Does Poverty Change Labor Supply? Evidence from Multiple Income Effects and 115,579 Bags*

Abhijit Banerjee†
Dean Karlan‡
Hannah Luk-Zilberman§
Christopher Udry¶

September 28, 2021

Abstract

The income elasticity of labor supply is a central parameter of many economic models. We test the response of labor supply and effort to exogenous changes in income using data from a randomized evaluation of a multi-faceted grant program in northern Ghana combined with a bag-making operation that we implemented. We find strong evidence of a *positive* "income effect" on labor supply. We argue that simple models with either labor or capital market frictions cannot explain the results, whereas a model that allows for positive physiological or psychological productivity effects from higher income fits with our findings.

Keywords: poverty, labor supply, income elasticity
JEL Classifications: H31, J22, O12

*Approval from the Yale University Human Subjects Committee, IRB 0705002656, 1002006308, 1006007026, and 1011007628; and from the Innovations for Poverty Action Human Subjects Committee, IRB Protocol 19.08January-002, 09December-003, 59.10June-002, and 10November-003.494. Thanks to the Ford Foundation, and 3ie for funding. Thanks to Nathan Barker, Caton Brewster, Abubakari Bukari, David Bullon Patton, Sébastien Fontenay, Angela Garcia, Yann Guy, Samantha Horn, Sana Khan, Hideto Koizumi, Matthew Lowes, Elizabeth Naah, Michael Polansky, Elana Safran, Sneha Stephen, Rachel Strohm, and Stefan Vedder for outstanding research assistance and project management, and in particular Bram Thuysbaert for collaboration. The authors would like to thank the leadership and staff at Presbyterian Agricultural Services (PAS) for their partnership. Thanks to Frank DeGiovanni of the Ford Foundation, Syed Hashemi of BRAC University, and Aude de Montesquiou and Alexia Latortue of CGAP for their support and encouragement of the research. No authors have any real or apparent conflicts of interest, except Karlan is on the Board of Directors of Innovations for Poverty Action, which participated in oversight of the implementation. All data and code will be available upon publication at the IPA Dataverse (doi pending).

†MIT, CEPR, NBER and Jameel Poverty Action Lab (J-PAL): banerjee@mit.edu
‡Northwestern University, CEPR, NBER, IPA, and J-PAL: karlan@northwestern.edu
§Hebrew University of Jerusalem: hannah.lukzilberman@mail.huji.ac.il
¶Northwestern University, CEPR, NBER, and J-PAL: christopher.udry@northwestern.edu

1 Introduction

The income elasticity of labor supply is one of the central parameters of economic models. Under the standard assumption that consumption and work are not strong complements, it is easy to derive the prediction that any increase in income will reduce labor supply. This has important implications for the design of social support policies, because, for example, such a reduction in labor supply would limit the net income gains.

The basic argument for why we should expect this negative labor supply response is well-known. Making the standard assumptions that the utility from consumption is $u(c)$, the disutility of labor supply is $v(l)$ and the relation between consumption and labor supply is $c = f(l) + t$, where f is income and is some increasing concave function of labor supply and t is a transfer, we immediately get a first order condition

$$u'(f(l) + t)f'(l) = v'(l)$$

from which it follows that any increase in t will reduce the marginal utility of income and therefore labor supply. Of course a number of important assumptions are being made here. First, as pointed out by Benjamin (1992) many years ago, we need that t does not directly raise the marginal product of labor. In other words, we cannot have $f(l, t)$ with $f_{lt}(l, t) > 0$. As Benjamin (1992) also points out, this is typically ruled out by either the assumption of perfect capital markets (in which case t should not enter $f(l, t)$) or by the assumption that household labor and market labor are perfect substitutes at the margin (in which case $f_l(l, t) = w$, where w is the market wage). However neither of these assumptions seem particularly plausible especially in the context of low income families in developing countries (LaFave et al., 2020). Therefore a transfer may actually directly raise the marginal product of labor, thus making this kind of *investment productivity effect* quite relevant.

A second reason why the expected income effect may be absent or even go the other way is that consumption (or income) and labor supply may be complements. In other words it is possible that the disutility of effort takes the form $v(l, c)$ with $v_{lc}(l, c) < 0$, at least for the very poor. The idea that a mechanical nutrition-productivity relationship generates complementarity between consumption and work lies at the heart of the earliest models of a poverty trap (Leibenstein, 1957; Dasgupta and Ray, 1986). In these models, a better-fed worker provides more effort. More generally, higher consumption may result in better health, which in turn may reduce the disutility of work (Strauss and Thomas 1998). We call this a *physiological productivity effect*.

More recently, psychological models of poverty traps have made a similar case for why, at low levels of psychological well-being, positive income shocks may boost labor supply—what we will call a *psychological productivity effect*. One body of work, summarized in Mullainathan and Shafir (2013), has suggested that people living under any form of scarcity exhibit "tunnel vision," focusing so intently to allocate their scarce resources that they neglect other

margins and make sub-optimal decisions as a result. There is evidence of such psychological effects of financial strain on productivity (Kaur et al., 2019; Fink et al., 2018). Another body of work, summarized in Haushofer and Fehr (2014), investigates the effects of poverty on risk-taking and time-discounting. Positive income shocks have been shown to reduce risk aversion (Tanaka et al., 2010), and negative income shocks have been shown to increase present-biased behavior (Haushofer and Fehr, 2019). Relatedly, poverty has been shown to cause negative affect and stress (Haushofer and Shapiro, 2016), which have in turn been shown to influence risk-taking and time-discounting (Kandasamy et al., 2014). A final body of work emphasizes the relationship between poverty and aspirations. Several theoretical papers explore how both individuals (Dalton et al., 2016) and economies (Genicot and Ray, 2017) can get trapped in poverty when aspirations and outcomes are jointly determined. There is also emerging evidence that economic success can indeed affect aspirations (Lybbert and Wydick, 2017), and that interventions that shift aspirations can lead to greater investment (Bernard et al., 2020). Taken together, these findings support the possibility that additional income might have a psychological productivity effect that plays off against the conventional income effect.

Consistent with this set of theories, the evidence from a number of recent field experiments suggests that the income effect on labor supply is often non-negative. Using data from a number of cash transfer programs around the world that had a built in randomized controlled trial, Banerjee et al. (2017) shows that cash transfers to low income households have no effect on labor supply, either at the intensive margin or at the extensive margin. Banerjee et al. (2015) and Bandiera et al. (2017) report on a six-country study and a one-country study, respectively, of the Graduation program, a multi-faceted program built around an asset transfer to very poor households, and both find that the intervention led to higher incomes and labor supply. The positive impact persisted when last measured, three years after the intervention began (and one year after it ended). While the graduation program potentially changes the entire life circumstances of the beneficiaries there is also evidence that much more temporary positive shocks also have a positive effect on labor supply of low-income households: in a field experiment with piece-rate workers in India, Kaur et al. (2019) finds higher levels of productivity on days when the workers are cash-rich, and argues this is indicative of improved cognition and focus due to lower levels of financial stress.

This evidence, while suggestive, has three important potential limitations. First there is concern with the measurement of labor supply. For example, if much of the labor supply response is in the form of reduced (unmeasured) effort on a job, it could be that the person is doing less and eventually will be fired, but we do not observe this long-term outcome.

Second, the Graduation experiments were not designed to shed light on mechanisms involved. They involve both the transfer of a productive asset to households who are very plausibly credit constrained (so an increase in t, which may shift the $f(l,t)$ function) and encouragement and hand-holding for the program recipients, intended to shift their $v(.)$ functions. The physiological or psychological effect of extra income may be an important part of what is

happening, but there is no way to tell from the data.

Finally, the productivity effects of temporary income shocks studied in Kaur et al. (2019) are indeed likely to be psychological: the immediacy of the impact makes it unlikely to be due to investment in a productive asset (unlike in Gertler et al. (2006)). However the (very) temporary nature of these interventions and the response to them raises questions about the durability of the impact.

With this context in mind, we make two contributions by building on our study of the Ghana Graduation program, (also called "Graduating the Ultra Poor", and here onward referred to as "GUP"), which was part of the set of studies reported on in Banerjee et al. (2015)). First, we provide better measurement of labor supply and still find a non-negative income effect on labor supply. Second, we provide evidence that what we call the psychological or physiological productivity effect is driving the observed departure from the traditional income effect, in a context of a large and persistent income shock.

A key to both contributions is a novel measurement exercise involving a bag-making operation. GUP treatment and control villages were randomly chosen to have bag production units. Those who were invited to work in these units were offered piece rate contracts to produce bags, and all inputs were provided. The number of bags they produced as well as their quality was carefully graded and the piece rate depended on quality, so we have a reliable measure of how much effort individuals put into bag-making. Each bag-making unit was also randomly assigned to produce either simple or more complex bags to test whether there is a differential effect for complex bags.

For those in the bags production sub-groups, the comparison of GUP and control households tells us that GUP increases participation in bags, bags production, and earnings from bags. These effects are individually statistically significant, and the q values after adjusting for multiple hypothesis testing are 0.10, 0.10 and 0.17. Moreover there is an increase in productivity in bags, which is not statistically significant overall but highly statistically and economically significant for complex bags, with GUP individuals spending a third less time per bag. If we interpret productivity as measuring effort per minute spent on producing bags, it represents an alternative dimension of labor supply. This increase in productivity cannot be attributed to complementary capital investment in bag making because all inputs are provided by the researchers.

Of course it could be that this increased effort reflects the fact that the household has cut back on other activities and therefore has extra time and and energy available to it. In terms of reported hours worked, we estimate that GUP-bags households supply only about two percent fewer hours to all forms of productive labor (producing bags, farming, business operations, animal production and home labor—including childcare, cleaning, cooking, collecting firewood/water, and shopping) than do control-bags households, and this difference is nowhere near statistically significant at conventional levels.

However, as mentioned above, labor hours are not all of labor supply—there could also be a reduction in effort in activities other than bag-making. To get at a measure of effort we start from the fact that there is essentially no wage labor in our context. Individuals either work on their own farms or run their

own businesses. In both of these cases the household is the residual claimant and the effective labor supply, including any differences in effort, should be reflected in the income from the activity, which we measure, in addition to the reported labor time on the activity. We do not see any evidence that GUP households are supplying less total effort in either of these occupations. Relative to control bags households, the average GUP-bags households spend 21 minutes fewer per day on farming but produce about 10% more.[1] Moreover we see little evidence that they are making labor-saving investments, which would allow earnings from agriculture to go up even when effort has gone down. Expenditure on herbicides (which is labor-saving) is marginally higher among GUP-bags households, but expenditure on (labor-using) fertilizer is also higher. There is no difference in hired labor between GUP-bags and control-bags households. GUP-bags households, relative to control-bags households, spend 33% more time on their businesses ($p=0.06$) and appear to earn more than twice as much, though this effect is not statistically significant ($p=0.16$). We do not have data on whether the business adopted labor saving innovations but given how small the businesses are, the absence of hired labor, and the simple technologies (shea butter production, petty trading), this seems unlikely. Finally, GUP households report spending a bit more time on livestock after two years, which makes sense given that most of them have additional goats to care for and revenue goes up a bit (but not significantly). Both these effects are small relative to the effects on farming and business revenue.[2]

It is striking that GUP-bags households supply more overall effort because these households earn substantially (and statistically significantly) more than the control bags households. Summing up across all the sources of earnings plus any cash transfers, during bag-making GUP-bags households earned \$20.9 more per month than control-bags households ($p<0.01$). This implies that they earned more than double the control-bags monthly earnings of \$17.9 while spending roughly the same amount of time on productive labor and delivering more effort to bags production and no less effort elsewhere. In other words there is prima facie evidence of a *positive* GUP effect on labor supply.

Turning to our second question, we argue that this is evidence for what we have called physiological or psychological productivity effects, rather than a investment productivity effect. This is because bag-making offered no scope additional investment by the households: all capital was provided by us, the researchers.

This still leaves the question of whether the GUP effect is merely an income effect. The issue is, as mentioned already, that the GUP program was multi-faceted and had a number of components that went beyond just providing an asset. However the experimental design included two arms that allows us to address this possibility.

One concern is that GUP had both a savings collection component and a pure

[1] For one key sub-treatment this effect is much larger and statistically significant–see below.

[2] Given that both time use in and earnings from livestock-rearing are small relative to other activities, and since we do not have data on livestock revenue for the period of the bags program, we will mostly ignore livestock activities in the remainder of the paper.

encouragement component. For the first 24 months of the program households were visited weekly by NGO staff who encouraged them to believe that they can and should aim higher. They also collected savings (for the GUP-savings households) and deposited it in the bank for them. A psychological productivity effect could have resulted from this encouragement, rather than from the extra income. The potential to save in a bank account, on the other hand, could have had an positive incentive effect on both labor supply and earnings.

To address these concerns we make use of the fact that the GUP households received weekly unconditional cash support during each lean season. For the bags households during the bags program, the amount of this unconditional support was randomly varied between $1.3 and $3.9 per week. Unlike the basic GUP effect, this is a pure income shock to the household, since all the GUP-bags households received the exact same set of interventions.[3]

The labor supply effects of this rather substantial pure income shock (which amounts to a 34% increase in total income in the lean season) align with our previous findings. The high UCT households are, unsurprisingly, richer than the low UCT households, but work roughly the same amount per day. They work slightly less on the farm and slightly more at their business, but neither difference is close to being statistically significant. The value of their harvest is higher while business earnings are similar. The high UCT households use more (labor-using) fertilizer and less (labor-saving) herbicide, and hire less outside labor than do low UCT households. In other words, there is no evidence of the high UCT households working less or putting less effort into non-bag-making activities. The high UCT households do participate more in bags production, produce more bags, earn more from bags and take fewer minutes per bag, though none of these differences are statistically significant. In other words, the effect of the pure income shock loosely parallels the GUP effect, suggesting that extra income through the psychological/physiological productivity effect plays an important role.

To rule out the potential role of savings collection, we make use of two additional features of our experiment. First, the experiment had a "savings only" arm which we called “SOUP” (Savings Only Ultra-Poor program), where households just received a weekly visit from a nonprofit organization to collect deposits into a bank account with a partnering financial institution. The bags intervention was then cross-cut with the SOUP treatment allowing us to test whether the observed complementarity between GUP and labor supply also shows up with SOUP. While the SOUP intervention by itself has an effect on household consumption comparable to the GUP effect, and also raises household earnings (though substantially—but not significantly—less than GUP), we find no evidence of a positive productivity effect on bag-making coming from SOUP. In fact the point estimates of SOUP on bag productivity are strongly negative (while the GUP effect is positive) and the difference between them is close to being statistically significant (p=0.13). Second, some GUP households randomly

[3]This is not strictly true, since some GUP-bags households did not get savings collection, but this was cross-randomized with the unconditional support and therefore does not pose a problem.

did not receive the savings collection service. These GUP-no-savings households also make more bags than control households (though the effect is smaller and marginally significant), and do not spend any less time on productive labor.

These findings confirm that the complementarity between GUP and bag productivity is not likely to be the result of the GUP savings component. They also suggest that the productivity effect is less likely to be driven by the physiological effect resulting from improved nutrition, because the SOUP intervention had a similar effect on consumption as the GUP intervention, but not the same effect on bags productivity.[4]

This paper contributes to a large literature on labor markets in developing countries (e.g. Lewis (1954); Rosenzweig (1988); Foster and Rosenzweig (1996); Goldberg (2016); Guiteras and Jack (2018)). It relates to work on the relationship between credit constraints and labor supply (e.g. Kochar (1999); Rose (2001); Jayachandran (2006); Fink et al. (2018)), and most directly builds on work understanding the effects of positive income shocks, through transfers or other mechanisms, on labor supply (e.g. Baird et al. (2018); Kaur et al. (2019)). Finally, it contributes to the large body of work that attempts to unpack the determinants of effort (e.g. Breza et al. (2018); Brune (2016); Brune et al. (2019); Kaur et al. (2015)), including the potential importance of psychological well-being and its link to income (Mani et al., 2013; Shah et al., 2012).

We start by presenting the overall experimental design in Section 2. Section 3 then presents the model that we use to interpret the results. Section 4 describes our data and empirical methods. Section 5 presents the results on the impact of GUP on standard economic outcomes and labor supply outside of bag-making. Section 6 presents the evidence from the bag-making program, first comparing GUP and SOUP, and then high and low unconditional transfers. We then use these results and the theory in Section 3 to try to make the case for a strong complementarity between consumption/income and labor supply/effort. We conclude in Section 7.

2 Experimental Design

We partnered with Presbyterian Agricultural Services (PAS), a local NGO in northern Ghana with prior experience doing extension work and promotion of savings groups, including a prior randomized controlled trial with Innovations for Poverty Action (Karlan et al., 2017). PAS field agents engaged in the direct field implementation, while Innovations for Poverty Action coordinated the implementation with senior management of PAS. PAS first identified poor communities in poor regions in northern Ghana, and in each identified community, staff members then facilitated a Participatory Wealth Ranking (PWR) in which members of the community ranked households by economic status. Finally, PAS staff members returned for a verification of the households judged to be

[4]It does not rule out the possibility that the GUP effect was at least partly the result of anticipated future consumption, since the households may have reason to think that GUP will have a more durable effect on household well-being than SOUP.

the poorest. In order to be eligible, households were required to have a female between the ages of 18-24, who we will call the female household head, to be the direct recipient of all treatment activities and the primary respondent for surveys.

We begin by describing the randomized design of the Graduation program in Ghana, and then move on to explain the sub-treatments within the bag-making exercise.

2.1 GUP and SOUP Treatment Designs

Table 1 Panel A shows the assignment of households and villages to GUP, SOUP and control, and the cross-cutting bags measurement village assignments. Each village was assigned GUP, SOUP, or control, and then within each treatment village, half of sample households actually received the treatment intervention, and half served as control households within treatment villages. Thus there is a two-level randomization: at the village level to assign the treatment arm, and then at the household level within village to assign treatment or control status to specific households.

In GUP villages, 51% of sample households were assigned to the GUP treatment, and the remainder were assigned to the GUP control group. The GUP program included six components: (1) the transfer of a productive asset; (2) skills training for the management of the asset, (3) life skills training and mentorship, via weekly household visits over two years, (4) a weekly cash stipend for consumption support, worth between \$6 and \$9 PPP depending on family size, during each lean season, (5) some basic health services and health education, and (6) access to a savings account at a local bank and deposit collection. The first component, the productive asset transfer, was provided at the beginning of the program, and households were permitted to choose a package of productive assets from a set list. Most households chose a package that included four goats.[5] The skills training, in which participants learned how to take care of the asset (e.g., when to vaccinate goats), took place at the start of the program, and then also as part of weekly household visits by the PAS field officer. The household visits also provided the backbone for delivering components three through six. The third component, a "hand-holding" or life-skills component, provided nudges to help the household focus on building productive assets to generate positive change in long-term outcomes, and more generally, to set aspirations and plans for coping with current problems and improving the future. The consumption support was explicitly intended to help this process in the short-run, by helping to absorb short-run shocks that could lead to households consuming the transferred assets. The fifth component, health, included basic education on health and hygiene as well as enrollment in the national health insurance scheme (about \$2 per month).

The sixth component of GUP, access to a savings account and weekly deposit collection, was the entirety of the SOUP intervention. In SOUP villages, 59%

[5] Other assets included hens, pigs, and inputs for the production of shea butter, maize, and sorghum.

of sample households were assigned to the SOUP treatment. These households received a visit from the field officer to collect savings, but did not receive any other components of the program.[6] The remaining households in SOUP villages were assigned to the SOUP control group.

For half of the households assigned to GUP, the weekly collection of savings deposits was randomly withheld. For the GUP households who did receive deposit collection services, the treatment is equivalent to the combination of GUP and SOUP.[7]

2.2 Bag-making

We designed an employment program offering wages for the production of cloth bags, and implemented it such that it cross-cut the three GUP treatment groups (GUP, SOUP, and control). Half of the villages (120) were then randomly selected to receive the Bags Program, as shown in Table 1 Panel A. In GUP and SOUP villages selected to receive the Bags program, all sample households assigned to GUP or SOUP were invited to participate. In control villages selected to receive the employment program, half of sample households were invited to participate. This amounts to 1098 households: 397 control, 313 GUP, and 388 SOUP. The employment program was implemented during the lean season and lasted for six months

Table 1 Panel B presents the details of two sub-treatments within the bags measurement exercise. First, we varied the complexity of the bag at the village level. Of the 120 villages, 60 were assigned to produce a simple bag, and 60 were assigned to produce a complex bag. The main difference between the complex and simple bag was that while the simple bag has basic "running stitches" on the hem and the strap, the complex bag alternates one "running stitch" with four "chain stitches," a slightly more complex stitch in a pattern that requires counting. Importantly, because of the difficulty of this pattern, it was harder to meet quality standards (discussed below).

Second, we varied the amount of unconditional consumption support, in the form of a cash transfer, received by GUP-bags households. This was varied at the village level, and was either USD 1.31 or USD 3.92. Since GUP-bags households also received earnings from bags, this was designed to be about half as much as what GUP-no-bags households received (between USD 6 and 9 depending on household size).[8]

[6] Among households assigned to SOUP, there was an additional sub-treatment: half received savings accounts and deposit collection without a match ("SOUP without match") and half received savings accounts and deposit collection with a 50% match ("SOUP-match"). Specifically, for every GHC 1 deposited, households in this group received a matching contribution of GHC 0.50. At the onset of the program, there was a maximum match of GHC 1.50 GHC per week (for a GHC 3 deposit) but this cap was eventually removed.

[7] We find no evidence that the presence or absence of savings collection makes a difference to the impact of GUP on consumption or income; see Banerjee et al. (2020).

[8] We also varied the wage at the village level over time. Every four weeks, villages were assigned a different baseline wage: USD 0.40 or USD 0.91. Women were informed of the payment per bag they would be receiving for bags made in a given week at the start of that

Immediately prior to the start of the program, field agents invited female household heads to participate in a community-level training, which lasted for four days. Bag production began immediately afterwards. During production, GUP, SOUP, and Control Field Agents visited each community on a weekly basis. At each visit, they collected new bags, distributed replacement fabric (according to the number of bags collected), and paid wages for bags submitted two weeks prior. Households could submit a maximum of ten bags per week. In the two weeks between when bags were collected and when wages were paid, quality checks were carried out by program facilitators. There are 18 quality standards for simple bags, and 25 quality standards for complex bags. Bags were assigned one point for meeting the quality standards at the "excellent" level, half a point for "satisfactory," and zero points for "unsatisfactory." At the end of the quality check, the final quality score was calculated and the bag was classified as high, mid, or low quality.

Wages were paid with a two-week lag. Each week, program facilitators informed households of the composition of high, mid, and low quality bags submitted two weeks prior, and distributed payment accordingly. Baseline wages were either USD 0.40 or USD 0.91. Bags judged to be high quality earned the baseline wage plus USD 0.13, bags judged to be mid quality earned the baseline wage, and bags judged to be low quality bags earned the baseline wage minus USD 0.13. The wage was not affected by whether the bag was simple or complex. Every four weeks, bags program facilitators returned to communities to give feedback and remedial training.

3 A model of labor supply

We provide a simple model to clarify the interplay between the GUP, SOUP and bags programs and to explain how we use them together to conclude that there was a non-negative (and perhaps positive) income effect on labor supply, drive by a psychological or physiological productivity effect. Result 3 introduces a potential confound: it shows that the savings component of the program has the potential to increase both labor supply and income through a pure incentive effect. Result 1 shows that absent this savings incentive effect, the finding that GUP recipients are better off and still work more requires that there is either an investment productivity effect or a psychological or physiological labor supply effect. Finally Result 2 is the key result for isolating the psychological or physiological labor supply effect. It shows that absent the savings incentive effect,

week. Bags produced in week 1 of a given wage rotation would be collected at the end of week 1 and inspected for quality over the course of weeks 2 and 3. Payment for the bags produced during week 1 would be given to the producer at the end of week 3. For this reason, there is a lag between when the wage rate changes and when individuals start receiving higher wages, and the data show that responsiveness to wage rate changes is lagged by three weeks (see Appendix Table 1 Panel A). Since the pattern and timing of responses to wage changes indicates that there were delays between the announcement of wage changes and full understanding of their effect, we do not focus on these results in the main part of the paper. Our estimates of wage elasticities are shown in Appendix Table 1 Panel B.

increased labor supply to the bags program, where capital inputs and the technology are exogenously fixed, has to come from a psychological or physiological labor supply effect of the GUP intervention. The other components might directly affect the marginal product of labor through technology or investment, or the marginal disutility of work through encouragement or better nutrition.

The utility from a certain income c is given by $\lambda u(\frac{c}{\lambda}))$, where λ is a shifter for the utility function. A higher λ is meant to capture the impact of the savings component of the GUP intervention, which makes it possible to spread the extra consumption over a longer future, hence raising the marginal utility of income. For results 1 and 2 we keep λ fixed. The household production function is $f(l,t)$, where the inclusion of t represents the possibility that the transfers raise the marginal product of labor. In other words we assume that $f_l(l,t) > 0$, $f_{ll}(l,t) < 0$, $f_t(l,t) \geq 0$ and $f_{lt}(l,t) \geq 0$. As noted, a necessary condition for this is that there are imperfections in both the capital market and the labor market. This framework captures the three components of the GUP program (the productive asset transfer, the skills training, and the consumption support) that have the potential to raise the marginal product of labor in this environment of incomplete markets.

The disutility of labor supply l is given by $v(l,T))$, where the inclusion of T is aimed to capture the relation between the various interventions and labor supply, via physiological or psychological channels. In other words it is possible that $T = t$, but we want to allow for possibility of interventions that shift labor supply without providing an income transfer (such as through encouragement). We assume that $v_l(l,T) > 0$, $v_{ll}(l,T) > 0$, $v_T(l,T) \geq 0$ and $v_{lT}(l,T) \leq 0$. One case where we might expect $v_T(l,T) < 0$ and $v_{lT}(l,T) < 0$, is when $T = t$, income transfers boost consumption and greater consumption raises labor supply. Another would be a coaching/encouragement treatment, where $t = 0$ but $T > 0$. Within this framework, two components of the GUP program (life skills training, and basic health services and health education) may have direct effects on the disutility of work via T. Finally we assume that $c = f(l,t)+t$. The asset transfer and consumption support components of GUP enter the budget constraint through t.

The first order condition for utility maximization is

$$u'\left(\frac{f(l,t)+t}{\lambda}\right) f_l(l,t) = v_l(l,T).$$

Suppose that $t = t(T)$ with $t'(T) > 0$. It is evident that $\frac{dl}{dT} < 0$ as long as $f_{lt}(l,t) = 0$ and $v_{lT}(l,T) = 0$. However $\frac{dl}{dT}$ can be positive if either $f_{lt}(l,t) > 0$ or $v_{lt}(l,T) < 0$. As before we call these two sources of a non-traditional income effects the investment productivity effect and the psychological/ physiological productivity effect.

Result 1: As long as λ is fixed, a necessary condition for the income effect on labor supply not to be negative is that there has to be either the investment productivity effect or the psychological/physiological productivity effect.

For our second result, we permit the household to have access to two pro-

duction technologies, so that

$$c = f^a(l^a, t) + f^b(l^b, t) + t,$$

where $f^a(.)$ represents the bag making opportunity.

The household now maximizes

$$\lambda u\left(\frac{f^a\left(l^a,t\right)+f^b\left(l^b,t\right)+t}{\lambda}\right)-v(l^a+\gamma l^b,T)$$

by choosing l^a and l^b. γ represents the relative cost of effort in the two tasks. Now suppose $f^a_{l^a t}(l^a, t) = 0$. The first order condition with respect to l^a yields

$$u'\left(\frac{c}{\lambda})\right) f^a_{l^a}(l^a, t) = v_l(l, T)$$

.

We wish to compare $l^a(T)$ with $l^a(T')$ where $t(T) > t(T')$. Now suppose $c(T) \geq c(T')$ and therefore $u'(c(T)) < u'(c(T'))$. Moreover let $l^b(T) \geq l^b(T')$. Then if it also true that $l^a(T) \geq l^a(T')$ then $l(T) \geq l(T')$. Now if $v_{lT}(l, T) = 0$, then $v_l(l(T), T) \geq v_l(l(T'), T')$. In this case the only way to satisfy the first order condition is for $f^a_{l^a t}(l^a, t) > 0$. Conversely, if $f^a_{l^a t}(l^a, t) = 0$ then it must be the case that $v_{lT}(l, T) < 0$. We summarize this as:

Result 2: As long as λ is fixed, if there is one activity where there is no investment productivity effect, and the labor supply to that activity is greater despite the fact the household is richer and is working no less, then there must be a psychological/physiological productivity effect on the disutility of labor.

The last observation is about λ. If λ goes up, say because of savings collection, the household's marginal utility of income goes up and therefore both its labor supply and its income must both go up.

Result 3: If λ goes up, the household's labor supply and its income must both go up.

4 Data and Empirical Methods

4.1 Data

The final sample was selected from the households identified as the poorest in their poor communities as described in Section 2. Participants come from three areas of Northern Ghana corresponding to three agricultural "stations" run by PAS: Tamale, Langbensi, and Sandema. We restrict all of our analysis to villages with more than 30 compounds, as for logistical reasons, we assigned all pure control villages with fewer than 30 compounds to no-bags. This leaves 93 bags villages and 72 non-bags villages.

We have three sources of data. First, we have weekly administrative data on labor supply (the number of bags submitted), the quality of each bag, and the resulting earnings. Second, we have time use surveys in which the female

household head reported how she spent her time the previous day. We administered these surveys five times monthly during the bags program, to 1051 bags households and 470 no-bags households.[9] Third, we have a series of standard and comprehensive household surveys that were part of the larger program evaluation of the Graduation program (Banerjee et al. 2015). These include a baseline survey, three shorter midline surveys, a two-year follow-up survey and a three-year follow-up survey. These surveys included questions about income, consumption, agricultural outcomes, business outcomes, and welfare. The second midline survey is used heavily, as it took place during the bags program. Midline surveys were conducted with about one third of the full sample, so for this survey, we have data on 1070 households, including 343 bags households and 727 no-bags households.

4.2 Orthogonality

Tables 2 and 3 show baseline survey data across treatment groups. We have baseline imbalance on average age, land area, monthly per capita consumption, monthly household income, and the food security index. We had intended to re-randomize, but due to a coding error, it did not happen. As a result, in every regression, we also control for the five aforementioned variables at baseline.

4.3 Method of Analysis

We use two main specifications for our three types of data: one for the analysis of individual-level outcomes measured in our two-year survey (Equation 1); and one for the analysis of individual-month level time use outcomes, or individual-week level bag-making outcomes, measured during the bags program (Equation 2). Any deviations from these specifications or additional details will be reported in table notes.

$$Y_i = \alpha + \beta T_i + \gamma Y_i^0 + W_i^{strata} + \theta^{interviewer} + \epsilon_i \tag{1}$$

$$Y_{it} = \alpha + \beta T_i + W_i^{strata} + \rho^{station*t} + \epsilon_{it} \tag{2}$$

$Y_{i(t)}$ is outcome Y for individual i at either month or week t, T_i is a treatment dummy, Y_i^0 is the baseline value of outcome Y for individual i (only used in Equation 1 since we do not have baseline data for time use or bag-making), W_i^{strata} is a vector of baseline controls that consists of the variables we used for re-randomization plus the five variables that were imbalanced at baseline, $\theta^{interviewer}$ are interviewer fixed effects, and $\rho^{station*t}$ are either $station * week$

[9] In our time use survey, rather than asking about time spent on bags directly, we asked only about "wage labor (including bags)" in order to maintain a strong separation between the evaluation team and the team that was implementing the bags program. We thus impute time on bags by taking the answer to a question about time on wage labor, and subtracting average time on wage labor from the control-no-bags, GUP-no-bags, and SOUP-no-bags households for each bags group, respectively. See Appendix Table 2 for details.

or $station * month$ fixed effects. We cluster standard errors at the village level, since both GUP/SOUP and bags were assigned at the village level.[10]

We use the Benjamini-Hochberg (Benjamini and Hochberg 1995) and procedures put forward in Anderson 2008 to compute q-values that correct for the multiple hypotheses within each table (and sometimes within panels). We do not extend these corrections beyond the boundary of an individual table (or panel) because the substantive aspects of the hypotheses we test change dramatically across tables. We decided to focus on theoretically related hypotheses, and our tables (panels) are organized exactly on such lines.

5 Impact Results for the Basic Treatments

5.1 Effects of GUP and SOUP

In Table 4 Panel A we report on the basic treatments, GUP and SOUP, including both bags and non-bags households. Columns 1-5 report data collected at two-years; columns 6 and 7 report time use data collected during the bags program, averaged over the five monthly surveys.[11] GUP and SOUP households spend the same amount time providing productive labor as do control households, and report the same amount of leisure time (each of the estimated treatment effects is smaller than four percent of the control mean, and statistically indistinguishable from zero at any conventional level of significance). The GUP treatment raised the value of livestock owned by the household by more than 30 percent relative to control (itt = \$73, s.e. = 16). SOUP households also acquire more livestock (itt = \$32, s.e. = 16), but the net increase is significantly less than that for the GUP households. On the other hand, as column 2 reports, SOUP has as large an effect on total asset value as GUP (and both are statistically different from control). The pattern for income (in column 3) is similar: both SOUP and GUP have positive point estimates, but the GUP effect is almost twice as large as the SOUP and is the only one that is significant. There are no statistically significant effects on consumption or health (columns 4 and 5).

To finish this section, we describe the results for the GUP-no-bags and SOUP-no-bags interventions, reported in Table 4 Panel B. This is of special interest because GUP-no-bags is the classic "graduation" intervention. GUP-no-bags households report statistically significantly lower amounts of leisure than control no-bags households, and also that they spend more time on productive labor (although this later effect is not statistically significant at conventional

[10] For some comparisons, this is conservative, since within GUP-no-bags, GUP-bags, SOUP-no-bags, SOUP-bags, and control-bags villages, each household in the sample was randomly assigned treatment. Comparing, say, GUP-bags to control-no-bags would not require clustering at the village level, because those GUP-bags households could have been individually assigned control-no-bags. But comparing GUP-bags to control-bags requires village-level clustering, because GUP-bags households could not have been individually assigned to control-bags.

[11] We use average time use data here so that we can use the specification from Equation 1, consistent with the rest of the table. In the remainder of the paper, when we report time use data we will not average over surveys, and will use the specification from Equation 2

levels). SOUP-no-bags households also report less leisure time and more productive labor supply than control no-bags households, but neither coefficient is statistically significant (nor can either be distinguished from its corresponding GUP effect). The effects of GUP-no-bags and SOUP-no-bags on livestock, total assets and income parallel those of GUP and SOUP overall: GUP-no-bags has a stronger effect on livestock than SOUP-no-bags, they have similar impacts on total assets, and GUP-no-bags has the largest and only statistically significant impact on income. Neither GUP-no-bags nor SOUP-no-bags has a noticeable impact on health, but SOUP-no-bags does increase consumption.

These program impacts indicate, first, that self-reported income was higher among GUP households, both with and without bags, at the end of the two-year program. Second, they show no evidence of a reduction in labor supply.

5.2 Are we missing the effect on effort?

We find no evidence so far that being a beneficiary of GUP, which raised household earnings, reduced household labor supply. However at this point it is useful to address one additional concern. Is it possible that GUP beneficiaries used their extra income to finance some substitute for their labor used in farming or other businesses and therefore are putting less direct effort into those, which allows them to work harder at the other occupations? As already noted, we do not see evidence of this in our measure of time spent on productive labor, but perhaps it shows up in measures of effort. To get at this we now examine GUP-induced changes in agriculture, which is the dominant household enterprise, and non-farm enterprises in Table 5.

We see that while GUP-bags recipients work somewhat less on their farms compared to control-bags households (column 5), there is no difference in the amount of hired labor they use (column 1). At the same time we see only minimal evidence of labor-saving expenditures, the most important of which would be herbicide. Column 2 shows that there is a statistically significant increase in expenditure on herbicide among GUP-bags household, which is large relative to the control mean, but the absolute magnitude is very small. As a point of comparison, the increase in herbicide equals about two percent of the average use by farmers in this region (calculated from data from the same agroclimatic zone from a representative set of farmers in villages with fewer than 50 compounds (Udry, 2019)). Moreover, there is a more sizable increase in fertilizer expenditure, which is labor-using rather than labor-saving because of its effects on weed growth and output (and here the increase equals 10 percent of the average use in the region, calculated from same regional data). Agricultural earnings are no lower for GUP-bags households—the point estimate is positive (column 8). Moreover GUP-bags has no impact on residual productivity, which is the residual from regressing harvest value on input expenditure, acreage and labor time, and is an attempt to measure the effort the household is putting into agriculture (column 9). In other words there is no evidence that the GUP-bags households are neglecting their agricultural business.

The same holds for their other businesses—the effect on business revenue

(column 10) and earnings (column 11) is positive, albeit not statistically significant—and the effect on time spent on the business is positive and statistically significant (column 6). We do not have ways to measure labor substitution for these businesses, but given the (tiny) scale of the businesses, this seems unlikely.

One other activity where there may be a related concern is household work. We do not have any measure of effort for household work but there is no difference in the time spent on household work by GUP-bags, Control-bags and SOUP-bags households. The last possibility, discussed in the introduction, is wage labor. Wage labor is extremely uncommon in our sample. In control-no-bags, average monthly wage labor earnings are USD 1.13, and only 16% of households have positive wage earnings in a month. In terms of time, in control-no-bags, average time spent on wage labor is 6.2 minutes daily, and only 4.8% of households spent any time on wage labor yesterday.[12] Appendix Table 2 shows that during the bags program, GUP-bags households earned $0.92 less in monthly wage income relative to control-bags. Thus there may be some substitution away from wage labor, but this is very small relative to the increases in earnings across the other sources.

Of course, it is possible that GUP households shift toward doing more relaxing or flexible labor in ways that would be difficult to measure. For example, GUP households might buy milled grain instead of grinding it themselves. Although we cannot rule such possibilities out, it seems to us equally likely that it could be the opposite: for example, GUP households might be exerting more effort in housework due to the presence of additional livestock. In net, these effects are likely to be second order. If they were first order and in the direction of reducing effort, we should also have seen increases in time spent on leisure, which we do not.

5.3 Summary at this point

Taken together these results suggest that GUP increases income (even without the consumption support), while not increasing leisure or reducing labor supply. From Result 1 in our theoretical model, these are consistent with either an investment productivity effect or an psychological/physiological productivity effect from the GUP intervention. The weak impacts on consumption and health shown in Table 4 Panel B suggest that nutritional or other physiological mechanisms are less likely to explain the observed increases in labor supply, especially given that SOUP generated similar impacts on consumption without corresponding labor supply effects. That said, we do not rule out a physiological channel, since even small amounts of measurement error in self-reported consumption may make it difficult to detect such effects.

At this point we also cannot yet rule out the investment productivity effect or, for example, the possibility that the incentive effect resulting from savings collection may be driving these results (as suggested by Result 3). In particular

[12] Demand for wage labor is also low: in control-no-bags, yearly expenditure on wage labor is USD 4.21 and only 10.4% of households demand any labor from the market in a year.

the SOUP outcomes reported above are not clearly enough differentiated from the outcomes of either the control group or GUP, making it difficult to interpret the mechanisms underlying the observed changes from SOUP. To make further progress we turn to the bags intervention.

6 The Evidence from Bags

6.1 Descriptive Statistics on Bags

Of the 1098 clients who were eligible to participate in the employment program, 91.3% chose to make bags at some point over the six months. Over the course of the study, we collected 116,488 bags. On average, the 1098 potential participants produced 4.2 bags per week. Among clients who participated in a given week, the average number of bags submitted was 7. Most people submitted either zero or 10 bags, as demonstrated in Figure 1. Over the course of the study, 35% of bags collected were low quality, 34% were mid quality, and 31% were high quality. Figure 2 shows the distribution of earnings, broken down by complex and simple bags, and holding wage rate constant. Both have a mode at zero (consistent with Figure 1), and the simple bags do show a slight shift towards higheser earnings (undoubtedly because the task was easier).

6.2 GUP Effects on Bags Production and Comparisons with SOUP

The positive effect of the GUP program on the supply of effort to bags production is shown in Panel A of Table 6. GUP participants are more likely to participate in bag production, produce a larger number of bags and earn more from bags production than control-bags. On the other hand SOUP participants are actually less likely to produce bags, produce less bags and earn less from bags production than control-bags. The difference with control bags is not significant, but SOUP-bags participants under-perform GUP-bags participants on almost every measure (for example, there is a twenty-three percentage point gap in bags participation rates). Indeed, improved access to savings is associated with substitution of labor towards household businesses and away from bags (Table 5), perhaps due to an improved ability to manage risk or the timing of working capital needs.

The differences become more stark when we focus on complex bag production in Panel B of Table 6. GUP households produce 43% more complex bags than control households on a weekly basis.[13] SOUP households, on the other hand,

[13]In Appendix Tables 3-5, we explore this effect further. In Appendix Table 3, we show that the additional complex bag production by GUP households does not come at the expense of quality. In Appendix Table 4 we show that this effect is consistent over the course of the bags program, suggesting that the difference in production between GUP and control households is not due to differential learning rates in the early weeks of the program. Finally, in Appendix Table 5, we show that the effect holds for both high-UCT and low-UCT GUP households.

produce 57% fewer complex bags than control bags and a fortiori than GUP-bags. In fact SOUP households spend much more time per dollar earned on complex bags than they do on simple bags, whereas there is no such difference for GUP households.

The dramatic gap in bag-making between GUP and SOUP households suggests that savings collection is unlikely to be the reason why GUP-bags participants earn more than control-bags participants and work no less hard. This conclusion is further supported by the fact that the impacts of GUP, including those on bag-making, are not significantly different for households with and without savings collection (Appendix Table 6).[14]

The fact that GUP-bags participants earn more from and work no less hours at non-bags occupations than control-bags households, and the fact they produce more bags, also sheds light on the possible mechanisms in operation. Specifically, as long as we rule out the savings collection effect, given that no investment is needed in bag production, Result 2 tells us that either a physiological or psychological productivity effect must be in operation.

Before moving on, however, we address the possibility that our treatments shifted intra-household allocation of labor (which our simple model did not account for). However, a close analogue to Result 2 holds for any collective household (Browning and Chiappori, 1998). The primary beneficiary of all of the GUP interventions was the female head of household, and she was the direct recipient of the bag-sewing training. If the receipt of the GUP intervention does not decrease the bargaining power (Pareto weight) of the recipient, and the recipient supplies more labor to the bags activity despite the fact that the household is richer and she is working no less, then there must be a psychological/physiological productivity effect on the bags activity. We have shown in Tables 4 and 5 that GUP increased household income and did not reduce the recipient's overall labor supply, nor is there evidence that it increased her leisure. Overall, 95% of bags were made by the female head of household and 96% were made by an adult female (over age 15). In the final column of Table 6 we find no evidence that GUP or SOUP affected the fraction of bags made by adult females within the household. Shifts in intrahousehold labor allocation, therefore, do not provide an alternative to a physiological or psychological effect of GUP on the supply of labor.[15]

6.3 High UCT versus Low UCT effects on Bags Production and What They Tell Us

What remains to be settled is the source of the physiological or psychological productivity effect. This is because an important part of the GUP intervention was encouragement and hand-holding of the beneficiaries and this could have

[14]The one exception is that GUP-no-savings households spend more time on leisure than GUP-savings households. They do not, however, spend more time on leisure than control households.

[15]The bags program itself may have shifted the intra-household allocation of labor but our evidence focuses on differences between GUP-bags and control-bags households.

directly shifted the cost of effort. While we do not rule out this possibility, what we can show is that there is a pure (positive) income effect operating through the effect of income on physiological or psychological states. For this we turn to the experimental variation in the unconditional cash transfer.

Table 7 compares the outcomes of GUP participants receiving a high level of unconditional cash transfers with those getting less. These households all received the same GUP program and high and low UCT households were balanced in terms of access to savings collection.[16] Column 1 shows that the bags production index is higher for GUP households receiving high UCT than for low UCT, but the difference between the two is not statistically different from zero at conventional levels of significance.[17] However, harvest value and residual productivity are statistically significantly higher for the high UCT households than the low UCT households, suggesting that if there is any crowd out of farming effort due to the GUP intervention, it is happening only for the low UCT households. The high UCT GUP households also spend less on hired labor and herbicide, which is labor-saving, and more on fertilizer (though this last estimate is not significantly different from zero) than low UCT households. While the high UCT households spend less time producing bags, they produce no less (in fact, they produce more) than low UCT households.

This is striking evidence of a physiological or psychological productivity effect resulting from higher income. The high UCT households are more productive at farming, and no less productive in business. They earn 34% more overall and yet produce more bags in less time. It appears that the fact of receiving the high UCT is encouraging those households to produce more from the same amount of time. It is of course possible that in addition there is an encouragement effect that is partly driving the differences between GUP-bags and control-bags.

This evidence also rules out the possibility that the entire effect is driven by the savings collection. High and low UCT households had similar access to savings collection, so that cannot explain the difference between them.

7 Conclusion

The idea that there may be positive rather than negative income effects on labor supply has a long pedigree. We provide support for this view based on a sequence of field experiments designed for this purpose.

Specifically, we find that GUP has a positive effect on income, but does not reduce labor supply, and in fact raises production of bags and especially production of complex bags. This is not driven by the savings component or (entirely) by the encouragement component of GUP. Rather it suggests that extra income provides a psychological productivity boost to ultra-poor house-

[16] The share of households with access to savings was 51.5% and 50% among low-UCT and high-UCT GUP-bags households, respectively.

[17] In Table 7 we show only the estimate for the bag production index; in Appendix Table 7 we report estimates for each component.

holds. This should strengthen the case for well-designed transfer programs to those living in poverty.

References

Baird, S., McKenzie, D., and Özler, B. (2018). The effects of cash transfers on adult labor market outcomes. *IZA Journal of Development and Migration*, 8(1):22.

Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I., and Sulaiman, M. (2017). Labor markets and poverty in village economies. *The Quarterly Journal of Economics*, 132(2):811–870.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., Shapiro, J., Thuysbaert, B., and Udry, C. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, 348(6236):1260799.

Banerjee, A., Karlan, D., Osei, R. D., Trachtman, H., and Udry, C. (2020). Unpacking a multi-faceted program to build sustainable income for the very poor. Working Paper 24271, National Bureau of Economic Research.

Banerjee, A., Rema, H., Gabriel, K., and Benjamin, O. (2017). Debunking the myth of the lazy welfare recipient: Evidence from cash transfer programs worldwide. *World Bank Research Observer*, 32(2):155–184.

Benjamin, D. (1992). Household composition, labor markets, and labor demand: Testing for separation in agricultural household models. *Econometrica*, 60(2):287–322.

Bernard, T., Dercon, S., Orkin, K., and Seyoum Taffesse, A. (2020). The future in mind: Long-run impact of an aspirations intervention in rural ethiopia.

Breza, E., Kaur, S., and Shamdasani, Y. (2018). The morale effects of pay inequality. *The Quarterly Journal of Economics*, 133(2):611–663.

Browning, M. and Chiappori, P.-A. (1998). Efficient intra-household allocations: A general characterization and empirical tests. *Econometrica*, 66(6):1241–1278.

Brune, L. (2016). The effect of lottery-incentives on labor supply: Evidence from a firm experiment in malawi. Working paper.

Brune, L., Chyn, E., and Kerwin, J. T. (2019). Pay me later: A simple employerbased saving scheme. Working paper.

Dalton, P. S., Ghosal, S., and Mani, A. (2016). Poverty and aspirations failure. *The Economic Journal*, 126(590):165–188.

Dasgupta, P. and Ray, D. (1986). Inequality as a determinant of malnutrition and unemployment: Theory. *The Economic Journal*, 96(384):1011–1034.

Fink, G., Jack, B. K., and Masiye, F. (2018). Seasonal liquidity, rural labor markets and agricultural production. Working Paper 24564, National Bureau of Economic Research.

Foster, A. D. and Rosenzweig, M. R. (1996). Comparative advantage, information and the allocation of workers to tasks: Evidence from an agricultural labour market. *The Review of Economic Studies*, 63(3):347–374.

Genicot, G. and Ray, D. (2017). Aspirations and inequality. *Econometrica*, 85(2):489–519.

Gertler, P., Martinez, S., and Rubio-Codina, M. (2006). Investing cash transfers to raise long term living standards. Working Paper 3994, The World Bank.

Goldberg, J. (2016). Kwacha gonna do? Experimental evidence about labor supply in rural Malawi. *American Economic Journal: Applied Economics*, 8(1):129–49.

Guiteras, R. P. and Jack, B. K. (2018). Productivity in piece-rate labor markets: Evidence from rural Malawi. *Journal of Development Economics*, 131:42–61.

Haushofer, J. and Fehr, E. (2014). On the psychology of poverty. *Science*, 344(6186):862–867.

Haushofer, J. and Fehr, E. (2019). Negative income shocks increase discount rates. Working paper.

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

Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575.

Kandasamy, N., Hardy, B., Page, L., Schaffner, M., Graggaber, J., Powlson, A. S., Fletcher, P. C., Gurnell, M., and Coates, J. (2014). Cortisol shifts financial risk preferences. *Proceedings of the National Academy of Sciences*, 111(9):3608–3613.

Karlan, D., Savonitto, B., Thuysbaert, B., and Udry, C. (2017). Impact of savings groups on the lives of the poor. *Proceedings of the National Academy of Sciences*, 114(12):3079–3084.

Kaur, S., Kremer, M., and Mullainathan, S. (2015). Self-control at work. *Journal of Political Economy*, 123(6):1227–1277.

Kaur, S., Mullainathan, S., Schilbach, F., and Oh, S. (2019). Does financial strain lower worker productivity? Working paper.

Kochar, A. (1999). Smoothing consumption by smoothing income: Hours-of-work responses to idiosyncratic agricultural shocks in rural India. *Review of Economics and Statistics*, 81(1):50–61.

LaFave, D. R., Peet, E. D., and Thomas, D. (2020). Farm profits, prices and household behavior. Working Paper 26636.

Leibenstein, H. (1957). *Economic backwardness and economic growth*. Wiley, New York.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *Manchester School*, 22(2):139–191.

Lybbert, T. J. and Wydick, B. (2017). Hope as aspirations, agency, and pathways: Poverty dynamics and microfinance in Oaxaca, Mexico. In *The Economics of Poverty Traps*, NBER Chapters, pages 153–177. National Bureau of Economic Research, Inc.

Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty impedes cognitive function. *Science*, 341(6149):976–980.

Mullainathan, S. and Shafir, E. (2013). *Scarcity: Why having too little means so much*. Times Books/Henry Holt and Co.

Rose, E. (2001). Ex ante and ex post labor supply response to risk in a low-income area. *Journal of Development Economics*, 64(2):371–388.

Rosenzweig, M. R. (1988). Labor markets in low-income countries. *Handbook of Development Economics*, 1:713–762.

Shah, A. K., Mullainathan, S., and Shafir, E. (2012). Some consequences of having too little. *Science*, 338(6107):682–685.

Tanaka, T., Camerer, C. F., and Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. *American Economic Review*, 100(1):557–71.

Udry, C. (2019). Information, market access and risk: Addressing constraints to agricultural transformation in Northern Ghana. *Draft Report*.

Table 1: Experimental Design

Panel A: Intervention and Bags Assignments

Intervention Village Assignment	**Bags Village Assignment**	**# Villages**	**Household Assignment**	**# Households**
control	no bags	34	untreated	526
	bags	42	untreated	376
			treated	397
GUP	no bags	39	untreated	328
			treated	353
	bags	39	untreated	314
			treated	313
SOUP	no bags	38	untreated	238
			treated	345
	bags	39	untreated	272
			treated	388
TOTAL		*231*		*3850*

Panel B: Bags Sub-Treatment Assignment

Intervention Village Assignment - Bags	**Bags Simple/Complex Sub-treatment**	**Bags UCT Sub-treatment**	**# Villages**	**# Households**
control-bags	simple	n/a	21	189
	complex	n/a	21	208
GUP-bags	simple	high UCT	10	69
		low UCT	10	90
	complex	high UCT	9	79
		low UCT	10	75
SOUP-bags	simple	n/a	19	202
	complex	n/a	20	186
TOTAL			*120*	*1098*

Panel A shows intervention treatment assignments (GUP, SOUP, and control) and assignment to the Bags program. Both were assigned at the village level. Within each village assigned to GUP or SOUP, about half of sample households were treated with GUP or SOUP, respectively. All treated households in bags villages received the Bags program. In control villages assigned to bags, about half of sample households were selected to receive the bags program. Panel B shows sub-treatments within the Bags program. All sub-treatments were randomized at the village level such that al individuals who received the Bags program received identical sub-treatment assignments. Control-Bags = intervention control villages assigned to Bags. GUP-bags GUP intervention villages assigned to bags. SOUP-bags = SOUP intervention villages assigned to bags. Simple = assigned to sew the simple bag. Complex = assigned to sew the complex bag. high UCT = GUP intervention households with Bags who received an unconditional cash transfer of USD 3.92 each week. low UCT = GUP intervention households with Bags who received an unconditional cash transfer of USD 1.31 each week. All monetary values are reported in 2014 USD, Purchasing Power Parity (PPP) terms.

Table 2: Household Level Descriptive Statistics for Baseline Indicators, Intervention Treatments
Means and Standard Deviations

	ctrl-no-bags	GUP-no-bags	SOUP-no-bags	ctrl-bags	GUP-bags	SOUP-bags	p-value, F-test joint sig
household size	7.16	7.61	7.28	7.18	7.29	7.48	0.69
	(3.88)	(4.16)	(3.72)	(3.50)	(3.62)	(3.51)	
average age, household	25.57	25.02	25.13	24.77	24.37	23.04	0.00
	(10.90)	(10.55)	(10.69)	(9.68)	(9.11)	(8.10)	
land area (acres)	4.54	4.50	4.84	4.65	5.03	4.82	0.56
	(3.92)	(4.02)	(4.12)	(3.76)	(3.88)	(3.94)	
monthly per cap cons. (USD)	58.96	54.00	62.61	58.58	59.24	58.55	0.15
	(38.68)	(34.49)	(40.88)	(39.14)	(41.87)	(36.97)	
monthly HH income (USD)	42.34	42.60	48.74	41.12	45.07	45.49	0.67
	(56.76)	(58.84)	(56.77)	(53.07)	(50.13)	(56.63)	
savings balances (USD)	2.40	1.63	2.78	1.65	2.46	3.11	0.47
	(11.98)	(8.54)	(13.91)	(9.86)	(11.06)	(13.97)	
food security index	0.00	0.03	0.16	-0.14	0.14	0.19	0.00
	(1.01)	(1.02)	(1.11)	(0.91)	(1.05)	(1.07)	
asset value index	0.03	0.03	0.09	0.03	0.00	0.02	0.95
	(1.07)	(1.01)	(1.12)	(1.02)	(0.73)	(0.77)	
financial inclusion index	0.00	0.00	0.05	-0.08	-0.08	0.19	0.08
	(0.96)	(0.94)	(1.07)	(0.84)	(0.76)	(1.36)	
physical health index	-0.06	0.00	0.00	0.03	0.13	0.00	0.12
	(1.02)	(1.07)	(1.04)	(0.98)	(0.93)	(0.99)	
mental health index	0.00	-0.04	0.14	-0.08	0.03	0.08	0.15
	(1.02)	(0.97)	(1.00)	(1.02)	(1.02)	(1.01)	
political involvement index	0.00	0.09	-0.16	0.06	0.00	0.01	0.06
	(1.00)	(0.99)	(1.01)	(0.99)	(1.00)	(1.00)	
female empowerment index	0.00	0.11	-0.09	0.02	-0.12	0.00	0.11
	(0.99)	(1.03)	(0.99)	(1.01)	(1.02)	(1.00)	

We report means and standard deviations for key indicators at baseline. Indices are centered around mean baseline value. The last column contains the p-value from an F-test of joint significance of all treatments. See Appendix for components of all indices.

Table 3: Household Level Descriptive Statistics for Baseline Indicators, Bags Sub-Treatments
Means and Standard Deviations

	ctrl-bags simple	ctrl-bags complex	GUP-bags simple, low UCT	GUP-bags simple, high UCT	GUP-bags complex, low UCT	GUP-bags complex, high UCT	SOUP-bags simple	SOUP-bags complex	p-value, F-test joint sig
household size	6.96	7.38	6.59	7.71	8.13	7.28	7.58	7.38	0.82
	(3.37)	(3.60)	(3.19)	(4.30)	(4.24)	(2.81)	(3.37)	(3.64)	
average age, household	24.20	25.28	26.87	23.02	24.45	22.74	22.72	23.36	0.34
	(8.50)	(10.61)	(11.36)	(7.37)	(9.12)	(6.72)	(7.43)	(8.73)	
land area (acres)	4.76	4.54	3.45	5.01	6.50	6.17	5.01	4.63	0.03
	(4.21)	(3.30)	(3.09)	(4.11)	(3.83)	(3.82)	(4.03)	(3.82)	
monthly per cap cons. (USD)	51.98	64.57	46.03	51.17	73.18	75.58	63.31	53.64	0.00
	(33.04)	(43.08)	(35.19)	(34.71)	(48.80)	(44.15)	(37.22)	(36.06)	
monthly HH income (USD)	41.37	40.89	25.69	43.46	71.51	56.17	53.74	36.99	0.02
	(58.20)	(47.93)	(30.06)	(43.75)	(77.74)	(48.24)	(66.85)	(42.03)	
savings balances (USD)	1.46	1.82	4.27	1.43	1.76	1.73	2.54	3.70	0.68
	(8.16)	(11.18)	(15.16)	(6.20)	(9.46)	(9.48)	(12.37)	(15.43)	
food security index	-0.28	0.00	-0.13	0.28	0.08	0.36	0.31	0.06	0.00
	(0.79)	(0.99)	(0.86)	(1.15)	(1.02)	(1.10)	(1.05)	(1.07)	
asset value index	-0.10	0.15	-0.12	0.00	0.12	0.10	0.06	0.00	0.59
	(0.87)	(1.13)	(0.66)	(0.71)	(0.96)	(0.67)	(0.78)	(0.75)	
financial inclusion index	0.00	-0.14	0.00	-0.05	-0.21	-0.10	0.08	0.30	0.14
	(0.89)	(0.79)	(0.91)	(0.75)	(0.57)	(0.66)	(1.18)	(1.51)	
physical health index	0.00	0.06	0.12	0.27	-0.09	0.11	0.00	0.03	0.30
	(0.97)	(0.98)	(1.01)	(0.61)	(1.12)	(0.94)	(1.02)	(0.97)	
mental health index	-0.20	0.02	-0.10	0.18	0.00	0.07	0.14	0.01	0.34
	(0.96)	(1.06)	(1.02)	(1.10)	(0.96)	(0.93)	(1.07)	(0.95)	
political involvement index	0.06	0.06	0.03	0.12	-0.14	-0.08	0.06	0.00	0.90
	(0.99)	(0.99)	(1.00)	(0.98)	(1.01)	(1.01)	(0.99)	(1.00)	
female empowerment index	0.05	0.01	0.10	-0.10	-0.39	-0.27	-0.06	0.00	0.49
	(0.98)	(1.04)	(1.07)	(0.95)	(1.07)	(0.96)	(1.02)	(0.97)	

We report means and standard deviations for key indicators at baseline. Indices are centered around mean baseline value. The last column contains the p-value from an F-test of joint significance of all treatments. See Appendix for components of all indices.

Table 4: Intervention Treatment Effects at Two Years

Panel A: Impacts of GUP and SOUP

		livestock value	asset value	monthly household income	monthly cons. per capita	physical health index	time prod. labor	time leisure
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
GUP	itt	72.92	74.85	7.49	1.13	0.04	3.55	-4.55
	se	(17.61)	(38.64)	(2.54)	(1.62)	(0.03)	(10.41)	(6.11)
	pval	0.00	0.05	0.00	0.48	0.24	0.73	0.46
	qval	0.00	0.16	0.03	0.58	0.36	0.80	0.56
SOUP	itt	32.09	83.33	4.00	2.91	-0.01	-9.44	1.97
	se	(14.65)	(39.53)	(2.49)	(1.64)	(0.05)	(11.87)	(7.59)
	pval	0.03	0.04	0.11	0.08	0.85	0.43	0.80
	qval	0.11	0.12	0.25	0.20	0.85	0.54	0.83
Ctrl Mean		240.17	589.48	36.59	44.15	-0.16	597.71	127.42
Ctrl SD		348.59	764.13	43.08	30.15	0.84	128.32	114.42
Obs		2909	2900	2907	2880	2767	1221	1221
GUP - SOUP = 0	pval	0.05	0.86	0.25	0.35	0.36	0.29	0.43

Panel B: Impacts by Bags Treatment

		livestock value	asset value	monthly household income	monthly cons. per capita	physical health index	time prod. labor	time leisure
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
control-bags	itt	-3.16	56.57	2.48	2.98	-0.02	44.44	-27.63
	se	(16.21)	(46.06)	(2.49)	(2.01)	(0.05)	(12.87)	(6.70)
	pval	0.85	0.22	0.32	0.14	0.65	0.00	0.00
	qval	0.85	0.34	0.45	0.29	0.74	0.01	0.00
GUP-no-bags, UCT $6-$9	itt	77.20	104.70	9.40	2.73	0.01	23.67	-30.43
	se	(23.75)	(47.92)	(3.33)	(2.07)	(0.04)	(18.43)	(10.39)
	pval	0.00	0.03	0.01	0.19	0.76	0.20	0.00
	qval	0.02	0.11	0.03	0.34	0.81	0.34	0.03
GUP-bags, UCT $3.9 or $1.3	itt	64.73	72.67	6.44	0.63	0.05	30.61	-18.22
	se	(21.75)	(54.46)	(3.32)	(1.63)	(0.05)	(11.32)	(7.50)
	pval	0.00	0.18	0.05	0.70	0.35	0.01	0.02
	qval	0.03	0.34	0.16	0.78	0.47	0.04	0.07
SOUP-no-bags	itt	26.77	107.61	4.97	5.09	0.03	28.11	-13.26
	se	(18.55)	(49.79)	(3.14)	(1.99)	(0.05)	(21.88)	(14.60)
	pval	0.15	0.03	0.12	0.01	0.63	0.20	0.37
	qval	0.30	0.11	0.25	0.06	0.73	0.34	0.48
SOUP-bags	itt	34.92	91.45	4.22	1.99	-0.07	14.01	-14.24
	se	(20.01)	(50.65)	(3.36)	(1.61)	(0.08)	(12.68)	(8.47)
	pval	0.08	0.07	0.21	0.22	0.37	0.27	0.09
	qval	0.20	0.20	0.34	0.34	0.48	0.39	0.22
Ctrl Mean		242.93	578.33	35.94	43.64	-0.17	575.48	137.46
Ctrl SD		356.19	760.35	42.14	29.43	0.85	125.66	111.80
Obs		2909	2900	2907	2880	2767	1221	1221

Panel A shows average effects of GUP and SOUP; the omitted group is control households (bags and non-bags) in any village. Panel B shows effects by bags sub-treatment; the omitted group is control non-bags households in any village. The sample is restricted to villages with more than 30 compounds. We include surveyor fixed effects and control for stratification variables, imbalanced variables (average household age, food security index, land area, monthly per capita consumption, and monthly household income), whether or not household was treated with bags (Panel A only), and baseline value of the outcome when possible. Standard errors clustered at village level. We use the Benjamini-Hochberg step-up method to compute q-values, considering all tests in the table. Columns 1-5 are taken from the two-year survey; Columns 6-7 are averages over the five monthly time use surveys administered during the bags program. Livestock value is the total number of livestock owned times the median reported price for each animal. Asset value is the total number of assets (including livestock, household and productive assets, and stocks), valued using asset prices relative to the price of goats from other countries. Monthly household income is monthly self-reported household income, computed as the sum of income from the household's business, farm, wage labor, and (revenue from) animals. Monthly consumption per capita is self-reported monthly consumption per capita, including both food and non-food expenditure. Physical health index includes two variables. The first is the average daily living score, which is the mean of four variables: capacity bathing, capacity lifting, capacity walking, and capacity working (each measured on a scale from 1 being easily done to 4 being unable to do). The second is sick day, which is 1 if the member did not miss a day of work due to illness in the last year, 0 otherwise. Time productive labor is minutes spent yesterday spent on bags or wage labor, agriculture, business, animals, and home labor (time spent on children, cleaning, cooking, collecting firewood, shopping, or fetching water). Time leisure is minutes spent yesterday on religious activities, social activities, ceremonies, traveling, personal care, and resting. All monetary values are reported in 2015 USD, Purchasing Power Parity (PPP) terms.

Table 5: Productive Non-Bags Labor during Bag-Making

		(1) expenditure on labor	(2) expenditure on herbicide	(3) expenditure on fertilizer	(4) time bags	(5) time field	(6) time business	(7) time home labor	(8) harvest value	(9) residual prod.	(10) business revenue	(11) business income
GUP-bags	itt	2.59	3.83	20.06	-2.12	-20.96	9.22	1.18	59.31	-4.52	14.97	8.76
	se	(3.54)	(2.11)	(13.35)	(11.64)	(9.80)	(4.80)	(11.79)	(55.62)	(64.00)	(11.66)	(6.05)
	pval	0.47	0.08	0.14	0.86	0.04	0.06	0.92	0.30	0.94	0.21	0.16
	qval	0.69	0.30	0.39	0.94	0.27	0.30	0.94	0.50	0.94	0.46	0.39
SOUP-bags	itt	-0.59	1.27	17.49	-21.85	-24.02	10.06	4.41	-22.20	-60.59	-9.47	-2.94
	se	(2.18)	(2.32)	(20.01)	(12.37)	(10.86)	(4.74)	(11.42)	(70.42)	(54.77)	(6.16)	(2.56)
	pval	0.79	0.59	0.39	0.08	0.03	0.04	0.70	0.75	0.28	0.13	0.26
	qval	0.92	0.81	0.61	0.30	0.27	0.27	0.91	0.92	0.50	0.39	0.50
Ctrl Mean		4.02	3.83	57.58	112.94	206.61	28.48	299.27	408.47	-13.50	13.94	6.54
Ctrl SD		15.67	10.82	88.11	106.61	185.65	78.23	155.09	461.91	349.39	55.61	24.83
Obs		272	272	272	1978	3442	3442	3442	272	266	287	287

Effects of treatments on productive non-bags labor outcomes for the bags sample, during the bag-making period. The omitted group is control-bags households. The sample is restricted to villages with more than 30 compounds. We include surveyor fixed effects and control for stratification variables, imbalanced variables (average household age, food security index, land area, monthly per capita consumption, and monthly household income), and baseline value of the outcome when possible. Standard errors clustered at the village level. We use the Benjamini-Hochberg step-up method to compute q-values, considering all tests in the table. Expenditure on labor, herbicide, fertilizer are expenditures in the last 12 months, measured in the midline survey during the bag-making period. Time on bags, field, business, and home labor are minutes spent yesterday on bags, agriculture, business, and home labor, respectively, measured in five monthly surveys during bag-making (on average, 78% of the 1098 bags households were found and surveyed each month). We compute time on bags by taking the answer to a question about time on wage labor, and subtracting average time on wage labor from the control-no-bags, GUP-no-bags, and SOUP-no-bags households for each bags group, respectively. See Appendix Table 1 for details. For this outcome we exclude one of the stations, Langbensi, that reported zero time spent on wage labor. Harvest value is the total revenue from harvest, measured in the midline survey six months after the bag-making period. Residual productivity contains the residuals from a regression of harvest value on input expenditure, acreage, and average time on field. Business revenue and income pertains to all non-agricultural businesses and is measured in the midline survey during the bag-making period. All monetary values are reported in 2014 USD, Purchasing Power Parity (PPP) terms.

Figure 1: Bags

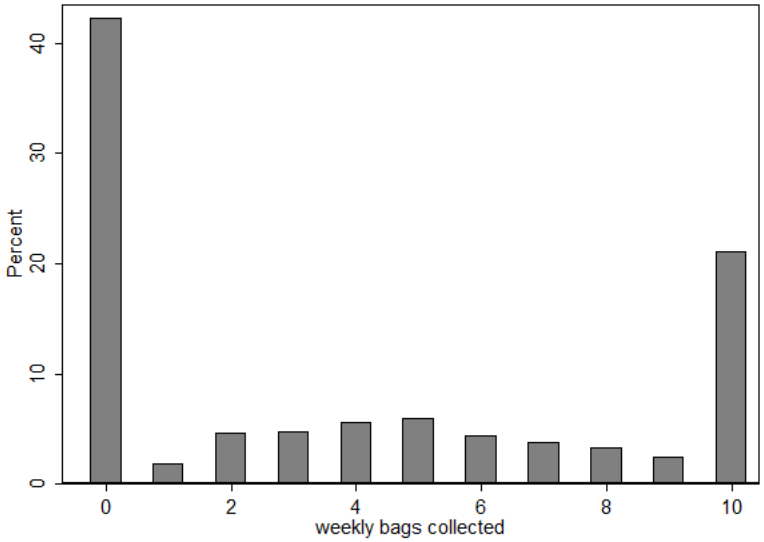


Figure 2: Bags Earnings by Complexity

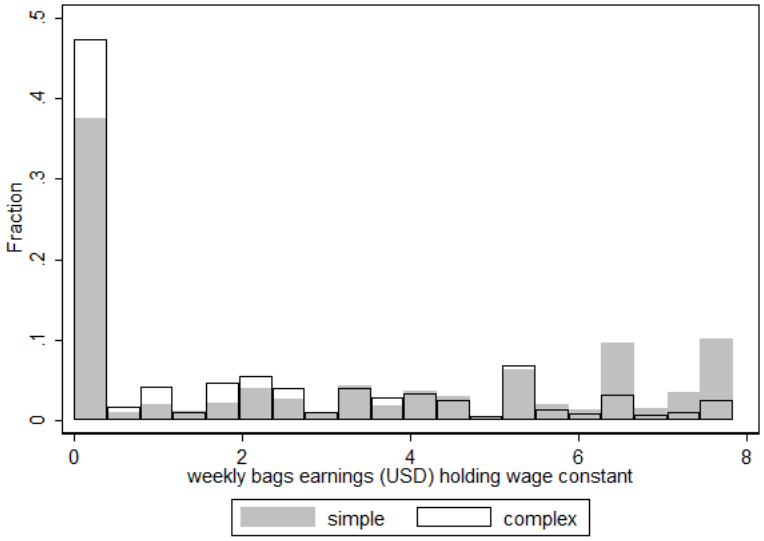


Table 6: Bag Production

Panel A: Effects of GUP, SOUP

		(1) bags production index	(2) number of bags	(3) participates (0/1)	(4) bags earnings	(5) minutes per dollar earned
any GUP	itt	0.28	1.14	0.12	0.66	-117.91
	se	(0.12)	(0.50)	(0.05)	(0.35)	(97.61)
	pval	0.03	0.03	0.02	0.06	0.23
	qval	0.10	0.10	0.10	0.17	0.41
any SOUP	itt	-0.17	-0.65	-0.11	-0.33	227.98
	se	(0.13)	(0.49)	(0.06)	(0.38)	(221.55)
	pval	0.19	0.18	0.06	0.39	0.31
	qval	0.37	0.37	0.17	0.49	0.45
Ctrl Mean		0.00	3.76	0.58	2.45	676.09
Ctrl SD		1.00	3.97	0.49	3.01	1316.80
Obs		18816	18816	18816	18816	1661
any GUP - any SOUP	pval	0.00	0.00	0.00	0.03	0.13

Panel B: Effects of Complex

		(1) bags production index	(2) number of bags	(3) participates (0/1)	(4) bags earnings	(5) minutes per dollar earned
control complex	itt	-0.18	-0.94	-0.06	-0.70	376.85
	se	(0.13)	(0.51)	(0.06)	(0.36)	(116.92)
	pval	0.18	0.07	0.30	0.05	0.00
	qval	0.37	0.19	0.45	0.17	0.02
GUP simple	itt	0.23	0.66	0.10	0.51	148.59
	se	(0.16)	(0.68)	(0.06)	(0.52)	(133.75)
	pval	0.16	0.33	0.12	0.33	0.27
	qval	0.37	0.45	0.31	0.45	0.45
GUP complex	itt	0.11	0.63	0.08	-0.02	90.75
	se	(0.20)	(0.82)	(0.09)	(0.48)	(127.30)
	pval	0.57	0.45	0.38	0.96	0.48
	qval	0.64	0.54	0.49	0.96	0.56
SOUP simple	itt	-0.08	-0.30	-0.09	-0.07	280.71
	se	(0.19)	(0.73)	(0.08)	(0.58)	(216.30)
	pval	0.68	0.68	0.29	0.90	0.20
	qval	0.73	0.73	0.45	0.93	0.37
SOUP complex	itt	-0.46	-2.08	-0.21	-1.39	854.84
	se	(0.12)	(0.44)	(0.07)	(0.31)	(377.08)
	pval	0.00	0.00	0.00	0.00	0.03
	qval	0.00	0.00	0.03	0.00	0.10
Ctrl Mean		0.17	4.59	0.62	3.07	472.56
Ctrl SD		1.07	4.24	0.49	3.30	918.84
Obs		18816	18816	18816	18816	1661
GUP complex - ctrl complex	itt	0.29	1.57	0.14	0.68	-286.09
GUP complex - ctrl complex	pval	0.14	0.05	0.14	0.15	0.03

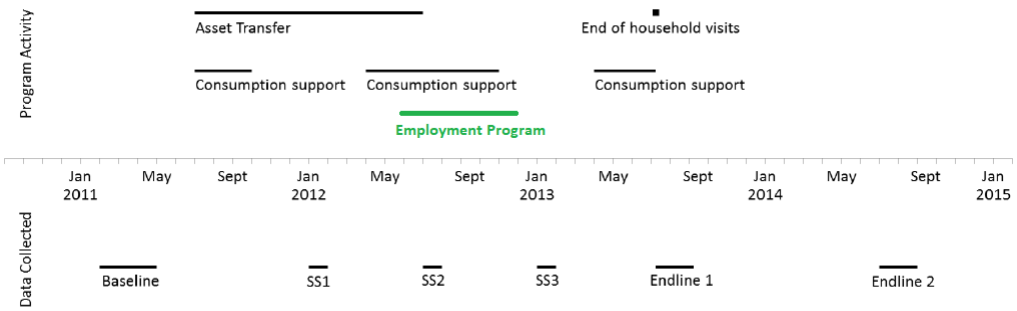
In Panel A, we show effects of GUP and SOUP on bag-making labor supply for bags households. The omitted group is control-bags households (i.e. those who received neither GUP nor SOUP but were assigned to the bags program). In Panel B, we show effects of being assigned the complex bag by treatment on bag-making labor supply for bags households. The omitted group is control-bags households with simple bags. In both panels, the sample is restricted to villages with more than 30 compounds. We control for stratification variables and imbalanced variables (average household age, food security index, land area, monthly per capita consumption, and monthly household income). Columns 1-4 report weekly data with station-week fixed effects (896 people over 21 weeks). The bags production index is a standardized index of the variables in columns 2-5, centered around the control-bags mean. Column 5 reports monthly data with station-month fixed effects, since this measure incorporates time use data (time use data was collected on only a monthly basis; on average, 78% of the 1098 bags households were found and surveyed each month). Standard errors clustered at the village level. We use the Benjamini-Hochberg step-up method to compute q-values, considering all tests in the table. We compute minutes per dollar earned by taking average daily earnings over the course of the month as the denominator, and time on bags (measured once in the month) as the numerator. We compute time on bags by taking the answer to a question about time on wage labor, and subtracting average time on wage labor from the control-no-bags, GUP-no-bags, and SOUP-no-bags households for each bags group, respectively. See Appendix Table 1 for details. All monetary values are reported in 2014 USD, Purchasing Power Parity (PPP) terms.

Table 7: Effects of High vs. Low UCT

		(1) bags prod. index	(2) harvest value	(3) residual prod.	(4) exp. labor	(5) exp. herbicide	(6) exp. fertilizer	(7) time bags	(8) time field	(9) time business	(10) time home labor	(11) business revenue	(12) business income
GUP, UCT $3.9	itt	0.37	138.60	94.92	-2.62	-0.32	26.20	-28.90	-25.13	14.85	11.73	10.58	6.42
	se	(0.17)	(64.83)	(60.56)	(3.48)	(1.72)	(14.72)	(12.55)	(11.82)	(5.95)	(11.38)	(12.44)	(5.36)
	pval	0.03	0.03	0.12	0.45	0.85	0.08	0.02	0.04	0.01	0.31	0.40	0.23
	qval	0.14	0.14	0.29	0.57	0.85	0.21	0.14	0.14	0.14	0.46	0.53	0.39
GUP, UCT $1.3	itt	0.17	-46.51	-136.67	9.45	9.36	11.91	18.19	-16.15	2.75	-10.95	21.01	12.00
	se	(0.15)	(79.86)	(70.00)	(5.09)	(3.20)	(19.20)	(13.66)	(12.74)	(6.07)	(17.66)	(21.26)	(9.31)
	pval	0.25	0.56	0.05	0.06	0.00	0.54	0.19	0.21	0.65	0.54	0.32	0.20
	qval	0.39	0.61	0.18	0.19	0.10	0.61	0.38	0.38	0.68	0.61	0.46	0.38
Ctrl Mean		0.00	408.47	-13.50	4.02	3.83	57.58	112.94	206.61	28.48	299.27	13.94	6.54
Ctrl SD		1.00	461.91	349.39	15.67	10.82	88.11	106.61	185.65	78.23	155.09	55.61	24.83
Obs		18816	272	266	272	272	272	1978	3442	3442	3442	287	287
high UCT - low UCT	pval	0.35	0.04	0.01	0.07	0.01	0.49	0.01	0.55	0.11	0.22	0.69	0.63

Differences in labor supply between GUP high UCT and GUP low UCT for bags households, during the bag-making period. The omitted group is control-bags households. The sample is restricted to villages with more than 30 compounds. We include surveyor fixed effects and control for stratification variables, imbalanced variables (average household age, food security index, land area, monthly per capita consumption, and monthly household income), and baseline value of the outcome when possible. Standard errors clustered at the village level. We use the Benjamini-Hochberg step-up method to compute q-values, considering all tests in the table. Bag Production Index is a standardized index of weekly bags submitted, weekly participation (0/1), bags earnings, and minutes per dollar earned, centered around the control-bags mean. Harvest value is the total revenue from harvest, measured in the midline survey six months after the bag-making period. Residual productivity contains the residuals from a regression of harvest value on input expenditure, acreage, and average time on field. Expenditure on labor, herbicide, fertilizer are expenditures in the last 12 months, measured in the midline survey during the bag-making period. Time on bags, field, business, and home labor are minutes spent yesterday on bags, agriculture, business, and home labor, respectively, measured in five monthly surveys during bag-making (on average, 78% of the 1098 bags households were found and surveyed each month). We compute time on bags by taking the answer to a question about time on wage labor, and subtracting average time on wage labor from the control-no-bags, GUP-no-bags, and SOUP-no-bags households for each bags group, respectively. See Appendix Table 1 for details. For this outcome we exclude one of the stations, Langbensi, that reported zero time spent on wage labor. Business revenue and income pertains to all non-agricultural businesses and is measured in the midline survey during the bag-making period. All monetary values are reported in 2014 USD, Purchasing Power Parity (PPP) terms.

Appendix Figure 1: Timeline



In the top part of the timeline we show program activities, and in the bottom part we show data collection. During the employment program we conducted additional time use surveys each month, over five months.

Appendix Figure 2: Simple Bag (left) and Complex Bag (right)

The simple bag has "running" stitches on the hem and strap. The complex bag has a more complicated pattern on the hem and strap: a sequence of four "chain" stitches alternating with one "running" stitch.

Appendix Table 1: Justifying Imputation of Time Spent on Bags

VARIABLES	(1) Monthly Wage Income (USD)	(2) Time Bags and/or Wage Labor
GUP-no-bags	-0.78**	-5.55
	(0.31)	(4.96)
SOUP-no-bags	-0.36	-2.54
	(0.65)	(3.65)
control-bags	0.21	77.98***
	(0.39)	(5.69)
any GUP-bags	-0.92***	74.96***
	(0.27)	(6.85)
SOUP-bags	-0.25	56.97***
	(0.58)	(6.57)
Observations	864	789
Ctrl Mean	1.130	6.220
any-GUP-bags = GUP-no-bags	0.590	0
SOUP-bags = SOUP-no-bags	0.890	0

This table shows levels of monthly wage income and time spent on bags and/or wage labor across treatment groups. In Column 1, we can see that within each treatment group—control, GUP, and SOUP—there is very little difference in wage income between bags and no-bags, despite large differences in time spent on bags and/or wage labor, as shown in Column 2. Therefore, we assume that any differences in time spent on "time bags and/or wage labor" within each treatment group, between bags and no-bags, can be attributed to time spent on bags. We thus impute time spent on bags by taking the time spent on "time bags and/or wage labor" for each bags participant, and subtracting the mean time spent on "time bags and/or wage labor" from the corresponding no-bags treatment group. For example, for a GUP-bags participant, we subtract the mean time spent on "time bags and/or wage labor" in GUP-no-bags to impute time spent on bags.

Appendix Table 2: Wage Elasticity Results

Panel A: Evidence of Responsiveness to Wages Received for Previously Submitted Bags

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	IHS(bags)	IHS(bags)	IHS(bags)	IHS(bags)	IHS(bags)	IHS(bags)	IHS(bags)
log(wage)	0.02	-0.00	0.10	-0.02	0.15**	-0.11***	0.19***
	(0.04)	(0.05)	(0.07)	(0.04)	(0.07)	(0.04)	(0.07)
Observations	23,058	14,822	8,236	16,470	6,588	13,146	9,912
consecutive		no	yes				
experience						no	yes
fourth week				no	yes		

Panel B: Elasticity Estimates with respect to the 3-Week Lagged Wage

	(1)
VARIABLES	IHS(bags)
log(wage(t-3))	0.16***
	(0.05)
Observations	19,764
experience	no

Panel A provides evidence that participants were responsive to wages they were receiving for bags submitted previously, as opposed to the correct relevant wage for the bags they were making. We examine elasticities by three sub-groups. First, we look at participants who were randomly assigned two consecutive high wage months and two consecutive low wage months (39/120 villages, and 363/1098 participants). Second, we look at participant-weeks that were the fourth week in the wage month. Participants were paid wages with a two-week lag. If participants only fully internalized the wage change upon receiving new wages, then they should take the new wage into account only for bags produced in the fourth week of the month. (The new wage is active in the first week of production; wages for these bags are paid in the third week, and thus only bags collected in the fourth week are produced with experience of new wage.) Finally, we define "experience" to mean either the fourth week of the month, or for "consecutive" participants, any week in the second consecutive month with the same wage. Given this evidence, Panel B shows elasticity estimates with respect to the 3-week lagged wage.

Appendix Table 3: Effects of High vs. Low UCT - Components of Bags Production Index

		(1) number of bags	(2) participates (0/1)	(3) bags earnings	(4) minutes per dollar earned
GUP, UCT $3.9	itt	1.46	0.16	0.86	-164.21
	se	(0.67)	(0.07)	(0.49)	(113.28)
	pval	0.03	0.03	0.09	0.15
	qval	0.13	0.13	0.23	0.31
GUP, UCT $1.3	itt	0.77	0.08	0.44	-82.20
	se	(0.66)	(0.07)	(0.41)	(130.70)
	pval	0.24	0.22	0.28	0.53
	qval	0.32	0.32	0.32	0.53
Ctrl Mean		3.76	0.58	2.45	676.09
Ctrl SD		3.97	0.49	3.01	1316.80
Obs		18816	18816	18816	1661
high UCT - low UCT	pval	0.44	0.41	0.49	0.59

Differences in labor supply between GUP high UCT and GUP low UCT for bags households. The sample is restricted to villages with more than 30 compounds. We control for stratification variables, imbalanced variables (average household age, food security index, land area, monthly per capita consumption, and monthly household income), and baseline value of the outcome when possible. Columns 1-4 report weekly data with station-week fixed effects (896 people over 21 weeks). Column 5 reports monthly data with station-month fixed effects. (Time use data was collected on only a monthly basis for roughly 60% of households over 5 months, and only about 60% of households were found each month.) Standard errors clustered at the village level. We use the Benjamini-Hochberg step-up method to compute q-values, considering all tests in the table. Standard errors clustered at the village level. We compute minutes per dollar earned by taking average daily earnings over the course of the month as the denominator, and time on bags (measured once in the month) as the numerator. We compute time on bags by taking the answer to a question about time on wage labor, and subtracting average time on wage labor from the control-no-bags, GUP-no-bags, and SOUP-no-bags households for each bags group, respectively. All monetary values are reported in 2014 USD, Purchasing Power Parity (PPP) terms.

Appendix: Variable Definitions and Construction

Variable	Definition
animal revenue, (monthly, USD)	Revenue from animal sales, births, slaughters, and animal product sales in the last 6 months. Each revenue stream is computed by multiplying the reported number of animals or products sold/born/slaughtered for revenue and multiplying it by the reported sale price (if available) or the median sale price. We sum over all four animal revenue streams and then divide by six.
asset value (USD)	We ask about all of the assets owned by the household. We calculate the price in goats for each asset by using price data from other countries, as in Banerjee et al. (2015). We then sum the total asset value in goats, and multiply by the price of goats (average of median buy price and sell price across villages).
asset value index	We ask about all of the assets owned by the household. We calculate the price in goats for each asset by using price data from other countries, as in Banerjee et al. (2015). We then sum the total asset value in goats, and standardize it into an index around the baseline mean.
bags earnings	Weekly earnings from bags submitted (not including any unconditional consumption support), paid with a two-week lag. Earnings depend on number of bags submitted, the relevant wage (which varied each month), and the quality of each bag. If the wage was GHC 0.70, mid-quality bags earned GHC 0.70, high-quality bags earned GHC 0.80, and low-quality bags earned GHC 0.60. If the wage was GHC 0.30, mid-quality bags earned GHC 0.30, low-quality bags earned GHC 0.20, and high-quality bags earned GHC 0.40. For the measure used in Figure 2, we hold the wage constant; we assume that all mid-quality bags earned GHC 0.50, low-quality bags earned GHC 0.40, and high-quality bags earned GHC 0.60.
bags production index	Standardized index of four variables, centered around control-bags mean. The variables are: number of bags, participates (0/1), bags earnings, and minutes per dollar earned.
business income (monthly, USD)	Monthly business revenues minus expenses. For each business within the household, we ask about how many months in the last year the business was operating, how many months were "normal" (neither higher nor lower than last month), how many months were "high profit," and how many months were "low profit." We ask about sales and profits in the last month, as well as profits in normal, high, and low months. We use the ratio of last month's profits to high and low profits to impute sales in high and low profit months, compute total sales in the last year by summing over sales in normal, high, and low months, and divide by the number of months the business was operating to get monthly revenue. We do the same exercise for expenses.
business revenue (monthly, USD)	For each business within the household, we ask about how many months in the last year the business was operating, how many months were "normal" (neither higher nor lower than last month), how many months were "high profit," and how many months were "low profit." We ask about sales and profits in the last month, as well as profits in normal, high, and low months. We use the ratio of last month's profits to high and low profits to impute sales in high and low profit months, compute total sales in the last year by summing over sales in normal, high, and low months, and divide by the number of months the business was operating to get monthly revenue.
crop income (monthly, USD)	Harvest sale value minus expenditure on inputs (annual), divided by twelve. To get harvest sale value, we ask about the quantity of each crop sold in the last year. If the units of harvested crops are the same as the units of sold crops and we have the sale price, then we use this price to compute the sale value of each crop, and then sum over crops. Otherwise, we use the median price for that crop. To get expenditure on inputs, we ask about expenditures on manure, fertilizer, labor, herbicide, insecticide, and other inputs in the last year, and then sum over all categories.

expenditure on fertilizer (USD)	Expenditure on fertilizer in the last year.
expenditure on herbicide (USD)	Expenditure on herbicide in the last year.
expenditure on labor (USD)	Expenditure on farm labor in the last year.
female empowerment index	Standardized index of five variables, centered around baseline means. Each variable is the answer to the question ""To what extent do you believe yourself able to make your own decisions concerning X?" The categories X are food, school expenses, health expenses, visiting friends, and purchases. They are measured on a scale from 1 to 3.
financial inclusion index	Standardized index of two variables, centered around the baseline mean. The first variable is the total amount received in loans by the household in the last year. The second variable is the total savings balances at the time of the survey.
food security index	Standardized index of three variables, centered around baseline means. The first two variables equal 0 if the household answered "all year" or "during the lean season only" to the following questions, about adults and kids, respectively: "Did adults/kids ever reduce number of meals per day or reduce portions over the past year?" The third variable equals 0 if the household answered "all year" or "during the lean season only" to the question "Did adults ever skip entire days without eating?"
harvest value (USD)	We ask about the quantity of each crop sold in the last year. If the units of harvested crops are the same as the units of sold crops and we have the sale price, then we use this price to compute the sale value of each crop, and then sum over crops. Otherwise, we use the median price for that crop.
livestock value (USD)	Sum of values of goats, fowl, pigs, sheep, and cows. To get the value of each animal, we first take the average of the median buy price and the median sell price across villages. We then multiple this price by the number of animals owned in the household.
mental health index	Standardized index of three variables, centered around baseline means. The first is economic satisfaction, measured on a scale from 1 to 5. The second is a standardized index of five measures: feeling sad, crying, not eating, not working, and feeling restless, measured on a scale from not at all, hardly ever, some of the time, or most of the time. The third is whether the individual was not worried in the last year (0 if the member experienced a period of worry in the last year, 1 otherwise).
minutes per dollar earned	We first compute the minutes spent on bags yesterday (see below). Since time use surveys were conducted once per month, we have one measure of this for each month. We then divide this by the average daily earnings for the month. (This variable is thus measured with error, as some households spent no time yesterday but had high average monthly earnings, while others spent lots of time yesterday but had very low average monthly earnings.)
monthly per capita consumption (USD)	For food consumption, we ask about household consumption of a variety of food products in the last month, and multiply the amounts by median prices to get expenditure. For non-food consumption, we ask about monthly and yearly expenses on a wide variety of goods and services that are typically consumed on a monthly (yearly) basis. Total monthly consumption per capita is the sum of monthly household food and non-food consumption, divided by the household size.
monthly household income (USD)	total of monthly business income, crop income, wage income, and animal revenue (we do not have animal income because we were unable to collect data on costs associated with animals).
number of bags	Weekly bags collected.
participates (0/1)	1 if the household submitted at least 1 bag that week, 0 otherwise.

physical health index	Standardized index of two variables, centered around baseline means. The first is the average daily living score, which is the mean of four variables: capacity bathing, capacity lifting, capacity walking, and capacity working (each measured on a scale from 1 being easily done to 4 being unable to do). The second is whether or not they did not take a sick day (1 if the member did not miss a day of work due to illness in the last year, 0 otherwise).
political involvement index	z-score of attendance at village meetings in the last year (1 if the person attended a village meeting in the last 12 months, 0 otherwise), centered around baseline mean.
residual productivity	The residuals from a regression of harvest value on input expenditure, acreage, and time field.
savings balances (USD)	We ask about the total amount saved in various locations (a private bank, at home, with a friend, etc.) and sum them all together.
time bags	We take the answer to a question about minutes spent on wage labor yesterday, and subtract average minutes spent yesterday on wage labor from the control-no-bags, GUP-no-bags, and SOUP-no-bags households for each bags group, respectively. In the appendix we confirm that there is no evidence that the bags treatment had an effect on wage income, so we assume it also did not affect time spent on wage labor. (We did not ask directly about time on bags because we were worried that it would compromise the separation between the evaluation and implementation teams.) We exclude one of the stations, Langbensi, that reported zero time spent on wage labor.
time business	Minutes spent on household business yesterday.
time home labor	Minutes spent on home labor yesterday. Home labor includes cleaning, cooking, collecting firewood, shopping, and fetching water.
time field	Minutes spent on the farm yesterday.
time productive labor	Minutes spent on wage labor (including bags), business, home labor, and field, yesterday.
time leisure	Minutes spent on leisure activities yesterday (resting, socializing, etc.)
wage income (monthly, USD)	We ask about each instance of paid labor in the last nine months within the household, the quantity of time worked, and the total earnings from that activity. We sum the total earnings and divide by nine to get monthly household wage income.