated with poverty or to increasing returns to scale. To explore the latter, Figure ID graphs a local polynomial regression of hourly returns on the value of livestock owned by households. While the estimated returns need to be cautiously interpreted given livestock holdings are endogenous, across the whole distribution the returns to livestock rearing are higher than for casual wage labor activities (that themselves do not vary with livestock ownership as expected). The vertical bars on Figure ID indicate the average value of livestock owned by the ultra-poor pre- and post- the TUP program intervention we evaluate. Over this range, the returns to livestock rearing are higher than for both forms of casual wage labor, and these returns are also clearly rising with livestock value, indicating there might be increasing returns to livestock rearing.[14] Evaluating the TUP program allows us to assess whether differences in returns can be explained by differences in innate ability, or reflect multiple barriers that the poor face in accessing labor activities that they are otherwise able to engage in.

Besides having different hourly returns and capital requirements, the two types of work activities also exhibit a different distribution of hours worked across days of the year. Table A.I shows that the average woman engaged in casual agricultural labor works in this activity for only 127

[13] A body of field experiments examining the returns to capital in developing country contexts find that these returns are higher than the returns to labor (de Mel *et al.* 2008, Blattman *et al.* 2014).

[14] That there are increasing returns to livestock rearing is in line with evidence from other settings in rural South Asia (Anagol *et al.* 2014, Attanasio and Augsburg 2014).

days of the year; engagement in domestic maid work is for only 167 days per year. In contrast, women engaged in livestock rearing work almost every day of the year. However, conditional on working, women employed in casual wage activities work many more hours per day: 7.6 daily hours for casual agricultural work, 7.0 for maid work, versus 1.8 daily hours for livestock rearing. [15]

Table II shows the implications of low demand for casual labor on the distribution of hours worked across wealth classes: over the course of a year, poor women bunch their work into *fewer* days of the year than wealthier women, but work *more* hours in the year overall. This bunching is driven by the concentration of poor women's labor supply into casual wage activities that are only available for less than half the year. In contrast, wealthier women specialize in livestock rearing, enabling them to smooth their labor supply over the year.

Taken together, the evidence suggests a clear correlation between poverty and labor market activities with poor women allocating most of their labor to low-return, irregular, casual jobs and richer women specializing in high-return, regular, livestock rearing. The key question is whether poor women would be better off engaging in the same activities as their wealthier counterparts but face barriers in accessing capital or skills that keep them in poverty. The beneficiaries' response to the TUP program, which simultaneously relaxes these capital and skills barriers, sheds light on this question. If ultra-poor women prefer employment in casual jobs they will sell (or rent out) the asset without changing their labor market choices. If they prefer livestock rearing but face asset and/or skills related barriers to engaging in such activities, they will retain the asset and work with it once barriers are removed.

[15] Absent large fixed costs of daily labor supply or concave daily costs of work effort, women should prefer to smooth their labor supply. The observed bunching of labor supply for casual wage activities into fewer days of the year is indicative of constrained or low aggregate demand for both forms of casual wage labor. This is not surprising for agricultural wage labor because of inherent seasonality in labor demand including the well documented pre-harvest lean season in the agricultural cycle in Bangladesh, during which the demand for labor is almost non-existent (Khandker and Mahmud 2012, Bryan *et al.* 2014).

III. Intervention and Research Design

III.A The Intervention: TUP

The TUP program is designed and implemented by BRAC to reach the very poorest women in rural Bangladesh who are not targeted by other forms of assistance. Pre-randomization, eligible households are selected by BRAC officers from the list of poor households produced by a village participatory wealth ranking.[16] To qualify for the program, the household needs to have an able adult woman present, not to be borrowing from a microfinance organization or receiving transfers from government anti-poverty programs, and meet three out of five inclusion criteria.[17] Eligibility is not conditional on participating in other BRAC activities.

The program targets the leading woman in eligible ultra-poor households. Women are presented with a menu of assets, each of which can be used in an income generating activity. These assets include livestock and those relevant for small-scale retail operations, tree nurseries and vegetable growing. Each asset is offered with a package of complementary training and support.

Of those households identified as ultra-poor at the outset, 86% eventually receive an asset. The other 14% either cease to meet the eligibility criteria when transfers are implemented, or choose not to take-up the program.[18] All the offered asset bundles are similarly valued at USD 560 in PPP terms. The scale of asset transfers corresponds to a near doubling of baseline wealth for the ultra-poor, values that are far higher than households could borrow through informal credit markets. All eligible women chose one of the six available livestock asset bundles from the asset menu and 91% of them choose an asset bundle containing at least one cow. Before the intervention, the value of livestock owned by the 47% of ultra-poor households with either a cow or a goat at baseline is just USD 49.7.

[16] For the participatory wealth ranking exercise, villages are asked to rank all households into wealth bins and reach a consensus on the wealth class of each household. People who own sufficient amounts of land, have a salaried job, live in a tin or paddy sheafhouse, own cows, goats or other livestock or own a power tiller, rice mill, etc. are considered wealthy and people who are landless and who own nothing outside their homestead, work as casual laborers, small traders or beg, do not own any livestock or assets and live in straw houses are considered to be poor (BRAC 2004). Alatas *et al.* (2012) show that, compared to proxy means tests, participatory methods result in higher satisfaction and greater legitimacy.

[17] The eligibility criteria are (i) total land owned including homestead land does not exceed 10 decimals; (ii) there is no adult male income earner in the household; (iii) adult women in the household work outside the homestead; (iv) school-aged children work; and (v) the household has no productive assets.

[18] It is likely most did not receive assets because they had become ineligible, not because of take-up refusal. For example, compared to those receiving assets, those who did not were twice as wealthy and more likely to own land.

Assets are typically transferred one month after choices are first made. Eligibles are encouraged by BRAC to retain the transferred asset for two years, after which they can liquidate it. Thus, whether the livestock asset is retained or liquidated by the time of our four-year follow up is itself an outcome of interest that ultimately determines whether the program impacts the long-run allocation of time across work activities, or just contributes to a potentially short run increase in household welfare.

The associated support and training package is also valued at around USD 560 per beneficiary. This component comprises initial classroom training at BRAC regional headquarters, followed by regular assistance through home visits. A livestock specialist visits eligibles every one to two months for the first year of the program, and BRAC program officers provide weekly visits for two years post transfer. As the ultra-poor have limited experience with large livestock (particularly cows), this assistance is designed to cover the life cycle of livestock. Ultimately, this training component is intended to mitigate earnings risks from working with livestock and to increase the overall return to livestock rearing.[19]

The program also provides a subsistence allowance to eligible women for the first 40 weeks after the asset transfer to help smooth any short-run earnings fluctuation due to adjustments across work activities. This allowance ends 15 months before our first follow-up and is therefore not part of the earnings measures reported. To empower ultra-poor women along non-economic dimensions the program also provides health support and training on legal, social and political rights. The program also sets up committees made up of village elites which offer support to program recipients and deal with any conflicts and problems they encounter. Finally, the program encourages saving with BRAC during the program and borrowing from BRAC microfinance at the end of the program, but neither is a pre-condition to obtain the asset-training bundle.

The program thus represents a bundle of asset and skills transfers. Given the economic circumstances and life experiences of the ultra-poor, there are good theoretical reasons why these components need to be offered together. The strong focus on continual training and support over

[19]Training is designed to help women maintain the animals' health, maximize the animals' productivity through best practices relating to feed and water, learn how to best inseminate animals to produce offspring and milk, rear calves, and to bring produce to market. The training is sufficiently long-lasting to enable women to learn how to rear livestock through their calving cycle and across seasons.

a two year period is one way in which the TUP program differs from previous asset transfer programs (Dreze 1990; Ashley *et al.* 1999). In short, the program can potentially change a number of dimensions of poor women's lives. Transferring assets has a large impact on their wealth and the program provides key asset and skill inputs needed to take on labor activities engaged in by richer women. Continued support during the period of learning can further improve their chances of being successful in taking on these activities. It may also make women more assured and confident that they can take on work activities other than casual labor (including those that are not encouraged by the program) and may change cultural attitudes toward these women. We evaluate the full impacts of the bundled version of the program, and thus do not aim to identify specific constraints on occupational change that the program may be operating through.

III.B Research Design

The TUP program evaluation sample comes from among the 13 poorest districts in rural Bangladesh, as described earlier. In most cases we randomly selected two sub-districts (upazilas) from each district and within each subdistrict we randomly assigned one BRAC branch office to be treated and one to be held as a control.[20] All villages within an 8 kilometer radius of a treated BRAC branch receive the program in 2007 while villages in control branches receive it after 2011. We randomize at the branch rather than village level to mitigate spillovers between treatment and control villages either through markets or through program officers. We are evaluating a scaled version of the TUP program: by 2014, this had reached over $360,000$ households containing 1.2 million individuals.[21]

For the purpose of the evaluation, the participatory wealth ranking is conducted in both treatment and control areas and BRAC officers identify eligible ultra-poor women in identical ways

[20]The average subdistrict has an area of approximately 250 square kilometers (97 square miles) and constitutes the lowest level of regional division within Bangladesh with administrative power and elected members. For each district located in the poorer Northern region we randomly select two subdistricts, and for each district located in the rest of the country we randomly select one subdistrict, restricting the draw to subdistricts containing more than one BRAC branch office. For the one district (Kishoreganj) that did not have subdistricts with more than one BRAC branch office, we randomly choose one treatment and one control branch without stratifying by subdistrict.

[21]A variant of the program where the poor have to repay the cost of the asset transferred to BRAC had reached an additional 1.1 million households containing 3.6 million members by 2014 (BRAC 2015).The TUP program started in 2002 and there was a second wave in 2004. The scale of these waves was smaller than the wave that started in 2007 and these were used, in part, to inform the design of the scale-up that took place in 2007. The 2002-2006 period therefore involved significant piloting and experimentation (Hossain and Matin 2004).

in both areas. To avoid anticipation effects, information about the availability of the program and eligibility status is not made public until program operations begin in a given area (in mid 2007 in treatment areas, after 2011 in control areas) and the participatory wealth ranking is presented as a part of regular BRAC activities rather than associated with a specific program.

Table A.II provides evidence on whether the characteristics of the ultra-poor are balanced between treatment and control villages. For each outcome considered, we report means and standard deviations in treatment and control villages (Columns 1 and 2), the p-value on a test of equality of means (Column 3) and the normalized difference of means (Column 4). For each family of outcomes we also report the average standardized difference following Kling *et al.* (2007). The samples are well balanced on outcomes: only one out of 22 tests yields a p-value below .05, and we cannot reject the null hypothesis of equal means for any of the average standardized differences. Furthermore, Column 4 shows that all normalized differences are smaller than 1/6th of the combined sample variation, suggesting linear regression methods are unlikely to be sensitive to specification changes (Imbens and Wooldridge 2009).

Over the four years from baseline to endline, 15% of ultra-poor households attrit, a rate comparable to other asset transfer program evaluations (Banerjee *et al.* 2015a). Table A.III estimates the probability of not attriting as a function of treatment status and baseline work activities. This shows: (i) attrition rates do not differ between treatment and control villages; (ii) women engaged in livestock rearing are more likely to be surveyed in all three waves; (iii) crucially, there is no differential attrition by baseline work activities between treatment and control individuals: the coefficients on interaction terms between treatment status and activity choice at baseline are all precisely estimated and close to zero. To ease comparability our working sample is based on those households that are tracked in both follow-ups, covering $6,732$ ultra-poor households.

IV. Treatment Effects on the Ultra-Poor

We evaluate the impacts of the TUP program on individual and household level outcomes exploiting the experimental variation caused by the random assignment of villages to treatment or

control. We estimate the following difference-in-difference specification:

$$y_{idt} = \alpha + \sum_{t=1}^{2} \beta_t \left(W_t \times T_i\right) + \gamma T_i + \sum_{t=1}^{2} \delta_t W_t + \eta_d + \varepsilon_{idt}, \tag{1}$$

where y_{idt} is the outcome of interest for individual/household i in subdistrict d at time t, where time periods refer to the 2007 baseline ($t = 0$), 2009 midline ($t = 1$) and 2011 endline ($t = 2$). W_t are survey wave indicators. $T_i = 1$ if individual i lives in a treated community and 0 otherwise. η_d are subdistrict fixed effects and are included to improve efficiency because the randomization is stratified by subdistrict. The error term ε_{idt} is clustered by BRAC branch, the unit of randomization. All monetary values are deflated to 2007 prices using the Bangladesh Bank's rural CPI estimates and converted into USD PPP.

β_t identifies the intent-to-treat impact of the program on ultra-poor individual/household i under the twin identifying assumption of random assignment and no spillovers between treatment and control villages. This estimate compares changes in outcomes among ultra-poor residing in treated villages pre- and post- intervention, to changes among counterfactual ultra-poor in control villages in the same subdistrict. As discussed earlier, the ultra-poor are identified in identical ways in treatment and control locations pre-randomization. To benchmark the magnitude of the effects we report the four year effects in percentage of the control mean in the same period throughout. Specification (1) controls for time-varying factors common to ultra-poor in treatment and control villages, and for all time-invariant heterogeneity within subdistrict. Tables A.VA and A.VB probe robustness to using an ANCOVA specification both pooling the survey waves and running each separately.[22] Table A.VI probes robustness to different inference methods that correct for the small number of clusters: the Young (2016) degrees of freedom correction and the Cameron *et al.* (2008) wild-bootstrap method. All results are quantitatively and qualitatively robust to both sets of changes.

The subsections below test the impact of the program at each step of the causal chain that links

[22]Table A.V.A reports the estimates of $y_{id} = \alpha + \beta T_i + \mu y_i^0 + \eta_d + \varepsilon_{id}$ run separately on the cross-section of eligible households in 2009 and 2011, where y_i^0 is the baseline (2007) value of y_i and all other variables are as defined above. Table A.V.B reports the estimates of $y_{idt} = \sum_{t=1}^{2} \beta_t \left(W_t \times T_i\right) + \sum_{t=1}^{2} \nu_t \left(W_t \times y_i^0\right) + \sum_{t=1}^{2} \delta_t W_t + \eta_d + \varepsilon_{idt}$
where $t \in [1, 2]$ (1=2009, 2=2011), y_i^0 is the baseline (2007) value of y_i and all other variables are as defined above.

choices over labor activities to earnings, consumption, savings and investment. The comparison between two and four year effects reveals whether the effects become stronger over time, which is important for understanding whether the program sets the ultra-poor on a sustainable trajectory out of poverty.

IV.A Labor Supply and Earnings

Table III shows program impacts on labor supply (Panel A) and earnings (Panel B) for the three main labor activities for women in Bangladeshi villages. Column 1 of Panel A shows that the program succeeds in its aim to induce ultra-poor women to take up livestock rearing: four years after baseline ultra-poor women allocate 415 more hours to livestock rearing per annum, a 217% increase relative to controls in the same time period. This corresponds to ultra-poor women working 172 days in this activity per annum representing an increase of 181% relative to controls (Column 2). Comparing two and four-year impacts we note that the change in hours devoted to livestock rearing is immediate, in line with the fact that beneficiaries move into livestock rearing as soon as they receive the assets. The increase represents 1.14 more hours per day which matches well with the time allocation to this activity observed at baseline (Table II).

In short, livestock rearing has become a central element in the working lives of ultra-poor women. The findings further indicate that beneficiaries continue to own livestock instead of liquidating it for consumption, despite the fact that the value of the transfer is equal to one year's worth of consumption for the average adult. They also indicate that beneficiaries are able to maintain the asset once assistance is removed as the effects are sustained after the two years mark.

Columns 3 to 6 show evidence that ultra-poor women start pulling out of casual wage labor activities. While the change in hours devoted to livestock rearing is immediate, the effect on casual labor hours is gradual. The reduction in agricultural labor (46 hours, 17% relative to controls) is not precisely estimated while the fall in maid hours increases in magnitude between two and four years and is significant only after four years (117 hours, 26% relative to controls). This is consistent with the fact that the wage rate for agricultural labor is higher than that for maid work (Figure IC and ID and Table A.I). Overall, ultra-poor women are dropping some of the least attractive casual labor hours but still hold on to the majority even as they significantly increase

livestock hours.[23]

Aggregating across labor activities, Columns 7 and 8 show that four years post-intervention total hours worked increases by 206 (17%) and days worked per year increase by 61 (22% more than in control). This suggests that the poor had idle labor capacity at baseline which they were able to successfully combine with the bundled asset-skills transfer as a result of the program. This improvement in the regularity of employment is a key labor market impact of the program. At baseline ultra-poor women, like many of the poorest women in rural parts of the developing world, were captive in occupations at the bottom of the employment ladder using labor, their only endowment. Significantly, demand for this labor was highly irregular. The opportunity to engage in livestock rearing that the program provides allows the women to fill in the days when they had previously been idle. The shift away in hours devoted to casual wage labor is more gradual. While economically significant, the magnitude of the reduction in hours devoted to casual wage labor implies that four years after the program ultra-poor women still engage in these activities so that differences in labor activities relative to middle and upper class women remain.

Panel B of Table III then focuses on earnings from work activities. In Column 9 we see that earnings from livestock rearing increase from USD 80 to USD 115 between years two and four post-intervention. The four year effect is significantly larger than the two year effect despite a modest drop in labor supply (Column 1) indicating that ultra-poor women are becoming more productive in this activity over time.

In Columns 10 and 12 we see that declines in supply of agricultural labor and maid services are associated with significant increases in wage rates in those activities after four years (by 12% and 21% respectively). These wage effects are insightful as they rule out that the aggregate supply of casual labor by ultra-poor women is perfectly elastic, as in Lewis (1954) and Fei and Ranis (1964). They are consistent with an upward sloping supply curve because as ultra-poor women

[23]The small scale of livestock rearing that ultra-poor women operate at, corresponding to keeping a couple of cows or a cow and several goats, may constrain both the labor input and returns to this activity, making continued engagement in casual wage labor necessary. In other settings, there is also evidence that even small-scale farmers resort to these occupations because they are unable to cover short-term consumption needs with savings or credit (Fink *et al.* 2014). The slightly smaller daily time allocation of ultra-poor women to livestock rearing relative to other women (Table II shows that pre-intervention, women allocated 1.8 hours per day to livestock rearing) might also be due to them operating at a smaller scale than middle and upper class women.

remove their labor from village labor markets for these activities, prices need to rise to clear the market (Rosenzweig 1978; Rosenzweig 1988; Rose 2001; Jayachandran 2006; Goldberg 2010; Kaur 2015).[24] The removal of ultra-poor labor from these activities and the consequent rise in wages therefore may have positive general equilibrium effects for the wages received by other women in other wealth classes who continue to work in these activities. We examine this issue in further detail in Section V.

Increased wages will of course also benefit the majority of ultra-poor women who continue to devote some hours to agricultural labor and maid services. For agricultural labor we see that the modest reduction in labor supply and the modest increase in wages cancel out so that there is no significant impact on earnings from this activity (Column 11). In Column 13 we see, however, that for maid labor the reduction in labor supply dominates the increase in wages and total earnings from maid labor fall by 22% after four years. This equates to a statistically significant loss of USD 25 from casual wage labor per annum after four years (Column 12). This, however, is modest relative to the gain of USD 115 from livestock rearing over the same period (Column 9).

Aggregating across activities, the reallocation of time from casual labor to a more-than-offsetting increase in livestock rearing leads to a significant increase in net annual earnings (earnings net of input costs of livestock rearing) of 21% relative to controls in the same time period (Column 14). A key impact of the program therefore is to make earnings from livestock a significant additional source of income for ultra-poor households. In short, the program allows women to both raise their net earnings, and to smooth their labor supply and earnings stream over the year. Taken together, these imply that the poorest women in these villages are able and willing to take on the same labor activities as their wealthier counterparts, suggesting that the program lifted barriers they must have faced to entering such work activities at baseline.[25]

It is possible that the program may affect the labor market choices of household members

[24]We can rule out that the wage increases are due to selection, namely to lower paid individuals dropping out of these activities. Indeed the estimated effect on wages is the same in the balanced sample of individuals that engage in these activities in all three waves of the survey (see Section V). This is consistent with these being low-skilled activities that pay similar wages across locations and across the wealth distribution as shown in Figure IC.

[25]The stability of the impact on net earnings at two and four years post-intervention suggests the ultra poor are not necessarily being exposed to more intertemporal risk in livestock rearing, even though 2009 was a low rainfall year in many parts of rural Bangladesh. This is of note given the findings in Attanasio and Augsburg (2014).

other than the targeted female and these must be taken into account to evaluate the effects on household welfare. In Table A.IV we show that, while all household members devote some more hours to livestock rearing, the effect is about one tenth of the size of that on ultra-poor women and does not crowd out other work activities or schooling. This allays the potential concern that the program increases women's earnings at the expense of the earnings of other family members, or children's education. Another possible channel through which the program might affect the labor market choices of other household members is by inducing some of them to migrate. We find no evidence that this occurs in our setting, likely because 47% of ultra-poor households have no adult members other than the main woman and her husband (if present) and 35% have just one, and because females do not typically engage in seasonal migration in Bangladesh for cultural reasons (Bryan *et al.* 2014). Given these null impacts on migration, migrant remittances are likely to play a minor role.[26]

IV.B Consumption Expenditures, Savings and Credit

Table IV analyzes the consequences of ultra-poor women reallocating their labor supply across activities, for the welfare of their households. Column 1 shows that relative to the controls, the share of households below the USD 1.25 poverty line drops by 8.4pp, or 14% after four years. In Column 2 we see that consumption expenditure per adult equivalent is 11% higher in treatment relative to control households after four years.[27]

Program effects are likely to be heterogeneous depending on unobservables such as the innate ability for livestock rearing and the underlying constraints faced. We test for heterogeneity by estimating the following quantile treatment effects (QTE) specification:

$$\text{Quant}_\tau(\Delta y_{id}) = \beta^i_\tau T_i + \vartheta_\tau \eta_d, \tag{2}$$

[26] On the migration channel we find that: (i) household size actually increases, rather than decreases, for treated households; (ii) this is partly driven by more adults remaining in the household; (iii) there is no significant change in out-migration.

[27] The consumption expenditure items covered are: food (both purchased and produced, accounting for the number of people taking meals in the household), fuel, cosmetics, entertainment, transportation, utilities, clothing, footwear, utensils, textiles, dowries, education, charity and legal expenses. Further decomposition of consumption expenditures into food and non-food reveals the effect is driven mostly by the latter but nutrition improves as the consumption of milk and meat increases.

where Δy_{id} corresponds to the difference between the four year and baseline values of outcome y for individual i in subdistrict d.

Figure IIA shows that treatment effects on consumption are non-negative at each centile, but they are significantly larger at higher centiles with the effect on the 5th centile being roughly one tenth of that at the 95th centile. Thus even within the narrow group of ultra-poor households, there is significant variation in the effect of treatment. Uncovering the root causes of these differences among the ultra-poor represents a key priority for future research.

In Column 3 of Table IV we see that, after four years, household assets (which include jewelry, sarees, radios, televisions, cell phones, bicycles and furniture) increase in value by 57% relative to control. The increase in the value of household assets is significantly larger after four years relative to two years. In Figure IIB we see that, although household asset effects are positive and significant for all centiles, asset accumulation is much more pronounced in the upper centiles.[28]

Columns 4 to 6 of Table IV analyze the impact of the program on financial assets. In Column 4 we see that household cash savings held with microfinance organizations, banks and saving guards increase significantly after two and four years. Given that ultra-poor household savings are negligible in the absence of treatment the increase in savings of USD 53 after four years is highly significant and represents a fourfold increase relative to controls. Though it remains a choice variable, households are encouraged to open and manage savings accounts during the first two years. The fact that the savings effect remains significant after four years indicates that households are choosing to save more two years after there is any encouragement to do so. Figure IIC shows that as with consumption expenditure and household assets, the program impact on savings is highly heterogeneous.

In Column 5 of Table IV we see that, after four years, households are 11pp more likely to receive loans which represents a 50% increase relative to controls. The program is thus enabling ultra-poor households to obtain access to credit two years after they are encouraged to do so as part of the program. On the other side of financial intermediation, at baseline only 1% of ultra-poor households give loans (Table I). Column 6 shows that they are 5pp more likely to do so after

[28]This is consistent with the pattern of effects on consumption although we cannot say whether those who experience the largest increases in consumption are the same as those who experience large increases in assets.

four years relative to controls.

The savings, borrowing and lending results all point to improved financial inclusion for ultra-poor households. Moreover, the enhanced lending by the ultra-poor to others is a key indicator that their financial position in the village has improved – a proportion of ultra-poor households now have surplus capital that they lend to others. This creates another channel through which the program can affect other households in the village, discussed further in Section V.

IV.C Productive Assets

Table V examines the program's impacts on the accumulation of productive assets, as this is central to whether the one-off asset and skills transfers lead to sustainable gains in welfare. Columns 1 and 2 analyze the effect on the value of assets transferred by the program, that is cows and goats. The first thing to note is that ultra-poor women mainly choose cows in their asset transfer package: the mean value of goats transferred is only 8.6% of the value of cows transferred. In Column 1 we see that, after four years, the value of cows owned by ultra-poor households has increased by 122% (net of the transfer value) relative to controls. At year four the value of cows is 16% larger than the value of the asset transfer: the value of cows has increased from USD 485 to USD 540 between years two and four where the original value of the cows transferred was USD 464. This signals that the majority of ultra-poor households have been able to grow the value of this productive asset via the enlargement of herds. Consistent with this, we can reject the null that the estimated effect on the number of cows (not shown) is equal to the number transferred.[29]

Column 2 shows that the value of goats held by ultra-poor households (net of the transfer value) actually declines after four years suggesting that some animals have been liquidated or have died. However, after four years, the cow value effect is 26 times the goat value effect so, overall, ultra-poor households experience a large and significant increase in the value of livestock held as a result of the program.

Land is the key asset in the densely populated rural areas of Bangladesh which are dominated by agriculture and ultra-poor households have very limited access to cultivable land (see Table I). In Columns 3-5 we see that the program impacts the access ultra-poor households have to land,

[29]Set against a backdrop where attempts to transfer cattle to the poor have a highly checkered history this is a significant finding (Dreze 1990; Ashley *et al.* 1999).

even though this is not an explicit aim of the program. Ultra-poor households become 11pp more likely to rent land after four years, representing a 139% increase relative to controls. In Column 4 we see that ultra-poor households are 2.6pp more likely to own land after four years representing a 45% increase and the value of land owned increases significantly by an average of USD 327 by four years post-intervention (Column 5). This accumulation of land takes place between years two and four with the four year effect being significantly higher than the two year effect. This indicates, importantly, that ultra-poor households are using part of the surpluses generated by their reallocation of labor supply towards livestock businesses, to invest in land acquisition.

The acquisition of assets also extends to other business assets such as livestock sheds, rickshaws, vans, pumps and trees: Column 6 shows that after four years the value of such assets held by the ultra-poor is 159% higher relative to the controls. As with land, accumulation of these assets accelerates between years two and four with the latter effect being significantly larger than the former. This is mostly driven by the acquisition of livestock sheds (an obvious complement to livestock) and means of transport such as rickshaws and vans.

Combining all productive assets – livestock, land and other business assets – the QTE estimates in Figure IID reveal considerable heterogeneity in gains across the productive asset holding distribution. No ultra-poor households reduce their holding of productive assets, but households in the lower centiles gain little. At higher centiles the gains increase markedly. Understanding the causes of this heterogeneity in returns is critical to comprehending how to reach all ultra-poor households, and is an important matter to take up in future research.

The materialization of asset accumulation and diversification after four years underlines the value of having longer run data to study poverty trajectories. We return to examine the issue in Section VII, where we exploit data tracking the same ultra-poor households seven years after the program first started.

IV.D Comparison with Program Effects in Other Contexts

The program evaluated in this paper was started by BRAC in 2002 in Bangladesh and is still the only fully scaled version of the program which, by the end of our study in 2014, had reached over $360,000$ ultra-poor households containing 1.2 million individuals. It has served as a

template for similar programs that have been implemented in a variety of contexts by different implementing partners. Results from randomized evaluations of pilots of these programs in six countries – Ethiopia, Ghana, Honduras, India, Pakistan and Peru – have recently been published (Banerjee *et al.* 2015a).[30] Our analysis differs from those in Banerjee *et al.* (2015a) in four respects: (i) we collect information on hours worked in every labor activity over the course of one year rather than the last 24 hours or week and this allows us to minimize measurement error due to the fact that most jobs are seasonal or casual; (ii) we survey all beneficiaries in the scaled up version of the program rather than a sample in pilot versions and this allows us to estimate the full distribution of treatment effects; (iii) we survey a representative sample of households across the entire wealth distribution rather than ultra-poor households only and this allows us to quantify general equilibrium effects as well as the distributional effects of the program; (iv) we track beneficiaries four and seven years after the intervention rather than three, and this allows us to study poverty trajectories.[31]

Using our data from Bangladesh we replicate the ten key outcome variables studied in Banerjee *et al.* (2015a). These are all index variables capturing changes along ten dimensions – consumption, food security, assets, financial inclusion, labor supply, income, physical health, mental health, political awareness and women's empowerment.[32]

Table VI contains a comparison of the effects we observe in our study after four years relative to those observed by Banerjee *et al.* (2015a) after three years. What is striking is how similar the pattern of effects is across the broad set of ten outcome variables. In all settings: (i) per

[30]The implementing partners, mainly NGOs some of which received state support (e.g. Pakistan, Ethiopia) visited or were visited by BRAC Bangladesh at least twice during the design phase to seek guidance on program design. Thus, though they had to be adapted to particular circumstances of a country these programs share many of the features of the Bangladeshi BRAC TUP program.

[31]In three sites, Ghana, Honduras and Peru, Banerjee *et al.* (2015a) randomize the treatment both within and across villages and thus measure spillovers on non-treated ultra-poor. Our design, in contrast, allows us to measure spillovers on households across the wealth distribution as well distributional changes. In one site, West Bengal, beneficiaries are resurveyed seven years after the intervention and a preliminary note (Banerjee *et al.* 2016) reports that, consistent with our evidence in Section VII, the program has lasting impacts.

[32]The online Appendix describes the construction of outcome variables that we use to compare with Banerjee *et al.* (2015a) and notes any differences in how our variables are constructed. Even though the survey instruments were designed independently, we are able to construct similar variables along each of the ten outcome dimensions. The exceptions are mental health and political awareness where we use variables that differ somewhat from Banerjee *et al.* (2015a). Furthermore, for labor supply we use annual labor supply converted to a daily measure to account for seasonal variation whereas Banerjee *et al.* (2015a) use labor supply as measured for the past 48 hours or week.

capita (non-durable) consumption and food security (which captures food adequacy and whether meals are skipped) is significantly increased by the program (Columns 1 and 2); (ii) households are accumulating more household and productive assets as well as saving, borrowing and lending more (Columns 3 and 4); (iii) adult labor supply, both for the main woman in Bangladesh (Column 5) and for all adults in the six pilots (Column 6) also increases; (iv) income and revenues received by the main ultra-poor woman are increased (Column 7).[33]

This comparison of studies bolsters the external validity of the scaled version of the program we have evaluated in Bangladesh. In a variety of settings the combined evidence suggests the arrival of livestock rearing opportunities for the ultra-poor, through asset and skill transfers and other components of the TUP approach, enables them to expand their labor supply, increase their income and accumulate assets. This, in turn, leads to improvements in welfare along consumption and food security dimensions. A key difference of the TUP program from cash or food transfer programs is this focus on occupational change. The fact that the program has proven to be effective in reducing poverty through occupational change in different contexts makes us more confident that this type of program can be successfully implemented in contexts other than Bangladesh and by organizations other than BRAC.[34]

In Panel B of Table VI we compare non-economic impacts of the program across studies. Physical health, covering ability to perform physical tasks, work interruptions due to ill-health and self-perception of physical health, is significantly improved by the program (Column 8). Mental health, captured by a happiness perception measure and measures of experiencing anxiety and worry, is also improved (Column 9), and in Column 10 we see that the program enhances political awareness, captured by political activity or awareness of political representatives at different levels of government. Women also exert greater influence over household decisions after they become beneficiaries of the program (Column 11)). Across contexts, the program thus seems to have

[33]Our estimated treatment effects are generally larger than those in Banerjee *et al.* (2015a). This is likely driven by the fact that the latter are an average across sites, some of which had small or zero treatment effects. Our estimated effects are similar in magnitude to Banerjee *et al.* (2015a)'s estimates for West Bengal, which is the most similar setting to ours.

[34]Despite being given a choice, livestock was the main asset taken up in all six pilots, as was the case in Bangladesh. The type of livestock, however, varied strongly – sheep, goats and oxen in Ethiopia, goats and hens in Ghana, chickens and pigs in Honduras, goats and cows in India, goats in Pakistan, guinea pigs and hens in Peru.

far reaching effects on physical and mental health, political empowerment and empowerment within the household for ultra-poor women. Economic and social empowerment are both key objectives of the program and may reinforce one another. Duflo (2012), for example, hypothesizes that improved mental health may (partly) be what gave ultra-poor women in the India pilot the energy to work more, save and invest in their children. Looking at these links and interactions to better understand the mechanisms behind the Table VI results represents a fertile area for future research.

V. General Equilibrium and Distributional Effects

The magnitude of the asset and skills transfers, and the fact that treated ultra-poor households comprise, on average, 6% of the village population imply that the program might also affect economic outcomes for households in other wealth classes through general equilibrium effects and other spillovers. In Section V.A we provide evidence on these indirect effects which could be negative or positive. For instance, the new engagement in livestock rearing activities started by the ultra-poor could compete away the financial returns to non-poor women already engaged in these activities. Alternatively, the additional income generated by the ultra-poor could allow them to increase financial intermediation, thus developing village credit markets to the benefit of all. Our partial population experiment also allows us to quantify distributional effects and, in Section V.B, we focus on the extent to which the program enables the ultra-poor to close the gap with the near-poor.

V.A *Indirect Treatment Effects on Ineligible Households*

To estimate the indirect treatment effect on ineligible households we can simply estimate the same difference-in-difference specification (1) on the sample of ineligibles (Angelucci and De Giorgi 2009). To estimate the indirect treatment effect (ITE) on each wealth class of ineligible households we further interact treatment and survey waves indicators with class indicators:

$$\begin{aligned} y_{idt} &= \sum_{t=1}^{2}\sum_{c=1}^{3}\beta_t^c\left(W_t\times T_i\times C_i^c\right)+\gamma T_{id} \\ &+\sum_{t=1}^{2}\delta_t W_t+\sum_{c=1}^{3}\vartheta^c C_i^c+\sum_{t=1}^{2}\sum_{c=1}^{3}\nu_t^c W_t C_i^c+\sum_{c=1}^{3}\rho^c T_i C_i^c+\eta_d+\varepsilon_{idt}, \end{aligned} \tag{3}$$

where C_i^c are dummies that take value 1 if i belongs to class c (near poor, middle and upper class) and all other variables are as defined previously. We thus evaluate the effect of the program on the ineligibles by comparing the change in their outcomes in treated villages to the change in their outcomes in control villages. To benchmark the magnitude of the effects we report the four year effects in percentage of the control mean for the same wealth class in the same period.

As the primary objective of the program is to induce occupational change of ultra-poor women by enabling them to shift their labor supply towards livestock rearing, Table VII first examines general equilibrium impacts on the livestock businesses of ineligible women. Panel A shows indirect treatment effects pooling all ineligible households, and Panel B breaks these out by wealth group. In Columns 1 and 2 we see that the program has no significant impact on the value of cows or goats held by ineligible households, and Column 3 shows that annual hours devoted by ineligible women to livestock rearing are also unaffected. The point estimates are small both relative to the effects on the ultrapoor and relative to ineligible households in control villages.[35] This is *prima facie* evidence that the entry of ultra-poor women into this work activity does not crowd out richer women who were the main participants in these markets at baseline. In line with this, village level regressions on the price of milk and the transaction value of cows show no significant reductions.

Part of the explanation for these muted general equilibrium effects is that the cows transferred to the ultra-poor through the program only constitute 7% of the baseline village level stock of cows. So although the gains in cow holdings brought about by the program are highly significant for the ultra-poor, they only have modest effects on the total number of village cows as the herds of wealthier women are much larger. Markets where livestock and livestock products are sold tend to cover a larger area than the area of operations of a BRAC office with sub-district and regional markets being particularly important in the Bangladesh context. Also important is the fact that the livestock transferred to the ultra-poor are procured in regional markets (and not from livestock owners within villages).

Although ultra-poor women have limited involvement in livestock rearing at baseline they are

[35]It should be noted that the standard errors are large, suggesting that effects are heterogeneous. This notwithstanding, even the largest effect we cannot reject is orders of magnitudes smaller than the effect on the ultra-poor. For instance, the program increases the value of cows by 540 for the ultrapoor while the largest decrease we cannot rule out on the ineligibles is 56.

heavily involved in casual wage labor activities, accounting for 47% (58%) of the aggregate hours supplied in agricultural labor (maid services). The changes in labor allocation of the beneficiaries residing in treatment villages might therefore have general equilibrium effects on ineligible households in the village, and these might differ by wealth class. In Columns 4 and 5 of Panel A in Table VII, we see that agricultural labor and maid wages for ineligible women rise significantly as a result of the program. This result was already observed for ultra-poor women in Table III as a result of them significantly reducing their casual labor supply. What Table VII illustrates is that ineligible women who continue to work in these labor activities also benefit from these wages increases.

When we break out the results by wealth class in Columns 4 and 5 of Panel B we see that upper class households do not participate in casual wage labor and that effects are similar across other ineligible wealth classes, consistent with the fact that these are unskilled activities where the return does not vary much across individuals. In Columns 6 and 7 we see that ineligible women respond to the wage increase by reducing hours worked, although none of the effects are precisely estimated. Given the muted responses of labor supply across the three main female work activities practised in these village economies it is not surprising that the yearly earnings of ineligible women are unaffected by the program (Column 8).

In Table VIII we estimate indirect treatment effects to gauge if there are spillovers of the program on the expenditures and asset accumulation of ineligible households. Columns 1 and 2 show no changes in poverty rates or consumption expenditure per equivalent adult. This is true for ineligible households taken as a whole (Panel A) and when we break out by wealth class (Panel B). All coefficients are small and precisely estimated. This is a key result as it shows that ineligible households are not being made worse or better off by the program. In Figure A.IA in the Appendix we graph out the four year quantile treatment effects on consumption for ineligible households. Unlike Figure IIA, which shows large positive effects for eligibles, this figure is flat and lies along the zero line for the entire consumption distribution.

Column 3 of Table VIII shows that there is no spillover effect of the program on the value of household assets held by ineligible households taken together (Panel A) but we do see a positive

effect that is significant at the 10% level for middle class households when we break out by wealth class (Panel B). In Figure A.IB we see some limited evidence of an effect in higher quantiles but this is very muted. Columns 4-6 of Table VIII show no significant changes in the value of savings for ineligibles, nor in the probability that these households give or receive a loan. Though imprecisely estimated there is some suggestion that middle and upper class households are less likely to give loans after the program.

Land is an important asset to examine as it is a fixed resource in the village. Column 7 shows that although it is not precisely estimated there is evidence that ineligibles are losing land as whole (Panel A) and this is almost entirely coming from upper class households (Panel B). The magnitude of the gain in value of land for ultra-poor households (Table V) is similar to the loss for upper class households (Table VIII). This provides suggestive evidence that land is transferred from the richest to the poorest in these villages but what are relatively large gains for the ultra-poor are relatively small losses for the upper classes.

Finally, Column 8 shows that the value of other business assets (livestock sheds, rickshaws, vans, pumps etc.) significantly increases overall (Panel A) and for the near-poor and middle class wealth classes (Panel B). The effect represents a 23% increase overall and a 34%, 34% and 6% increase for near poor, middle and upper class households respectively. This could be due to the ultra-poor channeling some of their newly accumulated resources to others in the village or to other households reducing support to the ultra-poor. These findings are consistent with earlier studies that have shown causal links between savings behavior of the poor and improved outcomes for the non-poor through greater financial intermediation (Angelucci and De Giorgi 2009; Dupas *et al.* 2015).[36] However, the value of these business assets is low relative to the value of livestock and land (see Table I), thus the indirect treatment effect on total productive assets is negligible. Figure A.ID, which plots quantile treatment effects for the combined value of all productive assets (livestock, land and other productive assets), shows that, although there is evidence of asset

[36]Dupas *et al.* (2015) estimate how access to bank accounts impacts household's financial engagement, where they vary the spouse within the household to whom the bank account is assigned. The spillover effects are estimated through how treated households report changes in transfers they send and receive from others. While this and other papers have used field experiments to estimate spillover and general equilibrium impacts our data also allows us to compare changes in outcomes for ultra-poor households relative to near poor households as is discussed in the next Subsection.

accumulation in upper quantiles, none of these effects are statistically significant.

V.B Distributional Effects

Table I documented that, at baseline, the near-poor were better off than ultra-poor households. The partial population experiment allows us to compare how the lives of ultra-poor households have changed relative to the near-poor after four years. To do so we estimate a triple difference specification between baseline and year four, treatment and control villages, and ultra-poor and near-poor households. All outcomes are divided by the average difference between ultra-poor and near-poor in treatment villages at baseline, thus an estimated triple difference ζ equal to one indicates that the gap has entirely closed between the two groups. We estimate:

$$y_{idt} = \alpha + \sum_{t=1}^{2} \beta_t^1 \left(W_t T_i\right) + \sum_{t=1}^{2} \beta_t^2 N_i W_t + \sum_{t=1}^{2} \beta_t^3 N_i T_i$$

$$+ \sum_{t=1}^{2} \zeta_t \left(N_i \times W_t \times T_i\right) + \gamma T_i + \gamma^2 N_i + \sum_{t=1}^{2} \gamma_t^3 W_t + \eta_d + \varepsilon_{idt}, \tag{4}$$

where N_i equals 1 if i belongs to the near poor class and all other variables are as defined previously. The results from this exercise are shown in Figure III. The first bar in the figure indicates that by four years post-intervention, ultra-poor households have closed the (small) gap with near-poor households in terms of consumption expenditure. More remarkably, the same is true for the value of household assets, as shown in the second bar, despite the value of household assets held by the ultra-poor being half of that held by the near-poor at baseline. When we examine savings in the third bar we see that financial savings held by ultra-poor households are four times those held by near-poor households, from a baseline ratio of 1/3. This is a striking result as this effect is measured four years after the program first starts, and so two years after BRAC's direct involvement and when there is no encouragement to hold savings. The result for productive assets in the final bar in Figure III is also striking as we see that ultra-poor households now hold twice the value of productive assets held by the near-poor, including in areas that are not covered by the program such as land and business assets.

This set of findings suggest the program has significant distributional impacts between the ultra-poor and near-poor, and that on many dimensions the ultra-poor can be classified as firmly

entrenched within or above the near-poor wealth class, four years after the program began.

VI. Cost-Benefit Analysis

Table IX makes use of the estimated program impacts to gauge the magnitude of the benefits relative to the program costs and to estimate its internal rate of return (IRR). The average cost per treated household for the two year program is USD 1120 in 2007 PPP terms. We initially set the social discount rate at 5% in line with World Bank guidelines and report sensitivity analysis to alternative rates.

Since the ultimate goal of the program is to reduce poverty, we follow Banerjee *et al.* (2015a) and use changes in household consumption as our core measure of benefits. These include yearly changes in consumption expenditure and a one-off change in household assets as measured in year four. The underlying assumption is that the effect of increased financial and productive assets is fully incorporated in consumption changes. To the extent that asset accumulation as of year four will lead to even greater increases in consumption in the future we will underestimate the benefits of the program. Moreover, we make no attempt to price the utility gains to the ultra-poor arising from a smoother allocation of labor hours across days of the year (as was shown in Table III).[37]

Rows 1-4 in Table IX report ITT estimates of the program on consumption, for every year after the intervention up to year four. The year two and four effects are estimated from our midline and end-line surveys, respectively, while the one- and three-year effects are imputed using linear interpolation. Row 5 reports the net present value of future consumption changes from year 5 onward, assuming that year four changes are repeated for twenty years from the transfer date (so 16 more years after year 4). Our choice of time horizon is dictated by three facts: (i) the average beneficiary was forty years old when she received the asset in 2007, (ii) women in these villages work the same number of hours at 60 and older as they do at 40, (iii) the female life expectancy at birth is 71 today. As these women were born when life expectancy was lower and they live in the poorest areas of the country, we assume they will be able to continue working with the

[37] We focus on the benefits accruing to the ultra-poor alone as the program had no effect on the consumption of ineligible households (Table VIII, Column 2). Table VIII however shows that after four years the program increases the business asset holdings of ineligible households. We therefore underestimate the benefits accruing to these households to the extent that this will allow them to increase future consumption.

assets until they are 60, so 20 years from the transfer date. Below we present sensitivity analysis to shorter time horizons. In every case we assume that the benefits cease with the death of the original beneficiaries, which is a lower bound if other family members inherit the asset or continue to benefit from it after the death of the beneficiary.[38] Row 6 reports the year four change in the value of household assets (i.e. durables) and Row 7 adds these up to compute the net present value of benefits. This is divided by the program cost to obtain the benefit/cost ratio in Row 8.

The estimates show that the average benefits of the program are 3.21 times larger than its cost.[39] Table A.VII in the Appendix uses our quantile treatment effects to compute the ratio at different quantiles – it shows that the ratio is above 1 throughout. Row 8 of Table IX investigates sensitivity to different values of the discount rate and different time horizons. The ratio of average benefits to costs remains above one in all cases except if we assume that benefits disappear the year after our endline, in which case the ratio falls below the break-even point for the average ultra-poor household. If benefits last two years after endline, that is six years after transfer, the benefit to cost ratio is 1.06.

Row 9 shows the IRR under alternative assumptions about outside options and time horizons. The average internal rate of return in our baseline specification is 22% and it is positive and clearly above the discount rate; it goes to zero only when we assume that benefits disappear altogether one year after our endline (five years after the transfer).[40] While these calculations take into account that beneficiaries substitute away from casual wage labor and hence lose some earnings from that activity (see Table III), they do not take into account that beneficiaries work 206 more hours and 61 more days over the course of a year. The value of this time depends on its opportunity cost. We consider two scenarios: (i) assuming aggregate demand constraints for wage labor bind so there is zero opportunity cost of spending additional hours in livestock rearing; (ii) assuming unconstrained demand in casual wage labor and so the lost hourly wage is USD 0.34 per hour, that for agricultural wage labor (which is higher than for causal maid work, as Table I shows). This is likely to be an upper bound as recent micro studies suggest the true opportunity cost of

[38]For instance, Roy *et al.* (2015) show that men belonging to the households of the treated women benefited indirectly by being able to purchase productive assets from the women's additional earnings.

[39]Using the same methods, Banerjee *at al.* (2015a) report an average benefit/cost ratio of 1.59 for the six pilots.

[40]This is also above the average internal rate of return of 12% reported in Banerjee *et al.* (2015a).

labor is likely below the prevailing wage rate (Foster and Rosenzweig 2010; Kaur 2015).[41] The final row of Panel B in Table IX reports lower bounds for the IRR under the latter assumption as we deduct the value of 206 hours at USD 0.34 per hour from estimated consumption benefits. With this adjustment the IRR falls from 22% to 16% but it remains positive and larger than the social discount rate of 5%.

Finally, Panel C of Table IX measures program benefits in terms of productive asset accumulation (livestock, land, agricultural equipment and other machinery used for production) and financial assets (savings plus net lending). Row 12 shows that four years after the asset transfers, the average household has further accumulated productive assets valued at almost twice as much as the original transfer. Financial assets are included in this calculation but they account for less that 10% of the total. The high rates of asset accumulation suggest that future consumption gains might be sustainable. The next section uses descriptive data from seven year follow-up on the same households to provide indicative evidence on this issue.

VII. The Ultra-Poor in the Long-Run

To assess whether the one-off asset and skills transfers provided by the program set the ultra-poor on a long-run trajectory out of poverty, we fielded a survey to the same ultra-poor households in 2014, seven years after the program's implementation. We were able to trace 93% of the households. As described above, the evaluation design was such that the program would be offered in control villages starting in 2011 (i.e. after the year four follow-up survey). By 2014, every control BRAC branch office had treated some villages within its radius. To choose which villages and which individuals to treat, BRAC program officers followed the same process as in 2007, namely they made a list of all villages in the branch ranked from poorest (i.e. with the largest number of poor households) to least poor, and then implemented a participatory rural appraisal (PRA) to identify the beneficiaries in each village starting from the poorest villages and stopping when they reached their target number of beneficiaries.

[41] Foster and Rosenzweig (2010) use data from rural India to document that various market imperfections such as supervision costs, credit market imperfections and scale economies lead to a surplus of labor on small farms: they quantify that 20% of the Indian agricultural labor force is surplus to requirement. Kaur (2015) finds that casual wage labor markets in rural India are well characterized by downward nominal wage rigidity (that are driven by fairness concerns of employers).

In practice this implies that 49% of the villages originally assigned to control have at least one woman treated and 20% of the originally selected beneficiaries plus 10% of the original "near poor" were treated. In 2014 we thus have three groups: the early treated (in 2007), the late treated (in 2011) and the untreated controls. The challenge in identifying the effect of the program in 2014 is that the selection of the late treated is correlated to the outcome of interest: poverty. Indeed, given BRAC's targeting strategy, the late treated have lower consumption expenditures, durables and other assets than those left untreated in 2011.

To provide evidence on the long run impact of the program we follow two strategies. The first simply extends specification (1) to include the 2014 survey wave and all control villages/individuals regardless of whether they are late treated or not. We note that to the extent that the program has some effect on the late treated after three years, this strategy yields a lower bound on the actual effect because 1/5 of the control group is actually treated. The second strategy requires making assumptions about the size of the effect on the late treated. To this purpose we exploit the QTE estimates on the original treated to create counterfactuals of the effect of the program on the late treated. Since by 2014 these have been treated for three years, we interpolate between our two- and four-year estimates of the ITT on the originally treated group to derive a counterfactual effect for the treated controls in 2014. Table X reports three difference in difference estimates derived by assuming that the effect on the late treated is equal to the median, 75th percentile and 25th percentile treatment effect on the early treated. Throughout we focus on the outcomes used in the cost-benefit analysis above: household consumption expenditures, household assets, savings and productive assets.

Table X reports difference-in-difference estimates at each survey wave (2009, 2011, 2014) using the two strategies above. As for the earlier estimates, Table A.VIII reports the equivalent ANCOVA specifications. The results are consistent across outcomes and specifications: the seven year effects are positive and precisely estimated. Moreover we never reject the null that the seven year effects on consumption are equal to the four year effects, thus reinforcing the conclusions of the cost benefit analysis. The only effect that is systematically smaller after seven years is that on savings, which falls by about 50%, depending on the specification. Further analysis shows

that this is coupled with an increase in land access through purchases, which are captured in the value of productive assets, and especially rentals, which are not. The most conservative estimate suggests that average quantity of land rented increase by 3.5 decimals after four years and by 4.4 decimals after seven. Given that agricultural land is a key asset in the villages we study and is also the asset which most clearly differentiates poor from non-poor households, this is a striking change.

Overall, while these seven year results must be interpreted with caution as the responses of the original beneficiaries might be an imperfect counterfactual for the responses of the late treated controls, a major difference would be needed to reverse the conclusion that a one-off transfer of assets and skills allows the ultra-poor to escape poverty in a sustainable way.

VIII. Conclusions

The question of how to eliminate extreme poverty by 2030 has now risen to the top of the development policy agenda and there is a growing realization that the poorest may be being bypassed both by economic growth and by current anti-poverty programs.[42] Our results suggest the labor activities the poor can access and their ability to exit poverty are intrinsically linked. The women we study possess no means of production other than their labor and lie at the bottom rung of the employment ladder in rural villages, facing low returns to and irregular demand for their labor. They live predominantly in the *monga* or famine areas of Bangladesh and in the work they do they are not very different from the majority of Indian famine victims in the 19th and 20th centuries (Dreze 1988).

We find that the TUP program enables these ultra-poor women to take on the labor market activities of better off women in the same villages as they dramatically expand overall labor supply, principally by working more hours in livestock rearing. As their labor supply expands and their employment becomes more regular, they experience a 21% increase in earnings which allows them to accumulate further productive and assets and set off on a sustainable trajectory out of poverty.

Our evidence demonstrates that enabling the poor to allocate their labor to the activities chosen by richer women in their villages may have a central role to play in eliminating extreme poverty.

[42] This was part of a longer set of Sustainable Development Goals agreed in 2015.

However, given the TUP program has multiple components bundled together, we understand little so far about which elements are critical to unleashing this process of change. Getting a better sense of this is therefore a key priority. Understanding why we observe heterogeneity in program returns is also critical for gaining a better understanding of the determinants of poverty.

After four years we find that the program was highly cost-effective with an IRR of 22%, and that a sizable fraction of ultra-poor households would have enjoyed positive returns had they been able to finance these investments from either the formal or microfinance sectors. Given these findings it would also seem worthwhile exploring versions of the program where households have to repay some fraction of the cost of the asset transfer as a means of reducing program costs.

What is also important is to understand how different ways of financing the program affect the cost benefit analysis. Buera *et al.* (2014) study the scale-up properties of TUP-style programs using a quantitative general equilibrium model of occupational choice with credit market imperfections to simulate the aggregate impacts of a one-time *redistribution* (not transfer from outside) of assets (ignoring skill transfers). Their simulations generate muted long run impacts because they find only the top quartile most productive individuals transition to capital intensive activities. This does not match our micro evidence where the TUP program appears well targeted so the share of ultra-poor engaged in livestock rearing rises by 48pp four years post transfer. More work needs to be done to bring together these macro and micro approaches, including developing models that incorporate the skills transfer component of the program and model transfers as coming from outside the village.

A key difference of the TUP program from most cash or food transfer programs is that it is a one-off, big push intervention. Though big push programs require large up-front investment, our evidence suggests they are cost-effective and lead to sustained increases in household welfare. Indeed, the observed pattern of asset accumulation between years two, four and seven indicates that, although the cost of the two year program is fixed, the benefits grow in the short term and stabilize in the medium term. This may be a key advantage relative to cash and food transfer programs which do not encourage occupational change, where annual costs are lower but need to be recurrent in order to exert an influence on consumption (see also Blattman *et al.* 2014;

Banerjee *et al.* 2015a, Banerjee 2016).[43]

Understanding whether and how governments can take up these programs and whether they can be adapted to urban settings are all unknowns that will have a critical bearing on whether this idea spreads and scales. The juxtaposition of the goal of eliminating extreme poverty by 2030 and the promising set of initial results in this and related papers does, however, suggest that taking up these research challenges would be a worthwhile endeavor.

Department of Economics, IGC and STICERD, London School of Economics
Department of Economics, IGC and STICERD, London School of Economics
Research and Evaluation Unit, BRAC
Department of Economics and IGIER, Universita' Bocconi
Department of Economics, UCL, IFS and IGC
Research and Evaluation Unit, BRAC International

Supplementary Materials

An Online Appendix for this article can be found at QJE online (qje.oxfordjournals.org).

References

[1] alatas, vivi, abhijit.v.banerjee, rema hanna, benjamin a. olken and julia tobias, "Targeting the Poor: Evidence from a Field Experiment in Indonesia," *American Economic Review,* 104 (2012), 1206-40.

[2] anagol, santosh, alvin etang and dean karlan, "Continued Existence of Cows Disproves Central Tenets of Capitalism?," Yale University Economic Growth Centre Discussion Paper No.1031, 2014.

[3] angelucci, manuela and giacomo de giorgi, "Indirect Effects of an Aid Program: How Do Liquidity Injections Affect Non-Eligibles' Consumption?," *American Economic Review,* 99 (2009), 486-508.

[43] At the same time, Banerjee *et al.* (2015b) analyze data from seven randomized controlled trials of government-run cash transfer programs in six developing countries and find no systematic evidence that cash transfer programs discourage work.

[4] ASHLEY, SAM, SARAH J. HOLDEN AND PETER B. S. BAZELEY, *Livestock in Poverty-focused Development*, (Crewkerne, UK: The IDL Group, In Development Ltd., 1999).

[5] ATTANASIO, ORAZIO AND BRITA AUGSBURG, "Holy Cows or Cash Cows?," IFS Working Paper No. W14/14, 2014.

[6] BANERJEE, ABHIJIT V., ESTHER DUFLO, NATHANAEL GOLDBERG, DEAN KARLAN, ROBERT OSEI, WILLIAM PARIENTE, JEREMY SHAPIRO, BRAM THUYSBAERT AND CHRIS UDRY, "A Multi-faceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries," *Science* 348: Issue 6236 (2015a).

[7] BANERJEE, ABHIJIT V., REMA HANNA, GABRIEL KREINDLER AND BENJAMIN A.OLKEN, "Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programs Worldwide," Working Paper, MIT, 2015.

[8] BANERJEE, ABHIJIT V., ESTHER DUFLO, RAGHABENDRA CHATTOPADHYAY AND JEREMY SHAPIRO, "The Long Term Impacts of a "Graduation Program: Evidence from West Bengal," Working Paper, MIT, 2016.

[9] BANERJEE, ABHIJIT V. "Policies for a Better-fed World," *Review of World Economics,* 152 (2016), 3-17.

[10] BARDHAN, KALPANA, "Work Patterns and Social Differentiation: Rural Women of West Bengal," in *Contractual Arrangements, Employment and Wages in Rural South Asia*, Hans P. Binswanger and Mark R.Rosenzweig, ed. (New Haven, CT: Yale University Press, 1984).

[11] BARDHAN, PRANAB, *Land, Labor and Rural Poverty* (New York: Columbia University Press, 1984).

[12] BLATTMAN, CHRIS, NATHAN FIALA AND SEBASTIAN MARTINEZ, "Generating Skilled Self-employment in Developing Countries: Experimental Evidence from Uganda," *Quarterly Journal of Economics,* 129 (2014), 697-752.

[13] BRAC, Stories of Targeting: Process Documentation of Selecting the Ultra Poor for CFPR/TUP Programme, BRAC RED CFPR-TUP WP Series No.1, 2004.

[14] BRAC, Annual Report, 2014, Dhaka, Bangladesh

[15] BRAC, CFPR Programme Annual Progress Report, 2015, Dhaka, Bangladesh

[16] BRYAN GHARAD, SHYAMAL CHOWDHURY AND A. MUSHFIQ MOBARAK,"Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh," *Econometrica* 82 (2014), 1671-758.

[17] BUERA, FRANCISCO, JOSEPH P.KABOSKI AND YONGSEOK SHIN, "Macro-Perspective on Asset Grants Programs: Occupational and Wealth Mobility," *American Economic Review* 104 (2014), 159-64.

[18] CAMERON, A. COLIN, JONAH B. GELBACH AND DOUGLAS L. MILLER, "Bootstrap-Based Improvements for Inference with Clustered Errors", *Review of Economics and Statistics*, 90 (2008), 414-27.

[19] CHAMBERS, ROBERT, *Rural Development: Putting the Last First* (Routledge: London and New York, 1983).

[20] DE MEL, SURESH, DAVID MCKENZIE AND CHRIS WOODRUFF, "Returns to Capital in Microenterprises: Evidence from a Field Experiment," *Quarterly Journal of Economics,* 123 (2008), 1329-72.

[21] DE MEL, SURESH, DAVID MCKENZIE AND CHRIS WOODRUFF, "Are Women More Credit Constrained? Experimental Evidence on Gender and Microenterprise Returns," *American Economic Journal: Applied,* 1 (2009), 1-32.

[22] DEATON, ANGUS, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy* (Johns Hopkins University Press, Baltimore, Maryland: The World Bank, 1997).

[23] DONALDSON, DAVE, "The Gains from Market Integration," *Annual Review of Economics*, 7 (2015): 619-647.

[24] DREZE, JEAN, "Famine Prevention in India," WIDER Working Papers No. WP45, 1988.

[25] DREZE, JEAN, "Poverty in India and the IRDP delusion," *Economic and Political Weekly*, 25 (1990), A95-A104.

[26] DREZE, JEAN AND AMARTYE SEN, *Hunger and Public Action* (Oxford University Press, Oxford, 1991).

[27] DUFLO, ESTHER, "Human Values and the Design of the Fight Against Poverty," Tanner Lectures on Human Values, 2012.

[28] DUPAS, PASCALINE, ANTHONY KEATS AND JONATHAN ROBINSON, "The Effect of Savings Accounts on Interpersonal Financial Relationships: Evidence from a Field Experiment in Rural Kenya," NBER Working Paper No. 21339, 2015.

[29] ESWARAN, MUKESH AND ASHOK KOTWAL, "A Theory of Contractual Structure in Agriculture," *American Economic Review,* 75 (1985), 352-67.

[30] FEI, JOHN C. H, AND GUSTAV RANIS, *Development of the Labor Surplus Economy: Theory and Policy* (Homewood, Illinois: Richard D. Irwin for the Econo- mic Growth Center, Yale University, 1964)

[31] FINK, GUNTHER, B. KELSEY JACK AND FELIX MASIYE, "Seasonal Credit Constraints and Agricultural Labor Supply: Evidence from Zambia," NBER Working Paper No. 20218, 2014.

[32] FOSTER, ANDREW D. AND MARK R. ROSENZWEIG, "A Test for Moral Hazard in the Labor Market: Contractual Arrangements, Effort, and Health," *Review of Economics and Statistics,* 76 (1994), 213-27.

[33] FOSTER, ANDREW D. AND MARK R. ROSENZWEIG, "Comparative Advantage, Information and the Allocation of Workers to Tasks: Evidence from an Agricultural Labor Market," *Review of Economic Studies,* 63 (1996), 347-74.

[34] FOSTER, ANDREW D. AND MARK R. ROSENZWEIG, " Is There Surplus Labor in India?," Working Paper, Yale University, 2010.

[35] GOLDBERG, JESSICA, "Kwacha Gonna Do? Experimental Evidence about Labor Supply in Rural Malawi," *American Economic Journal: Applied Economics*, 8 (2016), 129-49

[36] GOLLIN, DOUGLAS, "Getting Income Shares Right," *Journal of Political Economy,* 110 (2002), 458-74.

[37] GULESCI, SELIM, "Labor-Tying and Poverty in a Rural Economy: Evidence from Bangladesh," Working Paper, Bocconi University, 2016.

[38] HOSSAIN, NAOMI AND IMRAN MATIN, "Engaging Elite Support for the Poorest? BRAC's Expereince with the Ultra Poor Programme," BRAC RED CFPR-TUP WP Series No.3, 2004.

[39] IMBENS, GUIDO AND JEFFREY M.WOOLDRIDGE, "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature* 47 (2009), 5-86.

[40] JAYACHANDRAN, SEEMA "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries," *Journal of Political Economy,* 114 (2006), 538-75.

[41] KHAN, ZARINA R., *Women, Work and Values: Contradictions in the Prevailing Notions and the Realities of Women's Lives in Rural Bangladesh* (Centre for Social Studies: Dana Publishers, Dhaka, 1992).

[42] KAUR, SUPREET, "Nominal Wage Rigidity in Village Labor Markets," NBER Working Paper No. 20770, 2015.

[43] KHANDKER, SHAHIDUR R. AND WAHIDUDDIN MAHMUD, *Seasonal Hunger and Public Policies: Evidence from Northwest Bangladesh*, (World Bank, Washington D.C., 2012).

[44] KLING, JEFFREY R, JEFFREY B. LIEBMAN AND LAWRENCE F. KATZ, "Experimental Analysis of Neighborhood Effects," *Econometrica,* 75 (2007): 83-119.

[45] KREMER, MICHAEL, JEAN N. LEE, JONATHAN M. ROBINSON AND OLGA ROSTAPSHOVA, "The Return to Capital for Small Retailers in Kenya: Evidence from Inventories," Working Paper, Harvard University, 2010.

[46] LEWIS, W. ARTHUR., "Economic Development with Unlimited Supplies of Labor," *The Manchester School,* 22 (1954), 139-91.

[47] MOFFITT, ROBERT A., ""Policy interventions, low-level equilibria, and social interactions," *Social dynamics,* 4 (2001), 6-17.

[48] ROSE, ELAINA, "Ex Ante and Ex Post Labor Supply Response to Risk in a Low-income Area," *Journal of Development Economics* 64 (2001), 371-88.

[49] ROSENZWEIG, MARK R., "Rural Wages, Labor Supply, and Land Teform: A Theoretical and Empirical Analysis," *American Economic Review,* 68 (1978), 847-61.

[50] ROSENZWEIG, MARK R., "Labor Markets in Low-income Countries," in *Handbook of Development Economics Vol.1*, H.Chenery and T.N.Srinivasan (ed.), (Elsevier, 1988).

[51] ROSENZWEIG, MARK R. AND KENNETH I. WOLPIN, "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investment in Bullocks in India," *Journal of Political Economy,* 101 (1993), 223-44.

[52] ROY, SHALINI, JINNAT ARA, NARAYAN DAS AND AGNES R. QUISUMBING, "Flypaper Effects' in Transfers Targeted to Women: Evidence from BRAC's "Targeting the Ultra Poor" program in Bangladesh," *Journal of Development Economics,* 117 (2015), 1-19.

[53] SHABAN, RADWAN A., "Testing Between Competing Models of Sharecropping," *Journal of Political Economy,* 95 (1987), 893-920.

[54] YOUNG, ALWYN, Improved, Nearly Exact, Statistical Inference with Robust and Clustered Covariance Matrices using Effective Degrees of Freedom Corrections,Working Paper, London School of Economics and Political Science, 2016.

[55] WORLD BANK, *World Bank Annual Report,* 2014, Washington D.C..

Table I: Household Characteristics and Asset Holdings, by Wealth Class

	(1) Ultra-Poor	(2) Near-Poor	(3) Middle Class	(4) Upper Class
Household Characteristics				
Share of population in this wealth class	.061	.219	.585	.135
Primary female is the sole earner	.409	.250	.142	.120
Primary female is illiterate	.929	.832	.736	.489
Consumption and Assets				
Household is below the $1.25 a day poverty line	.530	.493	.373	.121
Consumption Expenditure (per adult equivalent)	627.8	645.1	759.5	1234.2
Household assets [USD]	36.5	68.1	279.9	1663.4
Household savings [USD]	7.9	22.1	84.5	481.9
Household receives loans	.191	.393	.498	.433
Household gives loans	.012	.018	.030	.067
Business assets (excl. livestock and land) [USD]	22.9	54.4	286.1	1569.8
Livestock				
Household owns cows	.055	.154	.469	.733
Household owns goats	.092	.142	.300	.425
Value of cows [USD]	33.8	120.2	633.8	1559.1
Value of goats [USD]	7.97	12.8	39.8	71.3
Household rents cows for rearing	.070	.148	.118	.030
Household rents goats for rearing	.111	.157	.102	.021
Land				
Household owns land	.066	.107	.487	.911
Value of land owned [USD]	200.0	491.2	6789.6	40125.1
Household rents land for cultivation	.060	.143	.276	.168
Number of sample households	6732	6743	6328	2036

Notes: 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: the ultra-poor are ranked in the bottom wealth bins (4th if 4 bins are used, 5th if 5 are used) and meet the program eligibility criteria, the near-poor are ranked in the bottom wealth bins and do not meet the program eligibility criteria, the middle class are ranked in the middle wealth bins (2nd and 3rd if 4 are used, 2nd, 3rd and 4th if 5 are used) and the upper classes are those ranked in the top bin. The number of sample households in each wealth class at baseline is reported at the foot of the table. The poverty line threshold used is $1.25 per person per day. Consumption expenditure is defined as total household expenditure over the previous year divided by adult equivalents in the household. The adult equivalence scale gives weight .5 to each child younger than 10. The expenditure items covered are: food (both purchased and produced), fuel, cosmetics, entertainment, transportation, utilities, clothing, footwear, utensils, textiles, dowries, education, charity and legal expenses. Household assets include jewelry, sarees, radio, television, mobile phones, furniture, etc. Household savings refer to value of savings held at home, at any bank, at any MFI and with saving guards. Loans are from both formal and informal sources. Business assets include pumps, livestock sheds, trees, rickshaws and others. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Table II: Labor Market Activities of Women, By Wealth Class

Means, standard deviation in parentheses

	(1) Ultra-Poor	(2) Near-Poor	(3) Middle Class	(4) Upper Class
Engaged in any income generating activity	.843	.810	.863	.903
Total hours worked in the past year	991	769	553	502
	(894)	(812)	(596)	(502)
Total days worked in the past year	252	265	302	325
	(137)	(142)	(123)	(103)
Casual Wage Labor:				
Hours devoted to agricultural labor	258	196	47.7	3.05
	(533)	(467)	(236)	(49.9)
Hours devoted to domestic maid	388	193	41.9	.648
	(708)	(516)	(251)	(22.7)
Capital-intensive activities:				
Hours devoted to livestock rearing (cows/goats)	121	221	366	404
	(265)	(341)	(390)	(370)
Number of sample households	6732	6743	6328	2036

Notes: 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: the ultra-poor are ranked in the bottom wealth bins (4th if 4 bins are used, 5th if 5 are used) and meet the program eligibility criteria, the near-poor are ranked in the bottom wealth bins and do not meet the program eligibility criteria, the middle class are ranked in the middle wealth bins (2nd and 3rd if 4 are used, 2nd, 3rd and 4th if 5 are used) 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 foot of the table. Engagement in any income generating activity covers all potential activities.

Table III: Treatment Effects on the Labor Supply and Earnings of Ultra-Poor Women

Sample: Ultra Poor Women
Standard Errors in Parentheses, Clustered by BRAC Branch Area

Panel A: Labor Supply	Livestock		Agriculture		Maid		All Activities	
	(1) Hours	(2) Days	(3) Hours	(4) Days	(5) Hours	(6) Days	(7) Hours	(8) Days
Program impact after 2 years	488***	205.5***	-42.3	-3.54	-57.4	-8.45	341***	72.4***
	(30.7)	(11.1)	(53.0)	(7.02)	(42.9)	(5.88)	(67.9)	(10.0)
Program impact after 4 years	415***	171.6***	-46.2	-4.77	-117**	-16.77***	206***	61.1***
	(38.9)	(10.9)	(42.7)	(5.43)	(45.0)	(5.82)	(73.0)	(12.5)
Control mean at four year follow-up	191.00	94.76	278.14	35.40	447.05	63.97	1217.00	277.40
Four year impact: % change	217%	181%	-17%	-13.5%	-26%	-26%	17%	22%
Two year impact = Four year impact [p-value]	.111	.023	.930	.831	.125	.125	.080	.179
Adjusted R-squared	.335	.367	.184	.183	.067	.061	.072	.069
Number of ultra-poor women	6732	6732	6732	6732	6732	6732	6732	6732
Number of observations (clusters)	20196 (40)	20196 (40)	20196 (40)	20196 (40)	20196 (40)	20196 (40)	20196 (40)	20196 (40)

Panel B: Earnings	Livestock	Agriculture		Maid		All Activities
	(9) Earnings	(10) Wage	(11) Earnings	(12) Wage	(13) Earnings	(14) Earnings
Program impact after 2 years	80***	.028	-9.99	.034	-11.48	62.3**
	(14.0)	(.021)	(13.98)	(.022)	(11.36)	(30.17)
Program impact after 4 years	115***	.053**	-3.89	.074***	-25.25**	87.8***
	(14.1)	(.024)	(13.97)	(.019)	(11.57)	(28.58)
Control mean at four year follow-up	18.48	.441	96.44	.354	112.84	410.92
Four year impact: % change	16%	12%	-4%	21%	-22%	21%
Two year impact = Four year impact [p-value]	.049	.219	.701	.080	.205	.455
Adjusted R-squared	.127	.486	0.178	.241	.095	0.088
Number of ultra-poor women	6732	6732	6732	6732	6732	6732
Number of observations (clusters)	20120 (40)	5227 (40)	19883 (40)	5833 (40)	19796 (40)	20135 (40)

Notes: *** (**) (*) indicates significance at the 1% (5%) (10%) level. Intent-to-treat estimates are reported based on a difference-in-difference specification estimated using OLS. This regresses the outcome of interest for woman i in village v in survey wave t on a constant, a dummy for whether the woman resides in a treated village, dummies for the two follow-up survey waves (two and four years post-intervention), the interaction between the treatment assignment dummy and each survey wave dummy, and a set of strata (sub-district) fixed effects. The coefficients shown are those on the treatment-survey wave interaction terms. Standard errors are clustered by BRAC branch area. All outcomes are measured at the individual level (for the ultra-poor woman in the household), and defined for the year prior to survey date. We report the mean of each dependent variable as measured at baseline in treated villages. In all Columns we report the p-value on the null hypothesis that the two and four year ITT impacts are equal. The number of ultra-poor is the number of eligible women that are observed at baseline and in both follow-up survey waves. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Table IV: Treatment Effects on Consumption, Household and Financial Assets of Ultra-Poor Households

DiD ITT Estimates: Household Level Outcomes
Sample: Ultra Poor Households
Standard Errors in Parentheses, Clustered by BRAC Branch Area

	Poverty and Consumption			Financial Assets		
	(1) Below Poverty Line	(2) Consumption Expenditure (per adult equivalent)	(3) Value of Household Assets	(4) Household Cash Savings	(5) Household Receives Loans	(6) Household Gives Loans
Program impact after 2 years	-.051	30.19	6.86	54.54***	.123***	.042***
	(.046)	(25.34)	(7.26)	(4.60)	(0.03)	(0.01)
Program impact after 4 years	-.084**	62.62***	39.65***	53.22***	.110***	.051***
	(.038)	(20.82)	(9.08)	(4.01)	(0.03)	(0.01)
Control mean at four year follow-up	.624	575.73	69.69	425%	.220	.016
Four year impact: % change	-13.5%	11%	57%	24%	50%	319%
Two year impact = Four year impact [p-value]	.379	.111	.000	.781	.714	.527
Adjusted R-squared	.032	.044	.082	.204	.086	.026
Number of ultra-poor women	6732	6732	6732	6732	6732	6732
Observations (clusters)	18882(40)	18838 (40)	20196 (40)	20179 (40)	20196 (40)	20196 (40)

Notes: *** (**) (*) indicates significance at the 1% (5%) (10%) level. Intent-to-treat (ITT) estimates are reported based on a difference-in-difference specification estimated using OLS. All outcomes are measured at the household level, using data on ultra poor households with an eligible woman resident in them at baseline. This regresses the outcome of interest for household h in village v in survey wave t on a constant, a dummy for whether the household resides in a treated village, dummies for the two follow-up survey waves (two and four years post-intervention), the interaction between the treatment assignment dummy and each survey wave dummy, and a set of strata (sub-district) fixed effects. The coefficients shown are those on the treatment-survey wave interaction terms. Standard errors are clustered by BRAC branch area. In Column 1, the poverty line threshold used is $1.25 per person per day, as measured in 2007 prices. In Column 2, consumption expenditure is defined as total household expenditure over the previous year divided by adult equivalents in the household. The adult equivalence scale gives weight .5 to each child younger than 10. The expenditure items covered are: food (both purchased and produced), fuel, cosmetics, entertainment, transportation, utilities, clothing, footwear, utensils, textiles, dowries, education, charity and legal expenses. In Column 3, household assets include jewelry, sarees, radio, television, mobile phones, furniture, etc. In Column 4, household cash savings refer to value of savings held at home, at any bank, at any MFI and with saving guards. We report the mean of each dependent variable as measured at baseline in treated villages. In all Columns we report the p-value on the null hypothesis that the two and four year ITT impacts are equal. The number of ultra-poor is the number of eligible women that are observed at baseline and in both follow-up survey waves. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Table V: Treatment Effects on Productive Assets Held by Ultra-Poor Households

DiD ITT Estimates: Household Level Outcomes
Sample: Ultra Poor Households
Standard Errors in Parentheses, Clustered by BRAC Branch Area

	(1) Value of Cows	(2) Value of Goats	(3) Rents Land	(4) Owns Land	(5) Value of Land owned	(6) Value of Other Business Assets
Program impact after 2 years	484.65***	28.11***	.069***	.005	39.80	23.84***
	(19.46)	(3.77)	(.020)	(.011)	(75.23)	(6.85)
Program impact after 4 years	539.66***	20.57***	.110***	.026*	326.98**	64.76***
	(45.16)	(4.12)	(.022)	(.012)	(131.27)	(11.91)
Control mean at four year follow-up	61.89	9.26	.079	.058	400.61	40.72
Mean value of asset transfer from program	464.03	39.9	-	-	-	-
Four year impact: % change (net of transfer if positive)	122%	-208%	139%	45%	82%	159%
Two year impact = Four year impact [p-value]	.148	.004	.054	.005	.002	.000
Adjusted R-squared	0.390	0.109	.077	.034	0.019	0.066
Number of ultra-poor women	6732	6732	6732	6732	6732	6732
Observations (clusters)	20182 (40)	20072 (40)	20196 (40)	20196 (40)	20195 (40)	20195 (40)

Notes: *** (**) (*) indicates significance at the 1% (5%) (10%) level. Intent-to-treat (ITT) estimates are reported based on a difference-in-difference specification estimated using OLS. All outcomes are measured at the household level, using data on ultra poor households with an eligible women resident in them at baseline. This regresses the outcome of interest for household h in village v in survey wave t on a constant, a dummy for whether the household resides in a treated village, dummies for the two follow-up survey waves (two and four years post-intervention), the interaction between the treatment assignment dummy and each survey wave dummy, and a set of strata (sub-district) fixed effects. The coefficients shown are those on the treatment-survey wave interaction terms. Standard errors are clustered by BRAC branch area. In Column 6, business assets include pumps, livestock sheds, trees, rickshaws and others. We report the mean of each dependent variable as measured at baseline in treated villages. In all Columns we report the p-value on the null hypothesis that the two and four year ITT impacts are equal. The number of ultra-poor is the number of eligible women that are observed at baseline and in both follow-up survey waves. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Table VI: Comparison with Pilot Results from Six Countries

Standard Errors in Parentheses, Clustered by BRAC Branch Area

Panel A	**(1) Total per capita consumption, standardized**	**(2) Food security index**	**(3) Asset index**	**(4) Financial inclusion index**	**(5) Total time spent working by main woman, standardized**	**(6) Total time spent working by both respondents pooled, standardized**	**(7) Incomes and revenues index**
Treatment effect - four year endline	0.314***	0.256***	0.327***	0.313***	0.122*	0.065	0.627***
	(0.034)	(0.079)	(0.029)	(0.040)	(0.065)	(0.047)	(0.074)
Treatment Effect in Banerjee et al (2015a) - three year endline	*0.120****	*0.113****	*0.249****	*0.212****	*n/a*	*0.054****	*0.273****
	(0.024)	*(0.022)*	*(0.024)*	*(0.031)*		*(0.018)*	*(0.029)*

Panel B	**(8) Physical health index**	**(9) Mental health index**	**(10) Political Awareness index**	**(11) Women's empowerment index**
Treatment effect - four year endline	0.108***	0.077*	0.269***	0.077
	(0.027)	(0.043)	(0.091)	(0.056)
Treatment Effect in Banerjee et al (2015a) - three year endline	*0.029*	*0.071****	*0.064****	*0.022*
	(0.020)	*(0.020)*	*(0.019)*	*(0.025)*

Notes: Following Banerjee *et al.* (2015a), we estimate ITT by regressing endline outcomes on baseline outcomes and randomization strata (sub-districts). We construct indices first by defining each outcome so that higher values correspond to better outcomes. We then standardize each outcome into a z-score, by subtracting the control group mean and dividing by the control group standard deviation (SD) at the corresponding survey round. We then average all of the z-scores, and again standardize to the control group within each round. The variables used for each index are described in detail in the Appendix. All indices but Mental Health and Political Awareness are directly comparable. Column 1 reports standardized total per capita consumption per month. The food security index in Column 2 is based on survey responses regarding whether the household had a food surplus or deficit, enough food to eat over the last month and could afford to have two meals per day most of the time during the last year. The asset index in Column 3 is constructed based on the total value of productive and household assets measured in terms of a numeraire asset and standardized. The financial inclusion index in Column 4 is constructed based on the amount borrowed in the last 12 months from all sources, informal sources and formal sources, and total savings at the time of the survey. Column 5 reports a standardized measure of the total time the main female household member spent in productive activities on a typical day during the past year, and Column 6 pools the same measure for both the female respondent and the male household head where applicable. The income and revenues index in Column 7 is constructed based on monthly household livestock revenue and income from agriculture, non-farm micro-enterprises and paid labor as reported by the main female respondent. The physical health index in Column 8 is constructed based on respondents' self-reported ability to perform physical tasks, whether any household member had an illness in the 15 days before the survey and whether this interfered with any income-generating activity, and the respondent's self-perception of her current health. The mental health index in Column 9 is constructed based on self-reported happiness and mental anxiety. The political awareness index in Column 10 is based on whether the respondent can correctly name politicians at different levels and is aware of the lowest legal age for voting. The women's empowerment index in Column 11 is based on women's responses to a series of questions regarding their influence over household decision-making in several scenarios. Our estimates are based on the sample of 6,732 eligible women used throughout the paper. The second row reports the endline 2 estimates from Table 3 in Banerjee *et al.* (2015a), based on a sample that varies from 9,482 to 9,508.

Table VII: Indirect Treatment Effects on Livestock and Casual Labor Markets of Women in Non-Eligible Households

DiD ITE: Program impact after 4 years
Sample: Non-Eligible Households
Standard Errors in Parentheses, Clustered by BRAC Branch Area

	(1) Value of Cows	(2) Value of Goats	(3) Hours devoted to livestock rearing (main woman)	(4) Wage-Maids (main woman)	(5) Wage-agriculture (main woman)	(6) Hours devoted to maid jobs (main woman)	(7) Hours devoted to agricultural wage jobs (main woman)	(8) Yearly earnings (main woman)
Panel A. Pooled								
Program impact after 4 years	-9.53	0.885	5.28	.044**	.043*	-16.10	-18.25	-28.75
	(23.02)	(2.49)	(43.89)	(.020)	(.024)	(18.99)	(25.47)	(31.26)
Four year impact: % change	-2%	4%	1%	11%	9%	-13%	-13%	-9%
Adjusted R-squared	.029	.050	.044	.208	.460	.021	.113	0.069
Number of observations (clusters)	48212 (40)	48303 (40)	48891 (40)	5055 (40)	6117 (40)	48891 (40)	48891 (40)	48094 (40)
Panel B. By Wealth Class								
Program impact on near-poor after 4 years	-24.27	1.72	51.97	.040**	.046*	-24.81	-35.45	-26.77
	(21.74)	(2.24)	(44.60)	(0.02)	(0.03)	(32.72)	(45.02)	(22.64)
Program impact on middle classes after 4 years	28.16	1.85	-30.41	.052*	.020	-20.36	-1.38	-14.16
	(30.88)	(3.37)	(46.01)	(0.03)	(0.03)	(14.22)	(12.08)	(49.43)
Program impact on upper classes after 4 years	-30.03	-1.23	-40.23	-	-	-	-	-63.05
	(72.65)	(6.03)	(54.23)	-	-	-	-	(69.50)
Four year impact on near-poor: % change	-16%	14%	18%	10%	9%	-13%	-14%	-8%
Four year impact on middle classes: % change	6%	7%	-7%	14%	4%	-28%	-2%	-5%
Four year impact on upper classes: % change	-3%	-3%	-8%	-	-	-	-	-18%
Adjusted R-squared	.213	.094	.089	.207	.462	.063	.150	.081
Number of observations (clusters)	48212 (40)	48303 (40)	48891 (40)	5055 (40)	6117 (40)	48891 (40)	48891 (40)	48094 (40)

Notes: *** (**) (*) indicates significance at the 1% (5%) (10%) level. The sample comprises all ineligible households who are present in the three survey waves. Panel A reports indirect treatment effect (ITE) estimates based on a difference-in-difference specification estimated using OLS in the whole sample. This regresses the outcome of interest for household h in village v in survey wave t on a constant, a dummy for whether the household resides in a treated village, dummies for the two follow-up survey waves (two and four years post-intervention), the interaction between the treatment assignment dummy and each survey wave dummy, and a set of strata (sub-district) fixed effects. The coefficients shown are those on the treatment-survey wave interaction terms four years post-intervention. Standard errors are clustered by BRAC branch area. Panel B reports the corresponding coefficients from a specification that allows treatment, survey waves and their interactions to vary according to social class. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Table VIII: Indirect Treatment Effects on Consumption, Household and Financial Assets of Non-Eligible Households

DiD ITE: Program impact after 4 years

Sample: Non-Eligible Households

Standard Errors in Parentheses, Clustered by BRAC Branch Area

	Poverty and Consumption			Financial Assets			Productive Assets	
	(1) Below Poverty Line	(2) Consumption Expenditure (per adult equivalent)	(3) Value of Household Assets	(4) Household Cash Savings	(5) Household Receives Loans	(6) Household Gives Loans	(7) Value of Land owned	(8) Value of Other Business Assets
Panel A. Pooled								
Program impact after 4 years	-.011	-.46	33.22	3.69	-0.002	-.013	626.14	63.55**
	(.05)	(29.90)	(28.60)	(6.03)	(0.04)	(.01)	(1182.80)	(29.35)
Four year impact: % change	-2%	-0.1%	8%	8%	-0.4%	-28%	5%	23%
Adjusted R-squared	.041	.038	.017	.007	.055	.029	.024	.018
Number of observations (clusters)	46046	45440	48200	48217	48891	48891	48201 (40)	48201 (40)
Panel B. By Wealth Class								
Program impact on near-poor after 4 years	-.015	5.31	11.13	2.52	0.007	-.003	-32.18	29.35**
	(.04)	(24.31)	(17.51)	(4.05)	(0.05)	(.01)	(282.52)	(14.43)
Program impact on middle classes after 4 years	-.030	11.17	53.72*	5.54	-0.003	-.024	51.16	97.23***
	(.05)	(36.57)	(31.80)	(8.06)	(0.04)	(.02)	(1425.75)	(34.01)
Program impact on upper classes after 4 years	.011	-27.06	55.03	6.34	-0.054	-.031	-566.68	63.75
	(.05)	(47.32)	(101.20)	(21.97)	(0.04)	(.02)	(3775.98)	(118.79)
Four year impact on near-poor: % change	-3%	1%	8%	11%	2%	-9%	-3%	34%
Four year impact on middle classes: % change	-6%	2%	12%	10%	-0.6%	-27%	0.3%	34%
Four year impact on upper classes: % change	4%	-3%	4%	6%	-9%	-30%	-1%	6%
Adjusted R-squared	.100	.156	.304	.066	.079	.046	.366	.204
Number of observations (clusters)	46046 (40)	45440 (40)	48200 (40)	48217 (40)	48891 (40)	48891(40)	48201 (40)	48201 (40)

Notes: *** (**) (*) indicates significance at the 1% (5%) (10%) level. The sample comprises all ineligible households who are present in the three survey waves. Panel A reports indirect treatment effect (ITE) estimates based on a difference-in-difference specification estimated using OLS in the whole sample. This regresses the outcome of interest for household h in village v in survey wave t on a constant, a dummy for whether the household resides in a treated village, dummies for the two follow-up survey waves (two and four years post-intervention), the interaction between the treatment assignment dummy and each survey wave dummy, and a set of strata (sub-district) fixed effects. The coefficients shown are those on the treatment-survey wave interaction terms four years post-intervention. Standard errors are clustered by BRAC branch area. Panel B reports the corresponding coefficients from a specification that allows treatment, survey waves and their interactions to vary according to social class. In Column 1, the poverty line threshold used is $1.25 per person per day, as measured in 2007 prices. In Column 2, consumption expenditure is defined as total household consumption expenditure over the previous year divided by adult equivalents in the household. The adult equivalence scale gives weight .5 to each child younger than 10. The expenditure items covered are: food (both purchased and produced), fuel, cosmetics, entertainment, transportation, utilities, clothing, footwear, utensils, textiles, dowries, education, charity and legal expenses. In Column 3, household assets include jewelry, sarees, radio, television, mobile phones, furniture, etc. In Column 4, household cash savings refer to value of savings held at home, at any bank, at any MFI and with saving guards. In Column 8, business assets include pumps, livestock sheds, trees, rickshaws and others. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Table IX: Cost-Benefit Analysis

Panel A. External parameters		
	Cost per household at year 0	1121.34
	Cost per household discounted at year 4	1363.00
	Social discount rate = 5%	
Panel B. Estimated Consumption Benefits		
1	Change in household consumption expenditure year 1	61
2	Change in household consumption expenditure year 2	106
3	Change in household consumption expenditure year 3	237
4	Change in household consumption expenditure year 4	345
5	NPV Change in household consumption expenditure from year 5 for 20 years	3581
6	Change in household assets year 4	40
7	Total benefits (1+2+3+4+5+6)	4369
8	**Benefits/cost ratio (assuming benefits last 20 years from transfer date)**	**3.21**
	Sensitivity to different discount rates/time horizons	
	Social discount rate = 10%	*2.50*
	Benefits last 10 years from transfer date	*1.86*
	Benefits last 5 years from transfer date	*0.82*
9	**IRR (assuming benefits last 20 years from transfer date)**	**0.22**
	Sensitivity to different outside options/time horizons	
	Wage jobs available all year at $.34 per hour	0.16
	Benefits last 10 years from transfer date	0.17
	Benefits last 5 years from transfer date	-0.01
Panel C. Estimated Asset Benefits		
10	Change in productive assets year 4	1030.50
11	Change in financial assets year 4	85.10
12	**Increase in assets /asset cost**	**1.85**

Notes: Household consumption includes: food (both purchased and produced), fuel, cosmetics, entertainment, transportation, utilities, clothing, footwear, utensils, textiles, dowries, education, charity and legal expenses. Productive assets include livestock, land, agricultural equipment and other machinery used for production. Financial assets equal the value of savings (held at home, at any bank, at any MFI and with saving guards) plus loans owed to the HH minus loans the HHs owes to others. The IRR is based on estimated non-durable consumption gains, assuming that these last for the expected productive life of the beneficiaries, set at 20 years. When we assume that wage jobs are always available at the observed agricultural wage we deduct the estimated increase in labor supply (206 hours) multiplied by wage the from consumption benefits. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

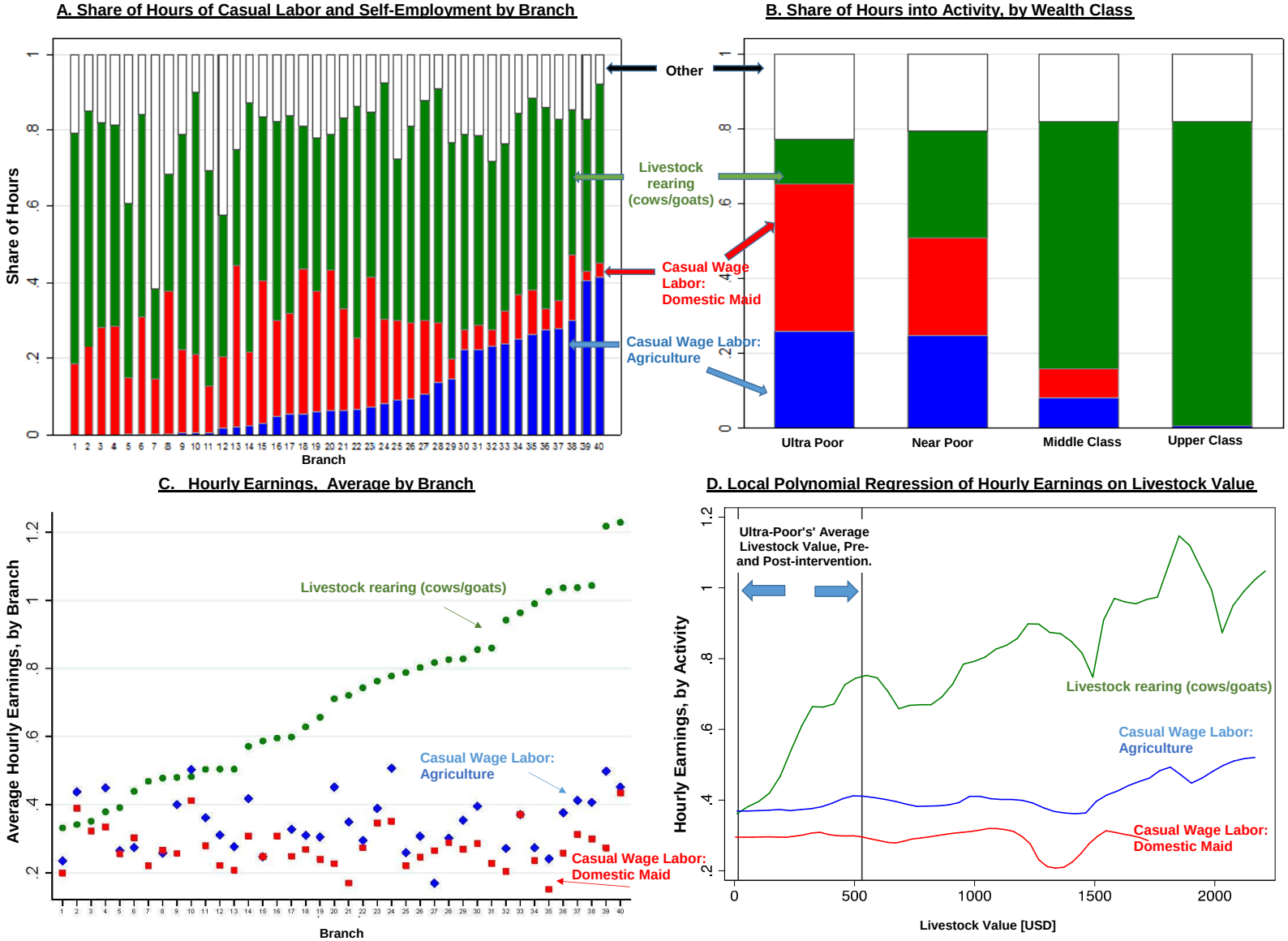
Table X. Seven-Year Treatment Effects on Consumption, Savings and Assets of Ultra-Poor Households

DiD ITT Estimates: Household Level Outcomes
Sample: Ultra Poor Households
Standard Errors in Parentheses, Clustered by BRAC Branch Area

	(1) Household Consumption Expenditure	(2) Value of Household Assets	(3) Household Cash Savings	(4) Value of Productive Assets
Program impact after 2 years	112.2*	6.860	54.69***	606.4***
	(62.62)	(7.262)	(4.601)	(92.05)
Program impact after 4 years	358.2***	39.65***	53.22***	972.6***
	(63.54)	(9.075)	(4.007)	(158.3)
Program impact after 7 years				
adjustment for program effect on the late treated:				
1. none	281.0**	27.09*	21.43***	662.0***
	(119.6)	(13.93)	(3.935)	(214.4)
2. = median 3Y treatment effect on the early treated	327.2***	30.36**	31.84***	782.8***
	(119.5)	(13.94)	(4.054)	(214.6)
3. = 75th ptile 3Y treatment effect on the early treated	338.9***	33.52**	36.34***	830.9***
	(119.6)	(13.96)	(4.222)	(215.0)
4. = 25th ptile 3Y treatment effect on the early treated	315.5**	28.36**	27.90***	751.1***
	(119.5)	(13.93)	(3.962)	(214.5)
P-values:				
Four year impact = Seven year impact (row 1)	.563	.354	.000	.052
Four year impact = Seven year impact (row 2)	.816	.496	.000	.233
Four year impact = Seven year impact (row 3)	.749	.409	.000	.374
Four year impact = Seven year impact (row 4)	.885	.652	.001	.164
Observations (clusters)	25176 (40)	26437 (40)	26437 (40)	26435 (40)

Notes: *** (**) (*) indicates significance at the 1% (5%) (10%) level. Intent-to-treat (ITT) estimates are reported based on a difference-in-difference specification estimated using OLS. All outcomes are measured at the household level, using data on ultra poor households with an eligible woman resident in them at baseline. We estimate 7 year treatment effects under different assumptions on the effect of the program on the late treated households in control villages. Row 1 assumes that the program effect on the late treated is zero and includes all control households regardless of whether late treated or not. Rows 2 (3,4) assume that that the program effect on the late treated is equal to the median (75th, 25th percentile) effect on the early treated at the same point in time. In these rows we adjust the seven year outcomes of the late treated by adding the estimated treatment effect of the early treated. In all specifications we regress the outcome of interest for household h in village v in survey wave t on a constant, a dummy for whether the household resides in a treated village, dummies for the three follow-up survey waves (two, four, and seven years post-intervention), the interaction between the treatment assignment dummy and each survey wave dummy, and a set of strata (sub-district) fixed effects. The coefficients shown are those on the treatment-survey wave interaction terms. Standard errors are clustered by BRAC branch area. In Column 1, consumption expenditure is defined as total household expenditure over the previous year. The expenditure items covered are: food (both purchased and produced), fuel, cosmetics, entertainment, transportation, utilities, clothing, footwear, utensils, textiles, dowries, education, charity and legal expenses. In Column 2, household assets include jewelry, sarees, radio, television, mobile phones, furniture, etc. In Column 3, household cash savings refer to value of savings held at home, at any bank, at any MFI and with saving guards. In Column 4 productive assets include livestock, land and business assets. We report the mean of each dependent variable as measured at baseline in treated villages. In all Columns we report the p-value on the null hypothesis that the four and seven year ITT impacts are equal. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

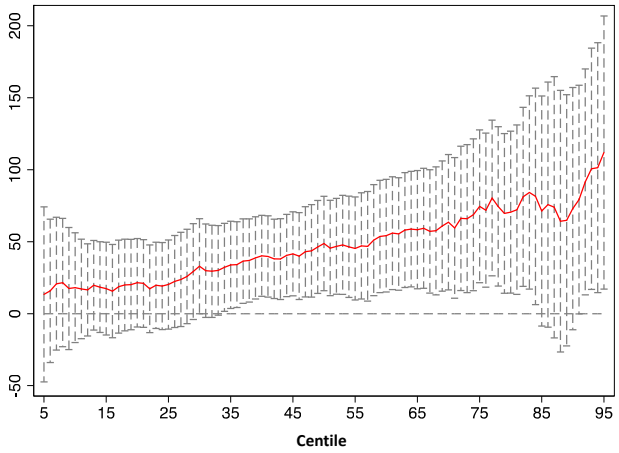
Figure I: Features of Rural Labor Markets for Women



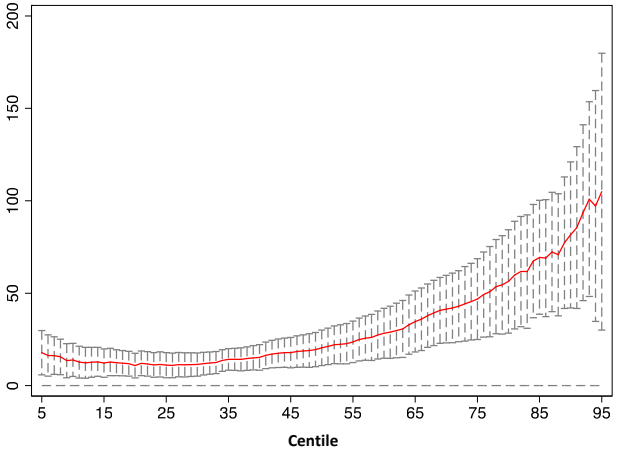
Notes: All figures are derived using the baseline household survey and present statistics on the three main occupations: domestic maid (red), agricultural labor (blue), livestock rearing (green), and other (white). Panel A shows the share of hours devoted to the different occupations by BRAC branch, ordered by the share of hours devoted to casual labor in agriculture. Panel B shows the share of hours devoted to the different labor market activities by wealth class. Panel C shows the hourly returns to the different occupations by BRAC branch, ordered by returns to livestock rearing. For each activity, earnings per hour are calculated as total earnings from that activity divided by total hours worked in the activity, both defined over the year prior to the baseline survey for individuals who had positive hours and non-missing earnings in that activity. Panel D graphs local polynomial regressions of the hourly returns to activities by the value of livestock owned. The vertical lines correspond to the average value of livestock owned by the ultra-poor pre- and post-intervention. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Figure II: Four-Year Quantile Treatment Effects

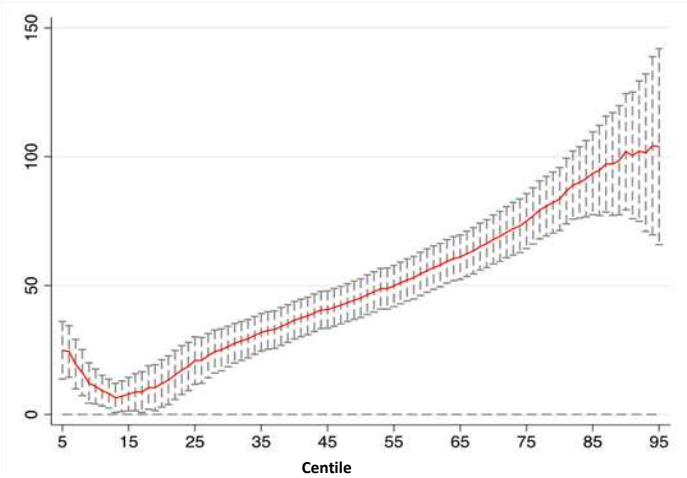
A. Consumption Expenditure (per adult equivalent)



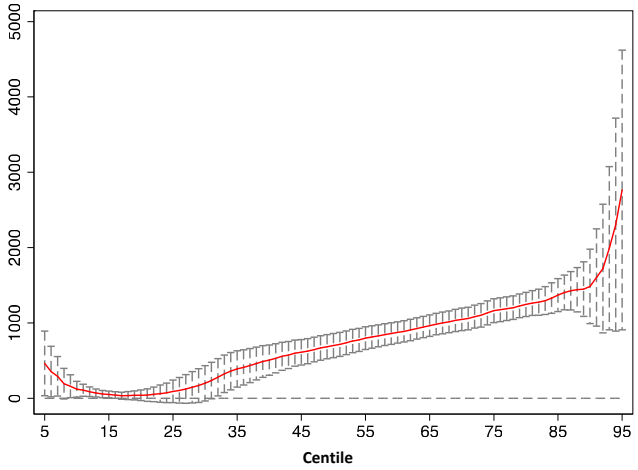
B. Value of Household Assets



C. Savings

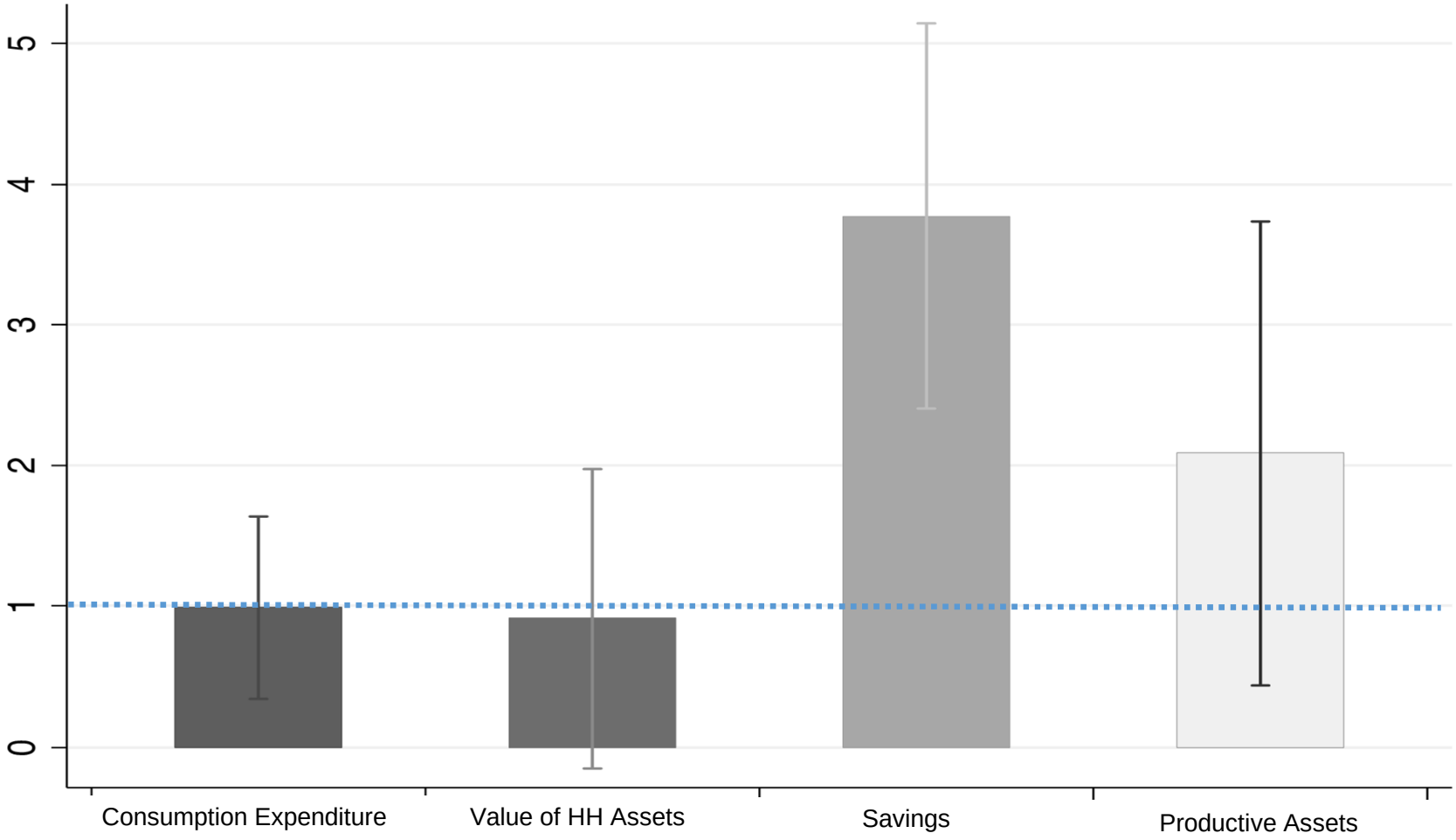


D. Value of Productive Assets



Notes: Quantile treatment effect (QTE) estimates of the differences in outcomes between four-year follow-up and baseline are presented in each panel. Each specification controls for randomization strata. Bootstrapped 95% confidence intervals (using 500 replications) are based on standard errors clustered by BRAC branch. Consumption expenditure includes: food (both purchased and produced), fuel, cosmetics, entertainment, transportation, utilities, clothing, footwear, utensils, textiles, dowries, education, charity and legal expenses. Household assets include jewelry, sarees, radio, television, mobile phones, furniture, etc. Productive assets include livestock, land, agricultural equipment and other machinery used for production. Savings equal the total value of savings held at home, at any bank, at any MFI and with saving guards. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Figure III: Four-Year Treatment Effects on the Gap between Ultra-Poor and Near Poor



Notes: Estimates are based on a triple-difference specification between baseline and year four, treatment and control, eligibles and non eligibles. estimated using OLS with standard errors clustered at the branch level. All outcomes are divided by the average difference between eligibles and non eligibles in treatment at baseline, thus a measured impacts of one indicates that gap has closed. Consumption expenditure includes: food (both purchased and produced), fuel, cosmetics, entertainment, transportation, utilities, clothing, footwear, utensils, textiles, dowries, education, charity and legal expenses. Household assets include jewelry, sarees, radio, television, mobile phones, furniture, etc. Productive assets include livestock, land, agricultural equipment and other machinery used for production. Savings equal the total value of savings held at home, at any bank, at any MFI and with saving guards. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

LABOR MARKETS AND POVERTY IN VILLAGE ECONOMIES

Oriana Bandiera, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, Munshi Sulaiman

ONLINE APPENDIX

Construction of Data Set for Comparison with Banerjee *et al.* (2015a)

The outcomes we replicate are indices corresponding to the ten primary outcome measures studied in Banerjee *et al.* (2015a). Each outcome is a composite index that combines outcomes for individual/household i related to outcome k, denoted Y_i^k. Following Banerjee *et al.* (2015a), we construct each index k by first defining every outcome within the relevant group of outcomes such that higher values correspond to better outcomes. We then standardize each outcome into a z-score by subtracting the control group mean and dividing by the control group standard deviation (SD) for the corresponding survey round. We then average all the z-scores and again standardize to the control group within each round. We convert all monetary values to 2014 USD PPP terms.

Following Banerjee *et al.* (2015a), we estimate the following specification:

$$Y_i^k = \alpha + \beta_1 assignment_i + Z_i^k + V_{strata} + \varepsilon_i, \tag{1}$$

where Y_i^k is the outcome k of interest for either household or adult i, $assignment_i$ is an indicator for having been randomly selected into the program, Z_i^k is the household's baseline value of the outcome variable k, V_{strata} is the vector of all variables included in stratification (i.e. subdistrict fixed effects). Standard errors are clustered at the branch level (unit of randomization).

The dependent variable in Column 1 of Table 6 is the standardized total per capita consumption per month. To ensure comparability with Banerjee *et al.* (2015a), this consumption measure differs from that used in the rest of the paper in the following ways: (i) expenditures on income-generating activities are excluded; (ii) expenditure is defined per household member (as opposed to adult-equivalent household member); (iii) monthly expenditure is used; (iv) monetary values are reported in 2014 USD PPP terms.

The dependent variable in Column 2 is a food security index. To build the food security index, Banerjee *et al.* (2015a) use five indicators: (i) everyone gets enough food every day; (ii) no adult skips meals; (iii) no one went a whole day without food; (iv) no child skipped meals; (v) everyone regularly eats two meals a day. We build the most comparable indicators we can using our survey instrument. In particular, to build a comparable measure for (i) we define a variable equal to 1 if the respondent reported that her household's status in terms of food availability was "neither deficit nor surplus" or "food surplus" and 0 if she said it was "always deficit" or "deficit sometimes". For (iii), we define a variable equal to 1 if the respondent reported that in the month preceding the survey, her household never had less than enough food to eat and 0 otherwise. Since our survey did not ask this question separately for adult versus child members, we cannot build indicators for measures (ii) and (iv). Finally, for (v) we define a dummy variable equal to 1 if the respondent reported that her household could afford to have two meals per day most of the time during the last year and 0 otherwise.

The dependent variable in Column 3 is an asset index, based on the total value of productive and household assets. To ensure comparability with Banerjee *et al.* (2015a), we construct the measure via the following steps: (i)

calculate the median unit value for each type of asset; (ii) calculate the value of each asset in terms of goats (the numeraire asset) by dividing the unit value of each asset by the median unit value of goats; (iii) calculate total asset value by multiplying the unit value of each asset (expressed in terms of goats) by the number of each asset owned; (iv) standardize the total asset value.

The dependent variable in Column 4 is a financial inclusion index. Banerjee *et al.* (2015a) use five indicators to construct this index: (i) total amount borrowed in the last 12 months; (ii) amount borrowed from informal sources (neighbor, friend, shopkeeper, family, work place, moneylender, etc.) in the last 12 months; (iii) amount borrowed from formal sources (MFI, NGO, government) in the last 12 months; (iv) total savings at the time of the survey; (v) total amount deposited in savings during the last 12 months. We have data on all but the last indicator, so we use (i)-(iv) to construct the index.

The dependent variable in Column 5 is a standardized measure of the total time spent by the main woman of the household in productive activities on a typical day during the past year. Banerjee *et al.* (2015a) measure individual labor supply as the total minutes spent on all productive activities in the day prior to survey day. To build this measure, they convert weekly or 48 hour labor supply (depending on survey/country) to minutes per 24 hours. We collected information on annual labor supply, asking respondents for the number of days they spent during the last year on each income-generating activity and the number of hours worked during a typical working day. Using this information, we build a measure of the number of hours worked during an average day during the last year in each activity, and multiply this by 60 to get minutes per day. Banerjee *et al.* (2015a) aggregate individuals' labor supply, however many adults were surveyed. Across countries, this ranges from one to seven adults per household. We collected individual labor supply information by work activity (separating self-employment from wage-labor) only for the main female respondent and (when applicable) for the male head of the household. Thus, we report the labor supply of the female respondent (in Column 5) and the pooled value for both respondents (in Column 6) for those households that had a male respondent. As in Banerjee *et al.* (2015a), we standardize each measure using the control group's mean for each survey wave.

The dependent variable in Column 7 is an income and revenues index, as reported by the main female respondent. Banerjee *et al.* (2015a) use five variables to construct this index: (i) household livestock revenue per month; (ii) household agricultural income per month; (iii) household non-farm micro-enterprise income per month; (iv) household income from paid labor per month; (v) self-reported economic status (0/1) which is defined based on the classification of household economic status on a ladder from 0 to 10. We collected information on all except the last indicator, so we use variables (i)-(iv) to construct the index. We did not ask for total household income by activity, but we did ask for each household member's income from each income-generating activity he/she was engaged in. In order to avoid double-counting of income from household businesses, we only use the earnings of the main female respondent.

The dependent variable in Column 8 is a physical health index consisting of three variables. Banerjee *et al.*

(2015a) use three variables to measure physical health: (i) activities of daily living scores based on respondents' self-reported ability (on a binary (0/1) scale) to perform the following physical tasks: lift a heavy object, work all day in the field, walk a certain distance without getting tired; these are averaged to give the daily living score; (ii) no adult member missed any work days due to illness; (iii) self-perception of physical health on a scale from 1-5 based on asking respondents about their satisfaction with their physical health. We build corresponding variables using our data as follows: for (i), we use information on whether the respondent would be able to perform five physical activities on a scale from 1 to 3 where 1=easily, 2=with trouble and 3=unable. We rescale these so that higher values imply better health status and take the average to build the index. The physical activities we asked about were: walking one mile at a normal speed, carrying a heavy load (e.g. 10 seer rice) for 20 yards, drawing a pail of water from a tube-well, standing up from a sitting position on the floor without help, using a ladder to climb to a storage place at least 5 feet high. For (ii), we use information on whether any household member had an illness in the 15 days before the survey, and if so whether this "interfered with any income-generating activity". For (iii), we use data on the respondent's self-perception of her current health on a 3-point scale (1=good, 2=average, 3=bad), scaled such that higher values imply better health status.

The dependent variable in Column 9 is a mental health index consisting of two variables. Banerjee *et al.* (2015a) use three indicators to construct this index: (i) self-reported happiness (in some countries based on satisfaction with mental health on a scale from 1 to 10, in others based on satisfaction with life on a scale from 1 to 5); (ii) a stress index (for which specific indicators vary across countries) which combines z-scores based on the number of times in the past week that the respondent felt sad, cried a lot, did not feel like eating, did not feel like working, had restless sleep, or whether the respondent had a period of worry lasting at least 30 days in a year; (iii) a dummy variable equal to 1 if the respondent did not experience worry that lasted for more than one month. For (i), we used a variable describing how the respondent considers her life in terms of happiness on a scale from 1 to 3 where 3=very happy, 2=happy and 1=unhappy. We do not have corresponding variables that can be used to construct indicator (ii). For (iii), we asked respondents whether they experienced any mental anxiety that "hampered their daily activities" during the past month (giving a binary variable). We rescaled indicators such that higher values imply better outcomes and then constructed the aggregate index using the same steps as Banerjee *et al.* (2015a).

The dependent variable in Column 10 is a political awareness index. Banerjee *et al.* (2015a) use four indicators to construct this index: (i) whether the respondent voted in the last election; (ii) whether the respondent was a member of a political party; (iii) whether the respondent attended a village meeting in the last year; (iv) whether the respondent has spoken with village leaders about village concerns in the last year. We do not have corresponding measures in our data. Instead, we build a measure based on information on whether or not the respondent knows politicians at different levels and the lowest legal age for voting. We have five binary variables, each equal to 1 if the respondent can correctly name the president, the prime minister, a parliamentary member from her area and a ward member, and whether she knows the lowest legal age for voting.

The dependent variable in Column 11 is a women's empowerment index. Banerjee *et al.* (2015a) use five indicators to construct this index: (i) female respondent has major say on food decisions; (ii) female respondent has major say on education decisions; (iii) female respondent has major say on health decisions (personal and family); (iv) female respondent has major say on home improvement decisions; (v) female respondent has major say on how to manage household finances. In our survey, we did not ask respondents whether or not they have the "major say" in the household, but we did ask whether they could influence household decision-making under various scenarios. In particular, we asked about the following scenarios: (1) If your household is going to buy land and you think it is not the right time, can you influence them to do it later?; (2) If your household is going to repair your house and you think it is not the right time, can you influence them to do it later?; (3) If your household is going to borrow from a source that you think is not the right source, can you influence them to change their decision?; (4) If you wish to be involved in a new activity would you need to gain permission from other household members?; (5) If you think your husband should take up a new activity, can you influence him to do that?; (6) If you think your son should take up a new activity, can you influence him to do that?; (7) If you think your daughter should take up a new activity, can you influence her to do that?; (8) Can you influence the decision on how far your son proceeds with his studies?; (9) Can you influence the decision on how far your daughter proceeds with her studies?; (10) If your husband is not spending as much on your children's clothing as you would like him to, can you make him spend more?; (11) If someone in the household is ill, would you be able to influence the decision about whether to seek outside treatment or not?; (12) If someone in the household is ill, would you be able to influence the decision about what type of treatment to seek?; We use the responses to these questions (all measured as binary (0/1) variables) to construct the women's empowerment index.

Table A.I: Features of Rural Labor Markets for Women

Village Level Statistics, Measured Pre-Intervention

Means, standard deviation in parentheses

	Casual Wage Labor		Self Employment		
	(1) Agriculture	(2) Domestic Maid	(3) Livestock Rearing [Cows, Goats]	(4) t-test [Col 1 = Col 3]	(5) t-test [Col 2 = Col 3]
Days per year	127	167	334	[.000]	[.000]
	(65.9)	(89.5)	(41.2)		
Hours per day	7.62	7.04	1.83	[.000]	[.000]
	(1.15)	(1.74)	(.771)		
Hourly earnings [USD]	.344	.268	.719	[.000]	[.000]
	(.102)	(.109)	(.779)		

Notes: All statistics are constructed at the village level, using baseline data from both treatment and control villages. The number of villages is 1309. In Column 3, livestock comprises cows and/or goats. To reduce sensitivity to outliers, the hours per day and hourly earnings variables are computed by first taking the median value for each activity in a village, and then averaging these across all villages. Columns 4 and 5 report p-values on a t-test of the equality of some of these outcomes between the two forms of casual wage labor (agriculture and domestic maid work) and livestock rearing. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Table A.II: Balance

Sample: Ultra-Poor Women and their Households

	(1) Treated Villages	(2) Control Villages	(3) t-test [Treatment=Control]	(4) Normalized Differences
A. Labor Market Outcomes				
Hours devoted to livestock rearing (cows/goats)	115	129	.584	-.036
	(258)	(275)		
Earnings from livestock rearing	7.85	8.90	.654	-.013
	(53.2)	(60.4)		
Hours devoted to agricultural labor	269	237	.740	.042
	(537)	(539)		
Hourly wage in agricultural labor	.330	.360	.431	-.195
	(.103)	(.114)		
Hours devoted to domestic maid	325	479	.013	-.152
	(651)	(774)		
Hourly wage in maid services	.256	.261	.823	-.028
	(.107)	(.113)		
Earnings from casual labor	164	191	.340	-.085
	(218)	(239)		
Total earnings	241	289	.172	-.117
	(275)	(300)		
Total days worked in the past year	247	259	.327	-.060
	(141)	(130)		
Average standardized difference (p-value)			*.207*	
B. Poverty, Expenditures and Financial Wealth				
Below the $1.25 a day poverty line [yes=1]	.556	.584	.524	-.040
	(.400)	(.398)		
Consumption expenditure, per adult equivalent	629	613	.501	.047
	(246)	(236)		
Value of household assets	36	37	.829	-.011
	(48)	(63)		
Household savings	6.2	9.2	.071	-.059
	(28)	(43)		
Household receives loans	.20	.18	.441	-.044
	(.40)	(.38)		
Household gives loans	.011	.014	.356	-.022
	(.10)	(.12)		
Average standardized difference (p-value)			*.849*	
C. Productive Assets				
Cows value	36	30	.575	.023
	(176)	(166)		
Goats value	6.5	8.5	.261	-.050
	(25)	(31)		
Household rents in land [yes=1]	.058	.061	.875	-.007
	(.235)	(.239)		
Household owns land [yes=1]	.068	.062	.738	.017
	(.252)	(.241)		
Value of land owned	175	238	.390	-.027
	(997)	(2190)		
Value of other business assets	23	23	.991	-.0004
	(79)	(101)		
Average standardized difference (p-value)			*.863*	

Notes: All data refers to the baseline survey. Columns 1 and 2 report means with standard deviation in parentheses, based on ultra-poor women/households in treatment and control villages respectively. Column 3 reports the p-value of the test of equal means, allowing for standard errors to be clustered by BRAC Branch. Column 4 reports normalized differences computed as the difference in means in treatment and control villages divided by the square root of the sum of the variances. The poverty line threshold used is $1.25 per person per day, as measured in 2007 prices. Household savings refer to value of savings held at home, at any bank, at any MFI and with saving guards. The household livestock value includes the value of cows and goats. Business assets include pumps, livestock sheds, trees, rickshaws and others. Consumption expenditure is defined as total household consumption expenditure over the previous year divided by adult equivalents in the household. The adult equivalence scale gives weight .5 to each child younger than 10. The expenditure items covered are: food, fuel, cosmetics, entertainment, transportation, utilities, clothing, footwear, utensils, textiles, dowries, education, charity and legal expenses. At the foot of each Panel we report the p-value associated with the average standardized difference, defined as in Kling *et al.* (2007). All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Table A.III: Attrition

OLS Estimates
Sample: All Ultra-Poor Women at Baseline
Dependent Variable=1 if Respondent is Surveyed in All Three Waves
Standard Errors Clustered by Village in Parentheses

	(1)	(2)	(3)
Treated village	.0139	.014	.012
	(.011)	(.011)	(.014)
Hours devoted to agriculture day labor		.000	-.000
		(.001)	(.001)
Hours devoted to domestic maid		-.000	-.000
		(.001)	(.001)
Hours devoted to livestock rearing		.009***	.008***
		(.002)	(.002)
Hours devoted to agriculture day labor x Treated village			.000
			(.001)
Hours devoted to domestic maid x Treated village			-.000
			(.001)
Hours devoted to livestock rearing x Treated village			.002
			(.003)
Subdistrict Fixed Effects	Yes	Yes	Yes
Attrition Rate: Baseline to Endline		14.6%	
Adjusted R-squared	.003	.007	.007
Observations (number of ultra-poor women)	7953	7953	7953

Notes: *** (**) (*) indicates significance at the 1% (5%) (10%) level. OLS estimates are reported based on the sample of ultra-poor women observed at baseline. The dependent variable is a dummy variable equal to one if the woman is observed in all three survey waves (baseline, two-year midline, four-year endline), and zero otherwise. All specifications control for the level effect of the treatment and sub-district fixed effects. Standard errors are clustered by village.

Table A.IV: Allocation of Labor of Household Members of the Ultra-Poor

DiD ITT 4-year Estimates
Standard Errors in Parentheses, Clustered by BRAC Branch Area
Each Coefficient Corresponds to a Separate Regression

	(1) Husbands	(2) Adult members (16 and older)	(3) Children (15 and younger)
Capital Intensive Activities			
Hours devoted to rearing livestock	59.0***	54.6***	41.3**
	(18.7)	(9.14)	(15.4)
Hours devoted to land cultivation	16.1	21.7***	7.67**
	(16.4)	(5.21)	(3.06)
Hours devoted to rickshaw driving	-38.5	.483	-11.0**
	(30.2)	(9.82)	(4.76)
Casual Wage Labor Activities			
Hours devoted to agriculture day labor	-85.4	6.22	11.1
	(123)	(24.8)	(12.9)
Hours devoted to domestic servant	-	-4.06	-3.53
	-	(10.8)	(22.1)
Total Hours Worked and Schooling			
Total hours worked	-18.1	116**	60.3
	(177)	(46.3)	(39.3)
Share enrolled in school	-	-	-.008
			(.025)
Number of households		6732	
Observations (clusters)	11731 (40)	12043 (40)	11407 (40)

Notes: *** (**) (*) indicates significance at the 1% (5%) (10%) level. Intent-to-treat estimates are reported based on a difference-in-difference specification estimated using OLS. This regresses the outcome of interest for individual i in village v in survey wave t on a constant, a dummy for whether the individual resides in a treated village, dummies for the two follow-up survey waves (two and four years post-intervention), the interaction between the treatment assignment dummy and each survey wave dummy, and a set of strata (sub-district) fixed effects. The coefficients shown are those on the treatment-survey wave interaction terms. Each coefficient corresponds to a separate regression. The sample includes individuals in the same household as an ultra-poor woman. Standard errors are clustered by BRAC branch area. All outcomes are measured at the individual level, and defined for the year prior to survey date. Livestock rearing refers to working with cows/goats.

Table A.V.A: Treatment Effects, ANCOVA Specification Separately for Each Survey Wave

ANCOVA ITT Estimates - Separate regressions for each survey wave
Sample: Ultra Poor Households
Standard Errors in Parentheses, Clustered by BRAC Branch Area

Panel A: Labor Supply	**Livestock**		**Agriculture**		**Maid**		**All Activities**	
	(1) Hours	**(2) Days**	**(3) Hours**	**(4) Days**	**(5) Hours**	**(6) Days**	**(7) Hours**	**(8) Days**
Program impact after 2 years	482.783***	196.672***	-23.821	-2.109	-157.780***	-21.349***	222.417***	61.122***
	(23.75)	(6.17)	(16.61)	(2.34)	(17.48)	(2.87)	(43.90)	(3.22)
Program impact after 4 years	420.271***	170.737***	-25.764	-2.770	-247.235***	-33.474***	84.428	54.149***
	(27.89)	(8.25)	(23.42)	(2.99)	(28.73)	(3.61)	(51.34)	(4.96)

Panel B: Earnings	**Livestock**	**Agriculture**		**Maid**		**All Activities**
	(9) Earnings	**(10) Wage**	**(11) Earnings**	**(12) Wage**	**(13) Earnings**	**(14) Earnings**
Program impact after 2 years	80.031***	0.011	-4.010	0.026***	-31.783***	25.105
	(10.30)	(0.02)	(7.80)	(0.01)	(5.12)	(16.23)
Program impact after 4 years	120.465***	0.041***	3.380	0.073***	-54.863***	58.178***
	(8.24)	(0.01)	(9.86)	(0.01)	(6.71)	(18.47)

Panel C: Consumption, Household and Financial Assets

	Poverty and Consumption			**Financial Assets**		
	(1) Below Poverty Line	**(2) Consumption Expenditure (per adult equivalent)**	**(3) Value of Household Assets**	**(4) Household Cash Savings**	**(5) Household Receives Loans**	**(6) Household Gives Loans**
Program impact after 2 years	-0.089***	53.905***	6.715**	50.230***	0.107***	0.039***
	(0.01)	(8.21)	(3.16)	(3.77)	(0.02)	(0.01)
Program impact after 4 years	-0.109***	80.333***	39.218***	50.926***	0.095***	0.047***
	(0.02)	(9.74)	(4.97)	(3.24)	(0.02)	(0.01)

Panel D: Productive Assets

	(1) Value of Cows	**(2) Value of Goats**	**(3) Rents Land**	**(4) Owns Land**	**(5) Value of Land owned**	**(6) Value of Other Business Assets**
Program impact after 2 years	493.088***	27.659***	0.071***	0.007	30.166	26.657***
	(13.24)	(2.47)	(0.01)	(0.01)	(52.69)	(4.91)
Program impact after 4 years	564.599***	19.795***	0.115***	0.028***	303.431***	64.064***
	(30.05)	(2.85)	(0.01)	(0.01)	(71.57)	(7.61)

Notes: *** (**) (*) indicates significance at the 1% (5%) (10%) level. Intent-to-treat (ITT) estimates from separate yearly ANCOVA specifications. These regress the outcome of interest on a constant, a dummy for whether the household resides in a treated village, the value of the outcome of interest at baseline, and a set of strata (sub-district) fixed effects. They do so separately using either the outcome data from the 2009 survey wave, or from the 2011 survey wave. The coefficients shown are those on the treatment dummy. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Table A.V.B: Treatment Effects, ANCOVA Specification Pooling All Survey Waves

ANCOVA ITT Estimates - Pooled

Sample: Ultra Poor Households

Standard Errors in Parentheses, Clustered by BRAC Branch Area

Panel A: Labor Supply	**Livestock**		**Agriculture**		**Maid**		**All Activities**	
	(1) Hours	**(2) Days**	**(3) Hours**	**(4) Days**	**(5) Hours**	**(6) Days**	**(7) Hours**	**(8) Days**
Program impact after 2 years	487.909***	200.584***	-21.072	-1.658	-169.546***	-22.788***	225.206***	63.365***
	(29.12)	(8.33)	(23.10)	(3.14)	(22.37)	(3.48)	(51.62)	(4.63)
Program impact after 4 years	415.145***	166.826***	-28.514	-3.222	-235.469***	-32.035***	81.638	51.905***
	(31.51)	(9.47)	(27.62)	(3.56)	(33.49)	(4.35)	(56.39)	(5.72)

Panel B: Earnings	**Livestock**	**Agriculture**		**Maid**		**All Activities**
	(9) Earnings	**(10) Wage**	**(11) Earnings**	**(12) Wage**	**(13) Earnings**	**(14) Earnings**
Program impact after 2 years	82.636***	0.021	-2.638	0.033***	-37.821***	27.125
	(12.24)	(0.02)	(9.71)	(0.01)	(6.42)	(20.97)
Program impact after 4 years	118.001***	0.030**	2.428	0.062***	-53.472***	55.011**
	(10.75)	(0.01)	(11.28)	(0.02)	(8.43)	(22.02)

Panel C: Consumption, Household and Financial Assets

	Poverty and Consumption			**Financial Assets**		
	(1) Below Poverty Line	**(2) Consumption Expenditure (per adult equivalent)**	**(3) Value of Household Assets**	**(4) Household Cash Savings**	**(5) Household Receives Loans**	**(6) Household Gives Loans**
Program impact after 2 years	-0.083***	51.099***	6.453	51.191***	0.107***	0.041***
	(0.02)	(11.07)	(4.14)	(4.09)	(0.03)	(0.01)
Program impact after 4 years	-0.116***	85.399***	39.479***	49.783***	0.093***	0.050***
	(0.02)	(11.70)	(5.58)	(3.77)	(0.03)	(0.01)

Panel D: Productive Assets

	(1) Value of Cows	**(2) Value of Goats**	**(3) Rents Land**	**(4) Owns Land**	**(5) Value of Land owned**	**(6) Value of Other Business Assets**
Program impact after 2 years	501.148***	27.622***	0.073***	0.007	24.210	24.905***
	(20.84)	(2.57)	(0.02)	(0.01)	(61.84)	(6.90)
Program impact after 4 years	556.569***	19.909***	0.113***	0.028***	309.408***	65.824***
	(30.80)	(2.92)	(0.01)	(0.01)	(83.37)	(8.15)

Notes: *** (**) (*) indicates significance at the 1% (5%) (10%) level. Intent-to-treat (ITT) estimates from ANCOVA specifications. These regress the outcome of interest on a dummy for whether the household resides in a treated village interacted with each of the two period dummies, the period dummies, the value of the outcome of interest at baseline interacted with each of the two period dummies, and a set of strata (sub-district) fixed effects. The coefficients shown are those on the treatment dummy interacted with each of the two period dummies. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Table A.VI: Inference Robustness

Sample: Ultra-Poor Women and Their Households

		(1) P-value based on clustered standard errors as in text	(2) P-value based on clustered standard errors adjusted for degrees of freedom (Young 2016)	(3) P-value based on wild-bootstrap clustered standard errors (Cameron et al 2008)
A. Labor Market Outcomes (Table 4)				
Livestock: Hours	**Program impact after 2 years**	0.000	0.000	0.000
	Program impact after 4 years	0.000	0.000	0.000
Livestock: Days	**Program impact after 2 years**	0.000	0.000	0.000
	Program impact after 4 years	0.000	0.000	0.000
Agriculture: Hours	**Program impact after 2 years**	0.430	0.438	0.394
	Program impact after 4 years	0.286	0.295	0.324
Agriculture: Days	**Program impact after 2 years**	0.617	0.623	0.579
	Program impact after 4 years	0.386	0.394	0.434
Maid: Hours	**Program impact after 2 years**	0.189	0.198	0.194
	Program impact after 4 years	0.013	0.015	0.018
Maid: Days	**Program impact after 2 years**	0.159	0.167	0.170
	Program impact after 4 years	0.006	0.008	0.008
All Three Activities: Hours	**Program impact after 2 years**	0.000	0.000	0.000
	Program impact after 4 years	0.008	0.009	0.012
All Three Activities: Days	**Program impact after 2 years**	0.000	0.000	0.000
	Program impact after 4 years	0.000	0.000	0.000
Livestock: Earnings	**Program impact after 2 years**	0.000	0.000	0.000
	Program impact after 4 years	0.000	0.000	0.000
Agriculture: Wage	**Program impact after 2 years**	0.187	0.216	0.236
	Program impact after 4 years	0.031	0.047	0.074
Agriculture: Earnings	**Program impact after 2 years**	0.479	0.487	0.436
	Program impact after 4 years	0.782	0.786	0.835
Maid: wage	**Program impact after 2 years**	0.130	0.142	0.188
	Program impact after 4 years	0.000	0.001	0.000
Maid: earnings	**Program impact after 2 years**	0.318	0.327	0.324
	Program impact after 4 years	0.035	0.040	0.030
All three activities: earning	**Program impact after 2 years**	0.046	0.051	0.052
	Program impact after 4 years	0.004	0.005	0.004
B. Poverty, Expenditures and Financial Wealth (Table 5)				
Below the $1.25 a day poverty line [yes=1]	**Program impact after 2 years**	0.283	0.292	0.324
	Program impact after 4 years	0.032	0.036	0.030
Consumption expenditure, per adult equivalent	**Program impact after 2 years**	0.241	0.250	0.270
	Program impact after 4 years	0.005	0.006	0.006
Value of household assets	**Program impact after 2 years**	0.351	0.360	0.384
	Program impact after 4 years	0.000	0.000	0.000
Household savings	**Program impact after 2 years**	0.000	0.000	0.000
	Program impact after 4 years	0.000	0.000	0.000
Household receives loans	**Program impact after 2 years**	0.000	0.001	0.000
	Program impact after 4 years	0.002	0.002	0.002
Household gives loans	**Program impact after 2 years**	0.003	0.003	0.000
	Program impact after 4 years	0.000	0.000	0.000
C. Productive Assets				
Cows value	**Program impact after 2 years**	0.000	0.000	0.000
	Program impact after 4 years	0.000	0.000	0.000
Goats value	**Program impact after 2 years**	0.000	0.000	0.000
	Program impact after 4 years	0.000	0.000	0.000
Household rents in land [yes=1]	**Program impact after 2 years**	0.002	0.002	0.004
	Program impact after 4 years	0.000	0.000	0.000
Household owns land [yes=1]	**Program impact after 2 years**	0.646	0.652	0.655
	Program impact after 4 years	0.053	0.058	0.070
Value of land owned	**Program impact after 2 years**	0.600	0.606	0.617
	Program impact after 4 years	0.017	0.020	0.020
Value of other business assets	**Program impact after 2 years**	0.001	0.002	0.000
	Program impact after 4 years	0.000	0.000	0.000

Notes: We report alternative p-values for the 4 year treatment effects estimated in Tables 4 , 5 and 6. Column 1 reports the p-value based on clustered standard errors as reported in the main text. Column 2 reports the p-value based on clustered standard errors with the degrees of freedom adjustment as in Young (2016). Column 3 reports the p-value based on clustered standard errors computed using the wild bootstrap method of Cameron *et al*. (2008).

Table A.VII: Quantile Cost-Benefit Analysis

Panel A. External parameters						
Cost per household at year 0	1121.34	Social discount rate = 5%				
Cost per household discounted at year 4	1363.00					
Panel B. Estimated Consumption Benefits	***q10***	***q25***	***q50***	***q75***	***q90***
1 Change in household consumption expenditure year 1	-3	30	44	107	194
2 Change in household consumption expenditure year 2	-5	51	76	184	335
3 Change in household consumption expenditure year 3	62	126	157	312	540
4 Change in household consumption expenditure year 4	123	188	223	410	694
5 NPV Change in household consumption expenditure year 5 and beyond-forever	1279	1955	2313	4256	7199
NPV Change in household consumption expenditure from year 5 for 10 years	625	956	1131	2081	3521
NPV Change in household consumption expenditure year 5 and 6	117	179	212	390	661
NPV Change in household consumption expenditure from year 5 for 20 years discount 10%	937	1433	1695	3119	5276
6 Change in household assets year 4	14	11	20	47	81
7 Total benefits (1+2+3+4+5+6)	2537	4174	4977	9260	15715
	2084	3277	3899	7243	12288
	1576	2500	2981	5553	9428
	2396	3753	4463	8281	14043
	1472.58	2331.87	2788.79	5209.01	8848.84
8 **Benefits/cost ratio**	***1.08***	***1.73***	***2.08***	***3.90***	***6.63***
Sensitivity to different discount rates/time horizons					
Social discount rate = 10%	0.83	1.35	1.62	3.07	5.22
Benefits last 10 years from transfer date	*0.60*	*1.00*	*1.21*	*2.30*	*3.94*
Benefits last 5 years from transfer date	*0.23*	*0.43*	*0.54*	*1.06*	*1.84*
9 IRR					
Sensitivity to different outside options/time horizons					
Wage jobs available all year at $.34 per hour	-0.03	0.05	0.08	0.21	0.35
Benefits last 10 years from transfer date	-0.01	0.07	0.10	0.24	0.39
Benefits last 5 years from transfer date	-0.26	-0.15	-0.11	0.07	0.26
Panel C. Estimated Asset Benefits					
10 Change in productive assets year 4	120.42	92.20	699.89	1162.95	1485.05
11 Change in financial assets year 4	53.95	9.53	30.93	61.00	112.08
12 **Increase in assets /asset cost**	**0.30**	**0.18**	**1.20**	**2.03**	**2.68**

Notes: Household consumption includes: food (both purchased and produced), fuel, cosmetics, entertainment, transportation, utilities, clothing, footwear, utensils, textiles, dowries, education, charity and legal expenses. Productive assets include livestock, land, agricultural equipment and other machinery used for production. Financial assets equal the value of savings (held at home, at any bank, at any MFI and with saving guards) plus loans owed to the HH minus loans the HHs owes to others. The IRR is based on estimated non-durable consumption gains, assuming that these last for the expected productive life of the beneficiaries, set at 20 years. When we assume that wage jobs are always available at the observed agricultural wage we deduct the estimated increase in labor supply (219 hours) multiplied by wage the from consumption benefits. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Table A.VIII: Seven-Year Treatment Effects on Consumption, Savings and Assets of Ultra-Poor Households-ANCOVA

ANCOVA ITT Estimates: Household Level Outcomes
Sample: Ultra Poor Households
Standard Errors in Parentheses, Clustered by BRAC Branch Area

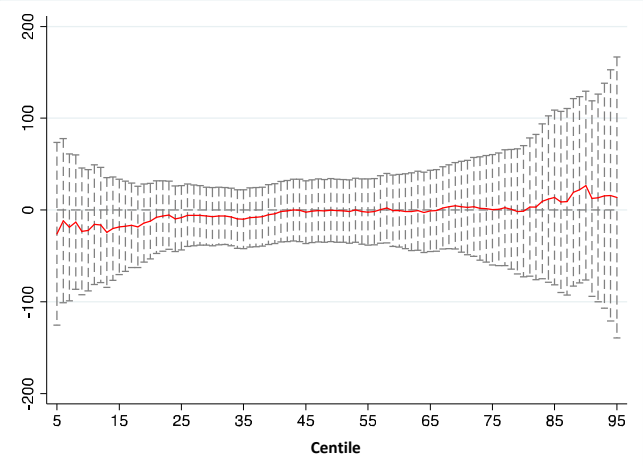
PANEL A: Each survey wave separately

	(1) Household Consumption Expenditure	(2) Value of Household Assets	(3) Household Cash Savings	(4) Value of Productive Assets
Program impact after 2 years	219.5***	6.031	45.79***	574.7***
	(24.24)	(3.664)	(5.607)	(106.9)
Program impact after 4 years	417.4***	40.18***	53.38***	1013.4***
	(24.01)	(5.353)	(4.006)	(149.6)
Program impact after 7 years	353.7***	29.25***	22.31***	708.3***
	(32.16)	(7.987)	(4.330)	(134.6)
PANEL B: All survey waves pooled				
Program impact after 2 years	198.7***	7.264	48.34***	556.2***
	(46.47)	(5.165)	(6.009)	(142.8)
Program impact after 4 years	441.7***	41.55***	51.35***	1017.0***
	(57.32)	(7.258)	(5.040)	(158.6)
Program impact after 7 years	347.9***	26.52**	21.68***	721.3***
	(83.85)	(9.839)	(4.496)	(184.8)

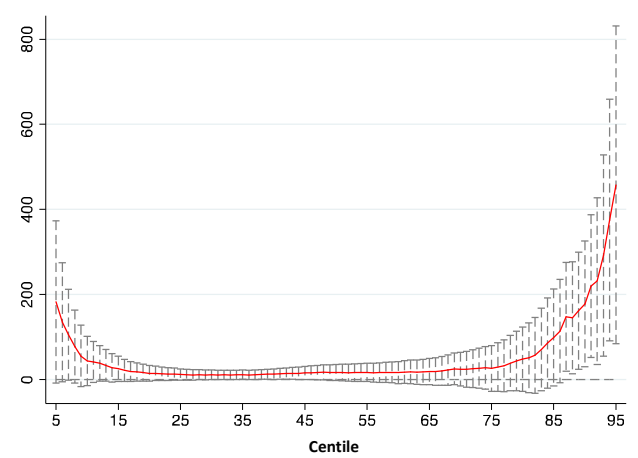
Notes: *** (**) (*) indicates significance at the 1% (5%) (10%) level. Intent-to-treat (ITT) estimates from ANCOVA specifications. In Panel A we restrict the sample to each of the three periods (2,4 and 7 years) and regress the outcome of interest on a constant, a dummy for whether the household resides in a treated village, the value of the outcome of interest at baseline, and a set of strata (sub-district) fixed effects. In Panel B we regress the outcome of interest on a dummy for whether the household resides in a treated village interacted with each of the two period dummies, the period dummies, the value of the outcome of interest at baseline interacted with each of the two period dummies, and a set of strata (sub-district) fixed effects. The coefficients shown are those on the treatment dummy. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.

Figure A.I: Four-Year Quantile Treatment Effects on Non-Eligible Households

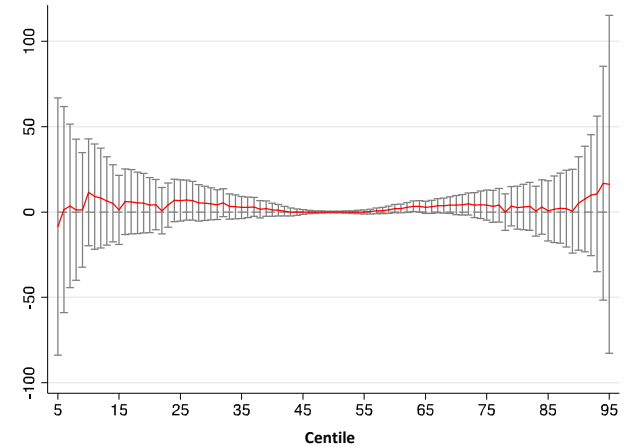
A. Consumption Expenditure (per adult equivalent)



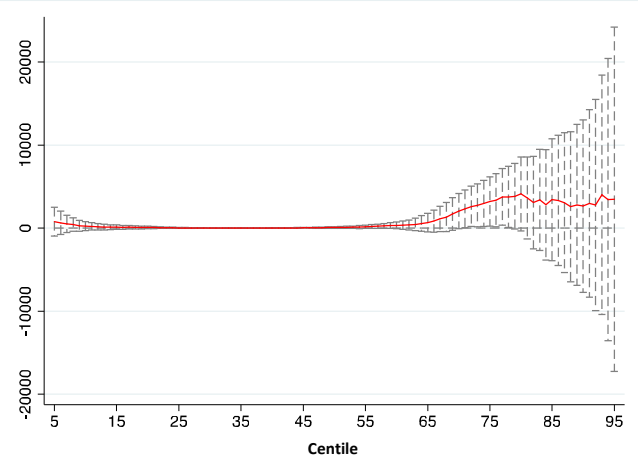
B. Value of Household Assets



C. Savings



D. Value of Productive Assets



Notes: Quantile treatment effect (QTE) estimates of the differences in outcomes between four-year follow-up and baseline are presented in each panel. Each specification controls for randomization strata. Bootstrapped 95% confidence intervals (using 500 replications) are based on standard errors clustered by BRAC branch. Consumption expenditure includes: food (both purchased and produced), fuel, cosmetics, entertainment, transportation, utilities, clothing, footwear, utensils, textiles, dowries, education, charity and legal expenses. Household assets include jewelry, sarees, radio, television, mobile phones, furniture, etc. Productive assets include livestock, land, agricultural equipment and other machinery used for production. Savings equal the total value of savings held at home, at any bank, at any MFI and with saving guards. All monetary amounts are PPP-adjusted USD terms, set at 2007 prices and deflated using CPI published by Bangladesh Bank. In 2007, 1USD=18.46TK PPP.