Aspiring to a Better Future: Can a Simple Psychological Intervention Reduce Poverty?*

Kate Orkin, Robert Garlick, Mahreen Mahmud,
Richard Sedlmayr, Johannes Haushofer, Stefan Dercon

January 17, 2023

Abstract

How do material and psychological constraints individually and jointly affect investment and poverty? To answer this, we design a workshop teaching techniques to raise aspirations and encourage long-term planning. We cross-randomise this with large unconditional cash transfers in a 415-village, 8,300-person, 1.5-year experiment in Kenya. The workshop increases aspirations and investment, and raises living standards, at least twice as cost-effectively as cash transfers. But combining the workshop with cash produces similar effects to cash alone, potentially because cash raises aspirations. This shows that alleviating psychological constraints can improve economic outcomes but these constraints may relax when material conditions improve.

*Orkin: University of Oxford (kate.orkin@bsg.ox.ac.uk); Garlick: Duke University (robert.garlick@duke.edu); Mahmud: University of Exeter (m.mahmud@exeter.ac.uk); Sedlmayr: University of Oxford (richard.sedlmayr@bsg.ox.ac.uk); Haushofer: Stockholm University (johannes.haushofer@su.se); Dercon: University of Oxford (stefan.dercon@economics.ox.ac.uk). The pre-analysis plans, questionnaires, and intervention materials are posted at https://www.socialscienceregistry.org/trials/996. We thank GiveDirectly, the Busara Centre for Behavioural Economics; Innovations for Poverty Action; and staff at the University of Oxford. Pat Olvera and Rita Wachera produced videos with the research team and Will Le, Stephanie Palla, Michael Cooke and Miriam Laker oversaw the GiveDirectly collaboration. We thank Charles Amuku, Paul Brimble, Rosie Chiang, Axel Eizmendi Larrinaga, Sam Fishman, Marta Grabowska, Michala Iben Riis-Vestergaard, Mungai Mwangi, Geetika Nagpal, Gabor Nyeki, Emma Riley, Giulio Schinaia, Victoria Schreitter, Catherine Thomas, and John Walker for outstanding research assistance. We thank Nava Ashraf, Jon de Quidt, Doug Gollin, Erica Field, Sarojini Hirshleifer, Travis Lybbert, Karen Macours, Edward Miguel, Paul Niehaus, Gautam Rao, Imran Rasul, Emma Riley, Frank Schilbach, Jessi Streib, Duncan Thomas, Chris Udry, Michael Walker, Bruce Wydick and many seminar participants for thoughtful input. The project is funded by Fidelity Charitable, the UKRI GCRF Accelerating Achievement for Africas Adolescents Hub, the Bill and Melinda Gates Foundation, Good Ventures, the JPAL PPE Initiative, and a donation from John Collenette. This study has ethics approval from Oxford (protocol # SSD/CUREC1A/BSG C1A_16-002), Maseno (protocol # MSU/DRPI/MUERC/00196/15) and Duke (protocol # 2018-0020) universities.

1 Introduction

Psychological constraints may prevent people living in poverty from taking up work or investment opportunities, despite potentially high returns.[1] One possible constraint is that people living in poverty may lack opportunities to learn "the capacity to aspire" (Appadurai, 2013): to set higher long-term aspirations to improve their socio-economic position and plan concrete steps to achieve them. They may not encounter institutions or traditions which encourage them to strive and plan for a better future or they may be less likely to encounter role models like them who have succeeded. They may set lower aspirations than they are able to attain and hence limit investments, contributing to persistence of poverty.[2]

This paper provides causal evidence consistent with this idea. We show that teaching rural Kenyan women simple techniques to set higher aspirations and plan to achieve them increases household investment and labour supply and improves living standards after 17 months. We find that teaching these scalable techniques is highly cost-effective in improving consumption-based measures of living standards compared to an unconditional cash transfer, a common anti-poverty intervention. We show the techniques work by boosting women's aspirations for their household's future economic position, not by providing information or other plausible psychological mechanisms.

We work with a sample of 8,300 rural women living in poverty in Western Kenya. We find their aspirations and expectations are closely linked to their investment and wealth. We measure their aspirations as the levels of assets, income, and children's education they *would like to reach* a decade or more in the future. Aspirations are typically higher than but correlated with expectations: respondent-specific distributions of these outcomes they *think they will reach* (Dominitz and Manski, 1997). In our observational data, aspirations and expectations are strongly associated with wealth and multiple types of investment conditional on wealth. They are more strongly associated than other psychological characteristics, such as beliefs about one's ability to achieve desired outcomes (e.g., self-efficacy or locus of control), time and risk preferences, or mental health.[3]

To examine whether boosting aspirations and expectations can *cause* improved investment and wealth, we develop a unique aspirations and planning workshop lasting roughly 80 minutes, administered by fieldworkers in small groups. Participants watch a video about two local fictional

[1] Work on this theme includes Angelucci and Bennett (2021), Ashraf et al. (2022), Baranov et al. (2020), Banerjee et al. (2020), Barker et al. (2022), Bhat et al. (2022), Blattman et al. (2017), Cecchi et al. (2022), Campos et al. (2017), Ghosal et al. (2020), Kaur et al. (2015), Kaur et al. (2021), McKelway (2021), and Rojas Valdes et al. (2022).

[2] This idea comes up in, for example, anthropology (Appadurai, 2013; Rapport, 2016), sociology (Willis, 1977, 1987), and economic theory on aspirations (Dalton et al., 2016; Genicot and Ray, 2017; Lybbert and Wydick, 2018) and neighbourhood and role model effects (Benabou, 1996; Durlauf, 1996; Loury, 2009; Streufert, 2000).

[3] Eble and Escueta (2022), Guyon and Huillery (2021), Janzen et al. (2017), La Ferrara (2019), Ross (2019) and Ross et al. (2021) document correlations between aspirations and investment in other settings.

female role models who improve their economic position by setting higher aspirations and medium-term goals; defining concrete, immediate steps to achieve goals; and planning for obstacles. They then do facilitated exercises to learn to use these techniques.[4] We study the impact of this workshop using a four-armed village-level randomised controlled trial with 8,300 women in 415 villages.

We present four main findings. First, we show that this "psychological" intervention has remarkably large effects on economic outcomes. We compare participants who are invited to the workshop (arm 1) to participants who are invited to a placebo intervention (arm 2). The placebo contains the same factual information as the workshop but without inspiring role models or teaching psychological techniques. We survey women roughly 17 months after the interventions to measure prespecified psychological and household economic outcomes. We find that the psychological intervention by itself substantially increases six prespecified economic outcomes. It raises expenditure on productive inputs by 22% of the placebo group mean, labour supply by 5%, education expenditure by 3% (not statistically significant), revenue by 11%, consumption by 4%, and non-land asset value by 6%. It raises an index combining these six outcomes by 0.11 standard deviations.

Second, we show these effects are of an economically important magnitude, by showing the workshop is highly cost-effective in improving living standards relative to a widely-used anti-poverty intervention, unconditional cash transfers. Participants in the third arm of the experiment are offered large, lump-sum cash transfers of 2,237 USD PPP, roughly 60% of mean annual household consumption, and the placebo intervention. The cash intervention increases the index of six economic outcomes by 0.23 standard deviations compared to the placebo, roughly double the size of the effect of the psychological intervention. We collect data on intervention costs and compare costs to accumulated benefits for recipients in terms of consumption expenditure and assets, an approach used in cost-effectiveness calculations for graduation programmes (Bandiera et al., 2017; Banerjee et al., 2015). The psychological intervention has a benefit-cost ratio at least twice and possibly up to six times as high as the cash transfer at the scale of this experiment, depending on the measure of costs, showing psychologically-informed interventions can achieve important reductions in poverty.

Third, we assess whether low aspirations continue to limit investment in economic activities, even when respondents receive a large income shock, by comparing participants in the cash arm to participants in the fourth, "combined", arm, who are invited to the workshop and offered cash transfers. Perhaps surprisingly, adding the psychological intervention to the cash transfer has limited effects: most economic outcomes are similar in the cash-only and combined arms. However, education expenditure is substantially higher in the combined arm.

Finally, we show that aspirations and expectations for one's future economic position are the

[4]Schools in high-income countries teach some related techniques (EEF, 2021a,b; US Department of Education, 2007) but our participants are unlikely to have been exposed to them, as discussed in Section 4.1.

most plausible psychological mechanism for the economic effects in all three arms. The psychological intervention has large positive treatment effects on aspirations and expectations 17 months after intervention. It has no effects on other psychological mechanisms we measure (self-efficacy, locus of control, time and risk preferences, and mental health). Effects are unlikely to be due to participants acquiring new information: we compare the workshop to an information equivalent-placebo and show it has few effects on beliefs about returns to economic activities and does not encourage mimicry of activities in the videos. We also show evidence against effects being driven by experimenter demand. We present a simple model of reference-dependent utility to explain these results. In this model, either aspirations (Dalton et al., 2016; Genicot and Ray, 2017; Lybbert and Wydick, 2018) or expectations (Kőszegi and Rabin, 2006, 2007) proxy for reference points and exceeding (missing) one's reference point for future consumption increases (decreases) utility. A higher reference point, induced by the psychological intervention, causes higher current investment to finance higher future consumption and meet the reference point, matching the treatment effects we observe.

Aspirations and expectations also provide a plausible mechanism for the similar economic effects of the combined and cash interventions. The cash transfer alone and the combined intervention have similar, large, positive effects on both aspirations and expectations and have near-zero effects on most other psychological mechanisms. This suggests that, in the combined group, the cash transfer crowds out the aspiration- and investment-promoting effect of the workshop, perhaps because of the transfer's windfall nature. Consistent with the windfall explanation, both cash alone and the combined intervention increase the share of expenditure allocated to investment more than standard consumption models would predict. We can reject the alternative explanation that the psychological intervention is less effective for wealthier households: the economic effects of the workshop by itself are not systematically different for respondents who are wealthier or poorer at baseline.

We contribute to a growing body of work on how psychological characteristics – beliefs, thinking styles, preferences, or soft skills – influence individual economic choices and outcomes; and how economic factors shape those psychological characteristics. We innovate relative to previous work in three ways. Our first contribution is to provide robust experimental evidence that low aspirations and expectations for one's future economic position constrain investment and reduce wealth. Research suggests that people from disadvantaged groups or neighbourhoods lack role models or exposure to institutions encouraging them to aspire to and plan for a better future (Benabou, 1996; Dalton et al., 2016; Durlauf, 1996; Genicot and Ray, 2017; Loury, 2009; Streufert, 2000). They thus may set lower aspirations than might be feasible for them to attain, which can lower investment and perpetuate poverty. Our work is unique in cleanly testing the full theoretical causal chain suggested by this literature, by running the first cluster-randomised placebo-controlled

study of how an intervention targeting aspirations affects aspirations, investment, and living standards. Prior work examines the effects of exposure to the life stories of role models (Beaman et al., 2012; Bernard et al., 2014; Jensen and Oster, 2009; Kipchumba et al., 2021; La Ferrara et al., 2012; Macours and Vakis, 2014; Porter and Serra, 2020; Riley, 2019) or role models and coaching (Carlana et al., 2022; Cecchi et al., 2022; Goux et al., 2017; Rojas Valdes et al., 2022). Unlike this work, we compare the workshop to an information-equivalent placebo which contains the same images and descriptions of economic opportunities as the psychological intervention. This isolates the effect of exposure to specific psychological practices – setting higher aspirations and planning to implement goals – and to role models who achieve success by applying these practices.[5] We rule out a broad range of other psychological mechanisms, such as self-efficacy, locus of control, and preferences. Our village-level randomisation largely eliminates spillovers relative to the within-school or within-geography randomisation of other papers in this literature, making our results robust to the risk of intervention content spilling over to the placebo group or behavioural responses from participants knowing others' treatment status. We also collect a large enough sample to measure a causal chain from effects on aspirations to investment to living standards, building on existing studies which examine specific investments in education or microenterprises.[6]

Our work complements but differs from a strand of work studying the effect of intensive interventions to change thinking patterns or non-cognitive skills through repeated exposure and practice, in multi-week training with skilled facilitators. For example, papers study interventions which build patience and self-regulation (Alan and Ertac, 2018; Alan et al., 2019; Blattman et al., 2017); slow down violent reactions (Heller et al., 2017); boost self-image and self-efficacy (Ghosal et al., 2020; John and Orkin, 2022; McKelway, 2021); build ability to simulate vivid images of alternative future scenarios (Ashraf et al., 2022); or encourage ambitious, innovative mindsets (Campos et al., 2017; McKenzie et al., 2022). We target a different psychological driver of behaviour which can be boosted through simple techniques by laypeople in one session, potentially because (as theory suggests), low aspirations result simply from lack of exposure to role models and practices of long-term planning, not thinking patterns, skills or underlying preferences. More broadly, our work provides empirical microeconomic foundations for a long economic tradition in both macro- and microeconomics studying how particular beliefs or preferences, such as a "long-term orientation" (Figlio et al., 2019), propensity for delayed gratification (Galor and Özak, 2014), or preferences for redistribution (Luttmer and Singhal, 2011), form, persist, and affect levels

[5]Genicot and Ray (2020) argue that inability to distinguish effects on aspirations from effects of new information has been a key weakness in this literature.

[6]Papers study effects on education (Beaman et al., 2012; Bernard et al., 2014; Carlana et al., 2022; Goux et al., 2017; Kipchumba et al., 2021; Macours and Vakis, 2014; Porter and Serra, 2020; Riley, 2019), microenterprises (Cecchi et al., 2022; Rojas Valdes et al., 2022) or fertility (Jensen and Oster, 2009; La Ferrara et al., 2012).

of economic development (Guiso et al., 2006; Spolaore and Wacziarg, 2013).

Our second contribution is to study the effectiveness and cost-effectiveness of a psychological intervention as a poverty reduction strategy. Our paper is the first in the literatures on aspirations, role models and thinking patterns to include a cash transfer comparison arm and collect cost data, drawing on work benchmarking non-psychological interventions to cash (McIntosh and Zeitlin, 2021, 2022). We study a general sample of women living in poverty and find positive treatment effects for most subgroups. Our intervention could be administered as a broad-based poverty reduction tool with limited targeting. This complements work on aspirations in specific populations whose aspirations may be particularly malleable, such as microfinance clients (Cecchi et al., 2022; Rojas Valdes et al., 2022) or school children (Beaman et al., 2012; Carlana et al., 2022; Goux et al., 2017; Kipchumba et al., 2021; Porter and Serra, 2020; Riley, 2019).

Our third contribution is to the literature on potential complementarity between alleviating psychological and material constraints. Theoretical work argues that psychological and material constraints can interact, generating poverty traps where psychological constraints reduce investment and increase poverty, with negative psychological consequences (Boswell-Dean et al., 2018; Dalton et al., 2016; De Quidt and Haushofer, 2018; Genicot and Ray, 2017). This raises the possibility of complementary effects from simultaneously alleviating material and psychological constraints. There is compelling evidence that treating mental health or substance abuse disorders alongside economic support generates complementarities (Angelucci and Bennett, 2021; Blattman et al., 2017, 2022). However, these studies are with recipients facing severe psychological challenges which can limit their ability to function in daily tasks (Lund et al., 2022). Indeed, they benefit little from the economic interventions offered in these trials without treatment. We examine another psychological constraint, low aspirations, in the general population. Instead of complementarity, we find that adding the psychological intervention to cash has little effect on aspirations or economic outcomes relative to cash alone, in part because of the novel finding that the unconditional cash transfer increases aspirations. These findings highlight the importance of studying the psychological effects of economic interventions and understanding each element of bundled anti-poverty programmes (Bandiera et al., 2017; Banerjee et al., 2015, 2018; Bossuroy et al., 2022; Sedlmayr et al., 2020).

Sections 2 and 3 describe the context and sample, including the non-experimental relationship between psychological and economic measures. In Section 4, we describe our experimental design and interventions. Sections 5 and 6 present results on the interventions' effects and cost-effectiveness. In Section 7, we explore mechanisms driving the economic effects.

2 Context and Sample

We study a group of people living in poverty in 415 villages in Homa Bay and Siaya counties, near Lake Victoria in Western Kenya. The area is rural but almost all villages are within a few hours' drive of Kisumu, Kenya's third largest city. Roughly half of the villages contain a primary school, one third contain a market, and one sixth contain a clinic (Table B.2).

We conducted a census of 32,921 households living in these villages. We used the census data to identify households that (1) included an adult female member and were either (2a) widow-headed or (2b) classified as poor by a multidimensional poverty index. The first criterion was chosen because the psychological intervention was aimed at adult women. The second criterion aligns with the eligibility criteria for the cash transfer programme we study, which targeted widows and poor households. Our sample consists of 8,309 households meeting these conditions. Per capita consumption is below the World Bank's current definition of extreme poverty in 55% of these households.

We conducted baseline surveys an average of five months before treatment and endline surveys an average of seventeen months after treatment, completing all surveys before the COVID19 pandemic. We surveyed the "primary adult woman" in each household: typically the female spouse in a couple-headed household or the widowed head of household. We surveyed 87.1% of the baseline sample at endline. Attrition is balanced across treatment arms and only weakly related to baseline household characteristics (Table B.1). See Appendix B for details on the census, eligibility rules, sampling, and endline tracking process.

The baseline and endline surveys covered three prespecified household-level investment measures – labour supply, expenditure on productive inputs and hired labour, and education expenditure – and three prespecified economic outcomes – revenue, consumption expenditure, and asset value. We chose these three investment types because they are covered in the psychological intervention and are available to almost all households. Most measures are adapted from the Living Standards Measurement Surveys and Indonesian Family Life Survey. We also measure multiple psychological characteristics, discussed in Section 3.1. See Appendix G for details on all measures.

We report summary statistics for the endline placebo sample so that we use the same time period and survey instrument as the treatment effects analysis; baseline statistics for the full sample are similar and balanced by treatment group (Table B.2). The average respondent in the placebo group is 41 years old with an interdecile range of 23–65 (Table 1). The average household contains 2.8 adult members and 2.8 children, which includes biological and non-biological children of younger respondents and grandchildren of older respondents. 59% of the respondents are married and 40% are widows, because widow-headed households are automatically eligible for our sample. Only 43% of respondents have completed primary education.

Table 1: Summary Statistics

	(1)	(2)	(3)	(4)
	Mean	Std dev.	10th percentile	90th percentile
Demographics				
Respondent Age	40.6	16.5	23.0	65.0
Household Size	5.56	2.53	2.00	9.00
Number of Children	2.77	1.86	0.00	5.00
Married	0.591	0.492		
Primary School Education	0.426	0.495		
Economic Investment				
Labour Supplied (days)	525	347	84	960
Expenditure on Productive Inputs	857	1,932	23	1,973
Education Expenditure	640	940	42	1,703
Enrolment Rate for Ages 6-13	0.978	0.124	1.000	1.000
Enrolment Rate for Ages 14-20	0.788	0.343	0.000	1.000
Economic Outcomes				
Revenue	2,101	3,204	165	4,708
Consumption	3,796	1,959	1,643	6,407
Consumption per Adult Equivalent	928	579	397	1,605
Value of Non-land Assets	1,529	1,506	289	3,219

Notes: This table shows endline summary statistics for the 1767 households of the placebo group. All currency values are measured in 2018 USD PPP. All flow measures except education expenditure are in annual terms. Details on measurement are given in Appendix G. Education expenditure is all fee and non-fee expenditure in the current and preceding school years summed over all household members aged 6-20. It is set to missing for households with no members aged 6-20. Labour supply is days of work on farm and non-farm household enterprises or supplied to the market, for all household members older than 15. Input expenditure includes purchase of inputs and stocks and inventory, rental, maintenance, and expenditure on hired labour, for household activities in crop agriculture and livestock rearing, and for non-farm enterprises owned or operated by household members. Revenue captures the value of production sold or consumed at home from these activities, valued at farm-gate prices. Consumption expenditure captures the value of purchased and home-produced food, non-durable and durable household goods, and social expenditures following Deaton and Zaidi (2002). Consumption per adult equivalent is adjusted for household demographic composition using adult equivalent scales for Kenya following Anzagi and Bernard (1977). Non-land assets are the estimated value, if sold, of durable assets, livestock, and stocks of dried maize, as well as cash savings. We value output and expenditure on inputs following the Living Standards Measurement Surveys (Grosh and Glewwe, 2000) and UN System of National Accounts (FAO, 1996).

We study a broad sample of poor households in a rural area, rather than restricting to groups with a particularly high propensity for investment, like microenterprise owners. Households in our sample have three main avenues for investment: labour supply, spending on inputs for home production, and children's education. The average household supplies 525 days of labour per year or 216 days per adult member per year, with almost all adult members doing some work, and spends 857 USD PPP per year on productive inputs (Table 1). Labour supply and input spending is typically spread over multiple economic activities: 95% of households grow crops, most commonly maize; 83%

raise livestock, most commonly chickens; 40% do casual or salaried work outside the household; and 44% operate a non-farm enterprise, most commonly retail (50%), manufacturing (23%), services or fishing (11% each). The average household spends 640 USD PPP per year on education (183 USD PPP per child), mostly on school fees. School enrollment is 98% and 79% for respectively primary school-aged children and teenagers, so there is limited scope for our treatments to affect enrollment.

Households' labour supply and input expenditure generates mean annual revenue of 2,101 USD PPP. They consume an average of 3,796 USD PPP annually, or 2.5 USD PPP per adult equivalent per day.[7] Households own non-land assets worth 1,529 USD PPP on average. The average household owns another 5,638 USD PPP of land and housing assets, although the values of land and housing are difficult to measure accurately as land transactions are rare.

3 The Psychological Correlates of Investment and Wealth

Here, we explore which psychological factors might constrain labour supply, investment, or wealth for the people we study. We identify and measure seven possible psychological predictors of investment and wealth – aspirations and expectations for one's future economic position, beliefs about one's ability to carry out actions or achieve desired outcomes, beliefs about returns to specific investment activities, time and risk preferences, and depression, which predict investment or wealth in other contexts. We describe their relationships with wealth and investment in our context. Appendix G provides more details on measurement.

3.1 Definitions and Measurement of Mechanisms

Aspirations for One's Future Economic Position: We define aspirations as the set of future outcomes that people prefer and aim for. We directly elicit a proxy measure for the extent to which individuals aspire to a better life and aim to achieve improvements in their economic circumstances. We measure aspirations as the level of outcomes respondents would like their household to attain. We measure them in three domains: the levels of assets and income the respondent would like their household to reach in ten years, and the level of education they would like one of their children to reach in adulthood. The child is the child aged closest to 14 from the roster of resident household members. We chose these dimensions because they are relevant to almost all respondents and because, in qualitative scoping work, most respondents described a better life in terms of improvements on these dimensions.[8] We aggregate across dimensions into

[7] We include the value of goods produced for home consumption in both the consumption and revenue measures. Our consumption measure is higher on average than our revenue measure, a common pattern in agricultural household surveys (e.g. Bandiera et al. 2017; Banerjee et al. 2020; Egger et al. 2022).

[8] Aspirations can refer to more specific long-term goals (Locke and Latham, 2002): for example, putting a tin roof on one's house. We do not use these as our primary measure of aspirations as they are difficult to compare

an inverse covariance-weighted average to capture a more general aspirational mindset.[9] Similar measures of aspirations predict higher savings and education investment (Janzen et al., 2017; Ross et al., 2021; Beaman et al., 2012) and small business investment (Dalton et al., 2018).

Expectations: We measure respondents' expectations about their possible distributions of assets and income over the same time period as aspirations and estimate the individual-specific means of these distributions, following Dominitz and Manski (1997) and Delavande et al. (2011). We also measure the level of education respondents expects a child to reach in adulthood (the same child as in the aspirations measure) and combine the three measures into an inverse covariance-weighted average.

Self-beliefs: "Self-beliefs" are people's beliefs about their ability to achieve desired outcomes (Locke and Latham, 2002). We measure three different self-beliefs. Generalised self-efficacy is a person's belief about their ability to carry out actions and achieve their goals across domains (Bandura, 1977). Locus of control is a person's beliefs about whether their actions or fate determine outcomes (Rotter, 1966). Growth mindset is a person's belief that their skills can be altered through effort (Dweck, 2012).[10] We measure these beliefs on psychological scales validated in other contexts, which we translated and back-translated into Luo and validate in a separate study to show they have appropriate psychometric properties (see Appendix G.3). We combine the three measures in an inverse covariance-weighted index, as they are correlated in our sample and in other work (Bong and Skaalvik, 2003). These measures predict labour supply, job search, saving, and educational effort in existing research (Caliendo et al., 2015; Dweck, 2012; Heckman and Kautz, 2012; John and Orkin, 2022; Lindqvist and Vestman, 2011; McKelway, 2021).

Beliefs about Returns to Investment: We measure respondents' beliefs about the returns to three potential investment activities: using more labour on their farm, using more fertiliser on their farm, and a university degree (vs secondary education) for their child. For example, we elicit beliefs about the levels of maize output if a farmer works 12 hours per week more for a given land size and set of inputs, and calculate the percentage change they expect relative to working their current hours. We picked these activities because most households face choices about whether to invest in them. Beliefs about these individual returns are broadly realistic, in that they are aligned with measures of actual returns from other work in this literature, as we discuss in Appendix

between respondents or over time. However, we measure goals in the psychological and combined groups after the psychological intervention and find that respondents with a specific goal in a domain also have higher levels on our quantitative metric of aspirations. See Appendix A.2.

[9] Averaging multiple measures of related concepts to improve precision is common in aspirations research (Beaman et al., 2012; Bernard and Taffesse, 2014; Janzen et al., 2017) and econometrics in general (Schennach, 2022).

[10] These are closely related to measures of self-confidence, self-esteem or identity as a competent person (Rosenberg and Kapland, 1982).

G.2.2. Beliefs about returns to investment and/or labour supply predict investment in education, migration and economic activity in other settings (see Delavande (forthcoming) for a review.)

Present Bias and Patience: We measure patience and present bias over money from an incentivised seven-choice multiple price list with choices over tomorrow vs 15 days and 15 vs 29 days (Andersen et al., 2008). Individuals who place higher utility weight on future relative to current consumption have higher savings and educational attainment in many contexts (Falk et al., 2018).

Propensity for Risk-Taking: We create an ordinal measure of risk-taking from an incentivised Eckel and Grossman (2002) measure. Propensity to take risks is correlated with higher levels of technology adoption and business ownership, potentially because it lowers the value of certain consumption and leisure today relative to uncertain future outcomes (Liu, 2013; Falk et al., 2018).

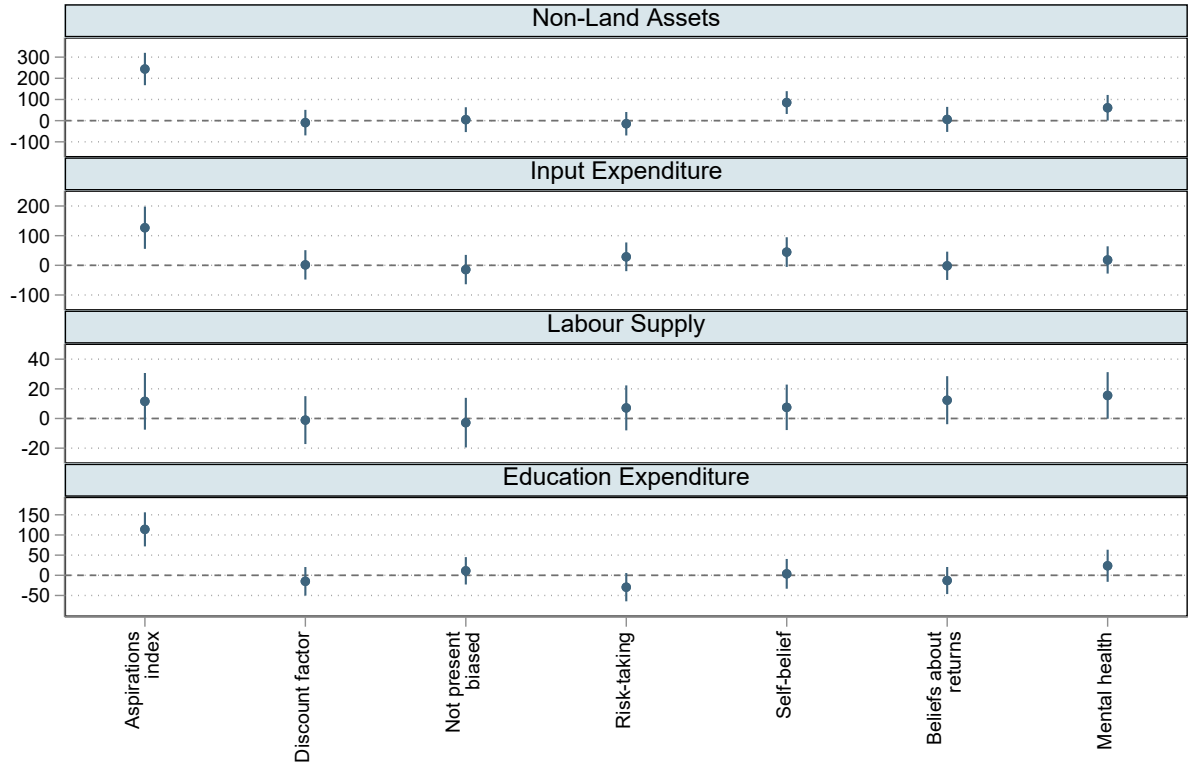
Mental Health: We focus on depression, a common mental health condition, measured with the standardised 10-item Centre for Epidemiological Studies-Depression (CESD) scale (Andresen et al., 1994). Symptoms include pessimistic beliefs, impairment in motivation and ability to perform everyday activities, and dejected mood (Beck and Alford, 2009). Treatments which reduce depression improve labour supply and investment, productivity at work, savings and education investment (Angelucci and Bennett, 2021; Baranov et al., 2020; Lund et al., 2022).

3.2 Relationships Between Psychological Characteristics and Investment and Wealth

Aspirations and expectations are stronger predictors of investment and wealth than the other psychological characteristics discussed above. To show this, we first use endline placebo group data to regress respondents' wealth on their aspirations, self-beliefs, beliefs about returns, time and risk preferences, mental health, age, education, marital status, household size, number of school-aged members, and county fixed effects. This regression shows two key patterns, in the top panel of Figure 1. First, wealthier people have higher aspirations: a one standard deviation increase in aspirations is associated with 247 USD PPP more non-land assets. Second, other psychological characteristics are more weakly associated with wealth: wealthier people have higher self-beliefs and mental health but these relationships are less than half as large as the aspirations-wealth relationship.

We next regress investment on the same controls listed above and consumption and non-land asset value. The two additional controls allow us to compare investment by households with different psychological characteristics but the same (proxies for) wealth. These regressions show two key patterns, in the bottom three panels of Figure 1. First, people with higher aspirations invest more. A one standard deviation increase in aspirations is associated with 128 USD PPP more input expenditure, 12 more days of household labour supply (not statistically significant),

Figure 1: Investment-Aspirations Relationship Using Non-experimental Variation



Notes: This figure shows coefficients and 95% confidence intervals from regressing different wealth and investment measures on psychological characteristics. The first row regresses non-land assets on respondents' aspirations, self-beliefs, beliefs about returns, time and risk preferences, mental health, age, education, marital status, household size, number of school-aged members, and county fixed effects. The second, third and fourth rows measure investment using respectively expenditure on productive inputs and hired labour, labour supply, and education expenditure. We regress each measure of investment on the same controls as in the first row, plus consumption and non-land asset value. All asset and investment measures are defined in the footnote below Table 1. The aspirations index and psychological measures are defined in Section 3.1. Here, each is standardised to allow for coefficient comparison. All regressions use the endline placebo group data with the top percentiles of aspirations, investment, assets, and consumption trimmed. Sample size is 1376 to 1747 depending on the choice of controls and investment measure. The smaller sample sizes are for education expenditure, as this is set to missing for households with no school-aged children. The confidence intervals are estimated using heteroskedasticity-robust standard errors.

and 108 USD PPP higher education expenditure.[11] Second, other psychological characteristics are again weakly associated with investment.

The same four patterns hold when we replace aspirations with expectations and repeat this analysis (Figure A.1). We run the aspirations and expectations analyses separately because the two are highly correlated ($\rho = 0.6$) and conceptually similar. The patterns are broadly similar when we remove all controls, drop the other psychological characteristics, or include respondent

[11]Separately, we find domain-specific aspirations are related to domain-specific investment. For example, education expenditure is positively related to education-specific aspirations ($\rho = 0.1$).

fixed effects using the panel structure (Figure A.2). The patterns hold using baseline data on aspirations, although the baseline analysis excludes time and risk preferences, which we did not measure at baseline (Figure A.3).[12] They are similar when we control for cognitive ability using a subsample of data.[13] Standardised ridge regressions, a method to measure predictive power, classify aspirations as the strongest predictor of both wealth and investment.

These four patterns motivate a focus on aspirations and expectations as key potential psychological influences on investment and wealth in this context. They are also consistent with a model of investment influenced by reference points, which we present in Section 7.3, where aspirations or expectations proxy for reference points for future consumption and people invest more to accumulate higher wealth to afford their reference-point level of consumption in the future. However, these findings don't prove a causal effect of aspirations on investment and wealth.

3.3 Relating Aspirations, Expectations, and Other Psychological Concepts

In our sample, aspirations are related to, but not the same as, standard economic measures of expectations, which elicit what respondents believe will occur. Our measures of aspirations capture something close to the upper limits of outcomes respondents believe are possible. The average respondent has assets and income aspirations at respectively the 71st and 66th percentiles of their individual expectations distribution. Only 30% of respondents have aspirations above the maximum of their expectations distribution and only 15% have aspirations below the mean of their expectations distribution. Patterns are similar for the level of children's education: 72% of respondents have equal aspirations and expectations and 21% of respondents have higher aspirations than expectations.

In our sample, respondents report aspiring to modest improvements to their current position. Aspirations are higher than respondents' estimates of their current assets and income, but not dramatically, with median ratios of aspirations to current levels of 3 and 4.2 respectively.[14] Respondents' aspirations are lower than their beliefs about the highest level of income and assets someone from their village could reach in ten years: the median ratio of own aspirations to this "village maximum" is 0.5 for assets and 0.4 for income.

[12] We also estimate the aspirations-investment and aspirations-wealth relationships using splines to allow for nonlinear relationships. We cannot reject the null hypothesis of linearity for any relationship. Some papers have hypothesised or documented inverse-U-shaped aspirations-investment relationships, which they interpret as high aspirations motivating higher investment and very high aspirations leading to frustration and lower investment (Genicot and Ray, 2017; Janzen et al., 2017). We find no evidence of an inverse-U-shaped relationship in our data.

[13] We only collected cognitive ability for a subsample to use as a robustness check because it is unlikely to respond to a short psychological intervention with adults (Heckman et al., 2006). We find that cognitive ability is positively associated with wealth and investment, echoing the finding from Laajaj and Macours (2021) in the same context.

[14] Respondents' summary estimates of their current assets and income are, in turn, similar to measures of their current assets and income we construct from their detailed survey responses. The median ratio of respondents' direct summary reports to the measures we construct is 0.75 for assets and 0.93 for income.

Aspirations are not simply close correlates of other mechanisms we measure: they have pairwise correlations ≤ 0.1 with all other psychological concepts besides expectations (Table A.10). Regressing aspirations on all the other psychological concepts explains only 2% of the variation in aspirations and a LASSO estimation approach gives similar conclusions (Table A.11).

4 Experimental Design and Interventions

Descriptive patterns in this sample suggest that low aspirations and expectations for future living standards might constrain investment and wealth accumulation. We hence run an experiment to test if a psychological intervention that boosts aspirations also increases investment and wealth. We randomly assigned each of the 415 villages to one of four treatments: a placebo workshop, the psychologically active workshop ("psych"), an unconditional cash transfer and placebo workshop ("cash"), or an unconditional cash transfer and psychologically active workshop ("combined"). See Appendix B for randomisation details and confirmation it achieves balance.

We use this four-armed design to make three comparisons. First, we compare the psychological intervention and placebo groups. This comparison captures the effect of the psychologically active elements of the psychological intervention, holding constant the effects of common elements in both the psychological and placebo interventions, such as information or interaction with facilitators. Second, we compare the groups receiving the psychological and and cash interventions to test if the psychological intervention changes investment and living standards by an economically important magnitude relative to a widely used anti-poverty intervention.[15] Third, we compare the cash and combined groups to capture the effect of the psychological intervention in the presence of the large wealth shock delivered by the cash transfer.

The psychological and placebo interventions are delivered to the "primary woman" in each household, who also completes our surveys. The cash transfer intervention is delivered to the mobile money account of one household member, chosen jointly by the adults in the household. In 85% of households in our sample, this is the primary woman. We study the allocation of resources between household members in a separate paper (Mahmud et al., 2022). All interventions are delivered in the same month to households in each geographic location.

The workshops involved two ten-minute videos shown on a laptop, followed by facilitated exercises, lasting a median of 33 minutes, for a total workshop time of roughly 60–90 minutes.[16] Work-

[15] Participants in the cash group also receive the placebo intervention. So this comparison identifies the effects of the psychological elements of the intervention relative to the cash transfer, holding constant any effects of the common elements of the psychological intervention and placebo intervention.

[16] The interquartile range of the duration of exercises is 16–47 minutes. The psychological and placebo videos are posted at https://mbrg.bsg.ox.ac.uk/aspirations-and-goal-setting-video-intervention. Appendix C contains a summary of the psychological intervention videos and exercise script.

shops were administered either individually or in randomly assembled groups of three to four people.

4.1 Psychological Intervention

We developed a unique psychological intervention to achieve three aims: to raise participants' aspirations, while keeping them attainable; to encourage them to form specific long- and short-term plans to achieve these aspirations; and to encourage them to take actions in the present to begin working towards their aspirations.

Aspirational Role Model Videos: The videos narrate the life stories of two fictional women from a similar area to the participants and from poor backgrounds. They succeeded in improving their socio-economic position by forming ambitious but attainable aspirations and working toward them over a number of years despite obstacles. For example, in one story, Judy and her husband Oyoo aspire to own a house with an iron roof and send their children to higher levels of education. In the medium term, they plan to sell vegetables, for Judy to train to be a tailor, and to save more. Judy struggles to learn to sew and her business faces heavy competition, but eventually she establishes the business and achieves her aspirations. The second story, of Josefine, describes her journey from being a day labourer to continuing her education. The characters also "model" the techniques for goal-setting and planning that the participants then learn in the exercises and link their use to achieving one's long-term goals or aspirations. The videos were filmed by a production company based on our scripts, with ordinary people who were paid to be amateur actors, near the study location. They were pilot-tested in focus groups and edited based on feedback.

We use life stories because health research suggests watching a relevant other "modelling" behaviour makes viewers more engaged and changes attitudes more than merely receiving information (La Ferrara, 2016; Murphy et al., 2014). Social learning theory in psychology suggests role models are particularly influential in the formation of aspirations because people form aspirations with reference to the outcomes of other similar individuals (Bandura, 1977). We make narratives inspirational and emotional, aiming to provide a "vicarious experience" (Bandura, 1977): a resonant and influential experience of the life of another similar person. Importantly, the characters are shown making modest progress over many years and facing difficulties and disappointments. This draws on psychology research which finds role models are most inspiring when people believe their success is attainable and are shown the process for achieving success (Lockwood and Kunda, 1997; Marx and Ko, 2012).

Best Possible Selves: After the videos, participants did a series of exercises. In the first exercise, participants are asked to think about and describe their lives in five years "after everything has gone as well as it possibly could" to inspire setting of aspirations (King, 2001). They are encouraged to think of "the realisation of their dreams" where they have "reached their

full potential". Fieldworkers recorded the goals participants described.

Personal Goal-setting: In the second exercise, fieldworkers relay participants' goals back to them. Participants select their most important and achievable goal and report how long it would take to achieve it (Morisano et al., 2010). Appendix A.2 describes the diverse topics of these goals: 78% of respondents report at least one goal related to farming (mostly raising a specific type of livestock) and the shares are 86% for housing, 50% for education, and 37% for non-farm enterprises. We did not promote goals in any specific domain.

Mental Contrasting, Implementation Intentions, and Anticipation of Obstacles: In a third exercise, participants formulate "implementation intentions": a concrete step they could take in the next week towards their goal (Gollwitzer and Sheeran, 2006). They identify potential obstacles to their goal and strategies to overcome these obstacles (Duckworth et al., 2013). They engage in "mental contrasting:" identifying how their lives would be improved and how they would feel if they achieve their aspiration (Oettingen and Gollwitzer, 2010).[17] We provide a reminder calendar, which shows the characters and sayings from each video, on which participants placed a sticker which reminded them of their goals.

We drew on local life stories about successful people we collected from focus groups in villages near the study area to write the video and exercise scripts. The intervention also draws elements from other psychological interventions. We reviewed the psychological literature, mainly searching the American Psychological Association's PsycNet database, to identify short exercises that targeted one of the goals of the intervention. Each of the exercises we use has, on its own, had effects on lab tasks, intentions, healthy behaviours or education effort after one to two months in richer countries (Cross and Sheffield, 2019; Duckworth et al., 2013; Loveday et al., 2018). Some related techniques are taught in schools in high-income countries (EEF, 2021a,b; US Department of Education, 2007) but our participants are unlikely to have been exposed to them. Goal-setting and planning and were included in the Kenyan primary school life skills curriculum in 2017 (KICD, 2017), but our participants left school before the curriculum was introduced and only 43% completed primary school.

4.2 Placebo Intervention

The placebo intervention includes two videos showing and describing all activities in the psychologically active video. They describe farming in rural Kenya and education and types of work, mirroring the information available in the stories of Judy and Josefine. We made these videos by including at least one shot of every scene and character from the psychological video but

[17] In our pre-analysis plan, we stated the videos have elements to encourage a growth mindset (Dweck, 2012). However, subsequent work suggests growth mindset interventions must be highly specific and focus on specific ideas about neuroplasticity of intelligence (Yeager and Dweck, 2020), which our interventions do not.

cutting out the "psychologically active" components – descriptions of characters or development of their stories, shots of people conveying emotion, or music – and replacing dialogue or narrative voiceovers with factual descriptions of the content of the scene. This is followed by exercises of similar length to the psychological techniques in which respondents recall and discuss the videos' content, accuracy and quality but without teaching the Best Possible Selves, goal-setting, mental contrasting, implementation intentions or anticipation of obstacles techniques. The placebo design rules out that effects are caused by participants learning about the economic activities from the videos. It also rules out effects of being selected for an intervention by people from outside the community or interacting with a facilitator, as this occurs in both interventions. Finally, it rules out the effects of meeting with a group, as both interventions have the same group structure.

4.3 Cash Transfers

We compare the benefits and costs of the psychological intervention to a large, lump-sum unconditional cash transfer, delivered by our implementation partner GiveDirectly, of 2,237 USD PPP, equal to, respectively, 59% of mean annual consumption expenditure and 146% of mean non-land asset value in our sample. The transfer was delivered in three tranches of 203, 1,017 and 1,017 USD PPP one month apart. We describe the transfer in detail in Appendix C. This specific cash transfer programme is a useful benchmark for the psychological intervention for four reasons. First, the same programme delivered in neighbouring areas affects investment and wealth, the same outcomes targeted by the psychological intervention (Haushofer and Shapiro, 2016; Egger et al., 2022). Second, cash transfers are increasingly used as benchmarks for evaluating other anti-poverty policies by researchers and policymakers (Banerjee et al., 2018; McIntosh and Zeitlin, 2021, 2022; Sedlmayr et al., 2020). Third, this cash transfer programme does not enforce any specific economic behaviours, like the psychological intervention. GiveDirectly emphasises that the transfer is unconditional and at endline only 3% of households report thinking that there were any conditions attached to the transfer. Finally, the transfer is large but, like the psychological intervention, it could be administered at scale. For example, the Kenyan government's Hunger Safety Net Programme pays out the equivalent of the GiveDirectly transfer in 21 months (Kenya National Social Protection Secretariat, 2022).

4.4 Rates of Treatment Receipt

Approximately 90% of endlined households in each of the four groups receive the psychological or placebo intervention they are assigned and no differences between groups are large or statistically significant (Table B.3). Approximately 80% of endlined households in the cash and combined groups receive cash transfers and the difference between the groups is small and not statistically significant.

We discuss the implications of the difference between the rates of receiving cash and receiving the psychological intervention in Section 6. We report statistics on intervention receipt in Appendix B.

4.5 Estimation and Inference

We estimate models of the form

$$Y_{iv} = \text{Cash}_v \cdot \beta_C + \text{Psych}_v \cdot \beta_P + \text{Combined}_v \cdot \beta_{CP} + \mathbf{X_{iv}} \cdot \mathbf{\Gamma} + \epsilon_{iv}, \quad (1)$$

where i and v index individuals and villages. Y_{iv} is the post-treatment outcome of interest measured at endline. Cash_v, Psych_v, and Combined_v are indicators for assignment to respectively cash transfers, the psychological intervention, and the combined intervention. Hence β_{CP} measures the combined effect of both interventions relative to neither, not the interaction effect. $\mathbf{X_{iv}}$ contains prespecified covariates and sublocation fixed effects.[18] The covariates make little difference to the estimated treatment effects but lower the estimated standard errors. We report heteroskedasticity-robust standard errors clustered by village, the unit of treatment assignment. We control the false discovery rate across multiple tests by reporting sharpened q-values for the effect of each treatment and for tests of equality of each pair of treatment effects (Benjamini et al., 2006). We run all analysis at the household level except some individual-level education analyses in the appendices.

Our estimation and inference methods and outcome measures are prespecified at https://www.socialscienceregistry.org/trials/996. We make a few departures from the pre-analysis plan to improve comparability across economic aggregates and remove components with high measurement error. We list these in Appendix F and highlight two here. First, we adjust for multiple testing across the six economic aggregates and across all the main psychological mechanisms. This is more conservative than the prespecified adjustment, which was only across the prespecified components of each aggregate or index. Second, we summarise results by constructing a non-prespecified inverse covariance-weighted average of the six economic aggregates, following Anderson (2008).

5 Effects on Economic Outcomes

Table 2 shows treatment effects of the three interventions on six prespecified economic aggregates. To help interpret these results, we also discuss treatment effects on some components of these aggregates and show the component results in Appendix A. None of the interventions affect household size or composition, including fertility.

[18] Sublocations are administrative units containing on average 10 contiguous villages. The prespecified covariates are month-of-endline fixed effects (to account for seasonality); the baseline values of Y_{iv}; baseline household size, asset value, a self-beliefs index (made up of locus of control and self-efficacy scales, defined in Appendix G), respondent education, respondent age; and an indicator for the endline being answered by a proxy respondent.

Effects of the Psychological Intervention: The psychological intervention increases multiple types of investment (Table 2, columns 2-4). It increases households' annual labour supply by 27 days per year (5% of the placebo mean) and respondent labour supply by 5% of the placebo mean, with no effect on school-aged household members' labour supply. It also increases annual expenditure on intermediate inputs for home production and hired labour by 230 USD PPP (27% of the placebo mean) and increases some measures of education investment. Annual household-level education spending increases by 22 USD PPP (3.5% of the placebo mean, not significant). A prespecified breakdown by age group shows that this is driven by an 11%, statistically significant, rise in spending per primary school-aged child (Table A.1). The treatment effect on expenditure for post-primary school-aged children is close to zero, potentially because secondary school fees are larger and lumpier (Jack and Habyarimana, 2018; Lucas and Mbiti, 2012). Treatment effects on both enrolment and attendance are negligible, perhaps in part because these are high at all ages.

The psychological intervention increases annual revenue by 260 USD PPP: 12% of the placebo mean (Table 2, column 5). The labour supply, input expenditure and revenue effects are concentrated in non-farm enterprises: households supply 9 days more labour per year to these enterprises, spend 174 USD PPP more on inputs, earn 284 USD PPP more revenue, and earn 109 USD PPP more profit (Table A.2).[19] Effects on revenue, inputs and labour supply in non-farm enterprises are all large and statistically significant – 8% of the placebo group mean for labour supply and roughly 35% for the other outcomes. The effect on profit is also large but less precisely estimated than the other outcomes ($p=0.12$). The psychological intervention also increases adoption of new business practices in non-farm enterprises, such as new/improved products and new customers/markets (Table A.2, column 4). In contrast, crop- and livestock-raising contribute little to the treatment effects on input expenditure and revenue, even though there are two growing seasons between treatment and endline. The fact that investment and revenue effects are concentrated in the same non-farm enterprises is consistent with a causal chain from treatment-induced investment to revenue.

The psychological intervention raises non-land asset value by 98 USD PPP: 6% of the placebo group mean (Table 2, column 6). The increase in asset stocks is mostly explained by small, non-lumpy items – durable assets and cash savings (Tables A.3). Effects on livestock are driven by small livestock, not lumpy purchases like large livestock that would be difficult to afford without a direct wealth infusion. There are changes in savings behaviour: respondents join more savings groups, used as commitment savings devices in this area (Gugerty, 2007), and make more contributions to them.

[19]Breaking down effects by type of household economic activity is prespecified. All measures in Table A.2 are prespecified except returns to factors and profit. We code activity-specific measures as zeroes for households that do not engage in that activity to avoid sample selection problems. For example, we code non-farm enterprise investment and revenue as zero for households without non-farm enterprises. Hence treatment effects for the roughly half of the sample who own an enterprise are even larger.

Table 2: Treatment Effects on Economic Behaviour

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Index Components					
	Economic Index	Labour Supplied (Days)	Inputs & Hired Labour	Education Expenditure	Revenue	Non-Land Assets	Consumption Expenditure
Psych	0.112***	26.8**	230**	22.2	260*	98**	142*
	(0.035)	(11.6)	(100)	(27.7)	(155)	(46)	(74)
	[.]	[0.069]	[0.069]	[0.127]	[0.069]	[0.069]	[0.069]
Cash	0.234***	27.2**	451***	44.8	465***	406***	322***
	(0.039)	(12.4)	(103)	(30.4)	(159)	(50)	(77)
	[.]	[0.012]	[0.001]	[0.036]	[0.003]	[0.001]	[0.001]
Combined	0.258***	9.0	653***	126.4***	546*	352***	232**
	(0.063)	(11.5)	(214)	(31.5)	(303)	(47)	(95)
	[.]	[0.096]	[0.004]	[0.001]	[0.030]	[0.001]	[0.012]
P: cash	0.003	0.972	0.040	0.503	0.230	0.000	0.032
= psych	[.]	[0.526]	[0.073]	[0.433]	[0.208]	[0.001]	[0.073]
P: cash	0.717	0.127	0.357	0.022	0.801	0.292	0.367
= combined	[.]	[0.466]	[0.581]	[0.150]	[0.788]	[0.581]	[0.581]
P: psych	0.004	0.118	0.007	0.001	0.217	0.000	0.327
= combined	[.]	[0.097]	[0.010]	[0.003]	[0.151]	[0.001]	[0.196]
P: cash	0.166	0.007	0.881	0.166	0.536	0.025	0.049
+ psych	[.]	[0.046]	[0.416]	[0.143]	[0.332]	[0.066]	[0.081]
= combined							
Placebo mean	0.000	525	857	640	2,101	1,529	3,796
# clusters	413	413	413	412	413	413	412
# obs	7,243	7,240	7,243	6,273	7,243	7,242	7,224

Notes: This table shows household-level treatment effects of the interventions on six prespecified economic outcomes and an index combining them. All currency values are measured in 2018 USD PPP. All outcomes except education expenditure are in annual terms. The outcome in column (1) is an inverse covariance-weighted average of the outcomes in columns (2)-(7), following Anderson (2008). The outcomes in columns (2)-(7) use the same definitions as in Table 1. Coefficients are from OLS regressions of each outcome on a vector of treatment assignments, the baseline outcome, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent and prespecified baseline covariates. For each outcome variable, we report the coefficients of interest and heteroskedasticity-robust standard errors, clustered at village level, in parentheses. Sharpened q-values controlling for the false discovery rate across all outcomes except the summary index are shown in brackets. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively. The number of observations varies slightly across columns because some respondents don't answer all questions, which also drops one small village.

The psychological intervention increases consumption expenditure by 142 USD PPP, 4% of the placebo mean (Table 2, column 7), typically interpreted as a rise in household welfare. This might be financed by the higher revenue generated by the treatment-induced rise in investment. The intervention raises both consumption and input expenditure, but it shifts the composition

of spending from consumption toward investment. The placebo group's mean ratio of input expenditure to consumption is 0.22, while the ratio of the treatment effects on input expenditure to consumption is 1.6. Hence the marginal resources generated by treatment are allocated more to investment than consumption, relative to the non-treatment allocation. This is consistent with psychological intervention's emphasis on working and investing towards long term aspirations.

These results provide strong evidence that the psychological intervention substantially shifts economic investment and outcomes. The intervention shifts a summary index combining of the six aggregates by 0.11 standard deviations (Table 2, column 1). Effects on five of the six economic aggregates are statistically significant, even after adjusting for multiple hypothesis testing over the six outcomes, a more stringent adjustment than we prespecified. These are large effects for a roughly 80-minute intervention occurring 17 months earlier.

Effects of the Cash Transfer: The cash transfer increases our summary index of the six economic aggregates by 0.23 standard deviations, roughly twice the psychological intervention's effect (Table 2, column 1). The relative magnitudes of the cash transfer and psychological intervention effects vary substantially across the different economic aggregates. Both interventions have similar effects on annual household labour supply.[20] Given the large resource injection of the cash transfer, its effects on all pecuniary outcomes are unsurprisingly larger than the psychological intervention: roughly double the size for input expenditure, education expenditure, and revenue; triple for consumption; and quadruple for assets, although not all differences are statistically significant.

Effects of the Combined Intervention: The combined intervention increases all economic aggregates we study. But it does not have systematically larger effects than the cash transfer alone. In particular, it increases the economic index by 0.26 standard deviations, only 0.02 standard deviations more than the cash transfer alone (Table 2, column 1). This is potentially surprising, as the psychological intervention alone increases the economic index by a substantially larger 0.11 standard deviations. Across five of the six aggregates, the combined-cash difference is smaller than the psych-placebo difference (columns 2-3, 5-7). And, for the same five aggregates, the psychological intervention alone produces statistically significant increases even after adjusting for multiple hypothesis testing, while the cash-combined difference is not statistically significant. We discuss potential explanations for these patterns in Section 7.4.

Education expenditure is the only outcome where the effect of adding the psychological intervention to the cash transfer significantly exceeds the effect of the cash transfer alone (Table 2, column 4). The combined arm increases education spending by 127 USD PPP: 30% of the placebo group mean, six times the effect of the psychological intervention, and triple the effect of the

[20]Most cash transfer evaluations find non-negative labour supply effects (Banerjee et al., 2017).

cash transfer. This is driven by spending on school fees for primary- and secondary school-aged children (Table A.1). It is not accompanied by any shift in school enrolment or attendance.

Heterogeneous Treatment Effects of the Interventions: None of the three interventions produce treatment effects that vary substantially across different types of respondents, showing that these interventions produce relatively broad-based economic improvements. We show this in two ways. We describe both methods and show detailed results in Appendix A.1 and briefly summarise the results here. First, we estimate heterogeneous treatment effects on the summary economic index using standard treatment-interacted regressions across eight prespecified baseline measures – respondent age, marital status, education, aspirations, expectations, self-beliefs, household assets and household size – plus the economic index itself. The estimated interaction effects are seldom large and the fraction of statistically significant estimates is no larger than would arise by chance. In particular, the psychological intervention's effects do not vary by baseline wealth or by baseline values of the targeted psychological concepts, aspirations and expectations.

Second, we test the null hypothesis of constant treatment effects on the economic index using causal forests (Wager and Athey, 2018). This is a data-driven method to flexibly partition the sample into subsamples defined by values of the nine baseline measures listed above and estimate treatment effects within each subsample. This approach estimates heterogeneous treatment effects simultaneously by all nine measures and does not impose a specific functional form. We do not reject the null hypothesis that treatment effects are equal across all subsamples.

6 Cost-Benefit Comparisons

The previous section showed that the psychological intervention produced substantial effects relative to the cash transfer. In this section we quantitatively compare the costs and benefits of the two interventions. At the study's scale, the average benefit-cost ratio of the psychological intervention is at least twice that of the cash transfer intervention, using a measure of benefits based on consumption and asset accumulation. Under plausible assumptions, the psychological intervention's ratio of benefit to marginal cost is substantially higher.

Panel A of Table 3 shows intervention costs, calculated following the principles in J-PAL (2016). We first consider the average total cost of the intervention: the fixed cost of intervention development and variable programme and delivery costs, divided by the number of people offered treatment (row 1). Using this approach, the psychological intervention, cash transfer, and the combined intervention have average total costs of respectively 353, 2,149, and 2,358 USD PPP. The cash and psychological interventions are not cost-equivalent since the cash transfer amount

Table 3: Benefit-Cost Comparison of All Interventions

	Psych	Cash	Combined
Panel A: Costs			
I: Average total cost[a]	353	2,149	2,358
III: Average intervention delivery cost[c]	54.3	1,931	2,059
Panel B: Accumulated benefits			
Non-durable consumption[d]	202	267	155
Education expenditure[d]	22.2	44.8	126
Housing and land expenditure[f]	18.0	487	465
Non-land non-housing asset stock at endline [e]	98.0	406	352
Total benefits	340	1,205	1,098
Panel C: Benefit-cost ratios (%)[g]			
Total benefits/average total cost	96.4	56.1	46.6
Total benefits/average intervention delivery cost	627	62.4	53.3

Notes: We present all figures in USD PPP (October 2018). See Appendix D for more detailed cost breakdowns. (a) Costs I includes all costs of delivering and developing the interventions, averaged across participants within each treatment arm. (b) Costs II includes only the costs of the cash transfers, plus delivering the cash,s or of delivering the media intervention (e.g. field officer time and travel costs). This line is most relevant for comparisons between different programme options within an organisation, where programme set-up (e.g., targeting) is complete and overhead and administration costs will be constant across interventions. (c) Accumulated consumption and education expenditure impacts are drawn from intent to treat estimates, using annualised flow impact estimates scaled for each participant by the time between their receipt of the intervention and the endline survey. Education expenditure is coded as zero for households without school-aged children, while it is coded as missing elsewhere in the paper. (d) This measure of asset stocks is observed in the endline survey. (e) This captures household expenditure on land and house assets, including repair, maintenance and construction work since the intervention. (f) Total benefits/total cost ratios are the accumulated impacts divided by the cost of each treatment arm.

was fixed by GiveDirectly.[21] More details on cost calculations are in Appendix D.

Panel B of Table 3 shows our estimates of intervention benefits. We define benefits as gains in a broad interpretation of consumption, as consumption is often used as a measure of living standards, which our interventions aim to improve (Deaton, 1997). We follow the principles used for cost-effectiveness by Bandiera et al. (2017) and Banerjee et al. (2015), who also study asset transfer programmes targeted at poor households. They measure benefits due to the programme as the sum of treatment effects on accumulated consumption and education expenditure between treatment and the endline and on asset holdings at the endline. The latter proxies for treatment effects of the intervention on future consumption. To implement these principles, we add together four treatment effects: consumption expenditure, education expenditure, expenditure on land and housing (from Table A.3), and the value at endline of durable non-land non-housing assets. We exclude the value of spending on durables from the effect on consumption because it is included

[21]Differences in average total costs are common but not universal in benchmarking exercises. For example, comparisons of food parcels to cash transfers can have equivalent values of food/cash delivered to participants or equivalent programme costs but not both, because delivery costs differ.

in durable asset stocks. We scale the consumption and education effects from annual terms to cover the individual-specific period from the interventions to endline. We use a flow measure of expenditure on land and housing, asked directly about the period from the intervention to endline, rather than the stock because land markets are thin and respondents who have not recently purchased land or housing struggle to value it.

Using this definition, the psychological intervention, cash transfer and combined intervention deliver average benefits per participant of respectively 340, 1,205, and 1,098 USD PPP by endline. These are arguably conservative estimates of benefits. Following Bandiera et al. (2017), we use the asset stock's resale value at endline as a proxy for the assets' contribution to future consumption, although assets might deliver future returns higher than their resale value. We don't explicitly model the potential lifetime benefits of changes in education, but value education using expenditure. Our definition of benefits also excludes any (dis)utility of treatment-induced labour supply.

We show the benefit-cost ratio of each intervention in Panel C of Table 3. At the scale of this experiment, the benefits accumulated by endline are 96% of the total costs for the psychological intervention and 56% for the cash transfer.[22] The psychological intervention delivers remarkably large gains in living standards given its low cost. This compares very favourably to unconditional cash transfers, an existing anti-poverty intervention widely regarded as cost-effective.

We compare the cost of the psychological intervention to the cost of the cash transfer, excluding the cost of the placebo intervention. If the placebo intervention delivers any benefits, our approach attributes extra benefits to the cash transfer and hence understates the cost-effectiveness of the psychological intervention relative to the cash transfer. Take-up of the psychological intervention is higher than the cash transfer (90 versus 80%). But this has a limited effect on our comparison of the two interventions' benefit-cost ratios because lower cash take-up lowers both the numerator – because the average benefits are intention-to-treat effects – and the denominator – because the average costs are per person *offered* each intervention – relative to full take-up.

The psychological intervention may be more cost-effective at larger scales or using different delivery mechanisms. It has scope for considerable scale economies because the fixed cost of development is large relative to the variable costs of the programme and delivery. The variable programme cost could be further reduced by adding it to existing development interventions such as agricultural extension. In contrast, the cash transfer has limited scope for economies of scale because most of the cost is the transfer itself. As a simple illustration, we calculate the average

[22] Our definition of benefits excludes the value of investment spending, so it is unsurprising that the benefit we measure is lower than the value of the cash transfer. Another approach is to value the cash transfer programme benefits as the amount of cash given to recipients, which results in a total benefit to average total cost ratio of 90% for cash. We cannot apply this approach to the psychological intervention. But the ratio for the psychological intervention using our preferred approach, 96%, still compares favourably to the ratio for cash using this alternative approach.

variable cost of just *delivering* each intervention. This excludes the costs of developing each intervention, targeting participants, and administering the program (row 2 of Panel A). The ratio of accumulated benefit by endline to this average variable cost of delivering each intervention is 627% for the psychological intervention but much smaller, 62%, for the cash transfer (row 2 of Panel C).

The combined intervention is marginally less cost-effective than the cash transfer alone, as the benefits are (not statistically significantly) smaller while its costs are higher. From a policy evaluation perspective, these results suggest that this psychological intervention should be added to cash transfer programmes only if policymakers place high weight on education spending. However, the long-term benefits depend on the return to education spending, which we cannot reliably estimate.

7 Mechanisms

In this section, we provide evidence that aspirations and expectations are the most plausible mechanism for the economic effects of the psychological intervention and the limited effects of adding the intervention to cash. First, we show the psychological intervention increases aspirations and expectations but has no effects on the other candidate mechanisms. Second, we rule out some alternative non-psychological mechanisms. Third, we present a simple conceptual framework to define aspirations and expectations in terms of economic concepts and show theoretically how they might affect investment and wealth. Finally, we show that the cash and combined interventions have similar effects on all psychological mechanisms, including substantial positive effects on aspirations and expectations, and test possible explanations for this finding. The similar effects of the cash and combined interventions on psychological mechanisms are consistent with their similar effects on economic outcomes.

7.1 Effects of the Psychological Intervention on Psychological Mechanisms

Table 4 shows effects of the psychological intervention on the candidate mechanisms. We code all measures so that higher values are theoretically associated with higher investment: e.g. we report the effect on mental health (the Z score on a depression scale multiplied by negative one). We report sharpened q-values that adjust for multiple testing across all mechanisms. This is highly conservative, as we prespecified time and risk preferences and mental health as unlikely mechanisms that we measured only to rule them out.

Aspirations and Expectations: Treatment increases an index of aspirations for the future by 0.092 standard deviations relative to the placebo group (Table 4, column 1). This effect is broad-based: it is non-negative at all quantiles, with no clear pattern of larger effects at higher or lower quantiles, and the average effect does not vary substantially by baseline aspirations or

other baseline characteristics (Figures A.6 and A.7). The treatment effect on aspirations is driven by a positive and significant increase in aspirations for children's education; there are positive but noisily estimated increases in aspirations for assets and income (Table A.7). The magnitude of changes is modest: for example, respondents aspire to have 265 USD PPP more assets and 127 USD PPP more annual income in ten years (Table A.7). Abstracting from inflation, this is approximately double the treatment effect on assets at endline (Table A.3) and similar to the treatment effect on profits (Table A.2). We focus on effects on the whole index, which we view as a measure of a general aspirational mindset. Differences across domains of aspirations may be due to differential effects of the workshop on different types of aspirations or the fact that the measure of education aspirations covers a period further into the future.

Treatment also increases an index of expectations for participants' future outcomes by 0.091 standard deviations relative to the placebo group (Table 4, column 2). The increases across domains, patterns across quantiles and lack of heterogeneous effects mirror those for aspirations. The similarity of aspirations and expectations effects suggests that participants view their higher aspirations as attainable. The aspirations and expectations effects are statistically significant at the 1 and 5% levels respectively. They remain significant when we adjust for multiple hypothesis testing across all eight mechanisms, while no effect on any other mechanism is large or statistically significant.

Self-Beliefs: Treatment has no effect on the index of three psychological scales capturing self-beliefs (Table 4, column 3) or on any of the individual scales (Table A.6). This is in line with studies in psychology, which find little empirical evidence that the individual psychological exercises change self-beliefs (Conroy and Hagger, 2018; Kwasnicka et al., 2013). While self-efficacy or similar traits can be altered, this is largely documented with more intensive interventions (Carlana et al., 2022; Ghosal et al., 2020; McKelway, 2021). Evidence on growth mindset suggests it mainly responds to specific interventions different from those we implement (Yeager and Dweck, 2020).

Information Acquisition, Beliefs about Returns, Mimicry or Recall: We find little evidence that the intervention changes beliefs about returns to specific activities. There is no effect on beliefs about returns to labour on the farm or to investment in university education, which we measured because the videos contain characters making these investments (Table A.4). This is unsurprising: we minimise the potential for learning about the returns to activities from the intervention by comparing the psychological intervention to the placebo group, who receive the same images and descriptions of economic activities. We also find no effect on beliefs about returns to fertiliser, an input not featured in the videos, nor on an index of all three beliefs (Table 4, column 4). We don't measure beliefs about returns to all possible investment activities, given

the number of activities featured in the videos.[23]

We also test if respondents simply copy specific activities depicted in the psychological and placebo videos. The psychological intervention does not increase an index of five dummy variables capturing if respondents mimic activities shown in the videos: weaving baskets, keeping savings in jar, taking a sewing class, training as a teacher, or growing vegetables for market sale (Table A.5, column 2). This is perhaps because the activities were common – 28% of placebo group respondents engage in at least one.

Finally, we examine if people are more likely to recall or pay attention to information about activities in the psychological intervention because it is more entertaining (Hanna et al., 2014; La Ferrara, 2016). In both the psychological and placebo videos, we include information about returns to education in neighbouring counties of Western Kenya, taken from Ozier (2018). The psychological intervention does not improve recall of this information relative to the placebo (Table A.5, column 1). Thus, while we cannot measure beliefs or recall over all information in the video, we find no direct evidence for a framework based on information acquisition, mimicry of particular activities or improved recall of information.

Time and Risk Preferences: The psychological intervention has negligible effects on risk preferences and on two proxies for patience over monetary payments: the discount factor and share of non-present biased people (Table 4, columns 5-7). We prespecified that we did not expect changes in these preferences and other light-touch interventions also find no effects (John and Orkin, 2022).

Mental Health: The psychological intervention has a very small, non-significant effect of 0.015 standard deviations on mental health (Table 4, column 8) and near-zero effects on indicators for whether an individual meets criteria for clinical depression (Table A.6). We prespecified changes in depression as unlikely, as the intervention is shorter than most therapies, which take 6 to 20 sessions, and omits many key elements of common therapies (Cuijpers et al., 2013).

7.2 Ruling Out Alternative Non-Psychological Mechanisms

Experimenter Demand: Participants in the psychological intervention may wish to give answers consistent with the behaviours they saw in the videos or what they reported in the exercise, and hence misreport their economic activities or outcomes. Evidence suggests this does not drive effects. First, we collect one objective economic measure. Enumerators manually check for the presence of seven easy-to-observe assets at the end of the survey, a measure we prespecified. The psychological

[23]We also measured beliefs about the profits from being a tailor at baseline and find treatment effects on investment do not vary by baseline beliefs about returns to tailoring, providing some evidence against these beliefs driving investment behaviour. We dropped the measure at endline because many respondents could not answer the question as they were unfamiliar with revenue and cost in this activity.

Table 4: Psychological Intervention Effects on Psychological Mechanisms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Aspir-ations Index	Expect-ations Index	Self-belief Index	Returns Index	Discount Factor	No Present Bias	Risk-taking	Mental Health Z-score
Psych	0.092***	0.091**	0.006	-0.011	-0.006	-0.023	-0.018	-0.015
	(0.035)	(0.039)	(0.046)	(0.037)	(0.011)	(0.016)	(0.038)	(0.035)
	[0.079]	[0.079]	[1.000]	[1.000]	[1.000]	[0.428]	[1.000]	[1.000]
Cash	0.130***	0.178***	-0.053	-0.002	-0.002	-0.015	-0.014	0.086**
	(0.036)	(0.038)	(0.042)	(0.038)	(0.011)	(0.016)	(0.036)	(0.034)
	[0.002]	[0.001]	[0.344]	[0.907]	[0.907]	[0.526]	[0.854]	[0.027]
Combined	0.178***	0.134***	0.025	-0.002	-0.003	-0.028*	-0.030	0.044
	(0.040)	(0.043)	(0.044)	(0.036)	(0.011)	(0.017)	(0.033)	(0.035)
	[0.001]	[0.006]	[0.668]	[0.908]	[0.788]	[0.235]	[0.594]	[0.339]
P: cash	0.324	0.041	0.198	0.830	0.728	0.617	0.927	0.006
= psych	[0.948]	[0.170]	[0.657]	[1.000]	[1.000]	[1.000]	[1.000]	[0.052]
P: cash	0.258	0.331	0.043	0.997	0.935	0.402	0.696	0.207
= combined	[1.000]	[1.000]	[0.528]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
P: psych	0.027	0.318	0.687	0.813	0.781	0.768	0.777	0.097
= combined	[0.277]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[0.515]
P: cash	0.398	0.019	0.224	0.825	0.728	0.652	0.958	0.581
+ psych	[1.000]	[0.180]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
= combined								
Placebo mean	0.000	0.000	0.000	0.000	0.698	0.753	0.000	0.000
# clusters	413	413	413	413	413	413	413	413
# obs	7,232	7,233	7,221	7,110	7,243	7,243	7,170	7,213

Notes: This table shows treatment effects of the psychological intervention on psychological outcomes that might explain the treatment effects on economic outcomes. The aspirations index and psychological measures are defined in the footnote below Figure 1. Most outcomes are inverse covariance-weighted averages of multiple measures so the treatment effects are in standard deviation units, following Anderson (2008). Measures are defined in detail in Section 3.1. Coefficients are from an OLS regression of each outcome on a vector of treatment assignments, the baseline outcome, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent and prespecified baseline covariates. For psychological outcomes where the order in which questions appear in the survey are randomised, a set of order indicator variables are also included as additional controls. For each outcome variable, we report the coefficients of interest and heteroskedasticity-robust standard errors, clustered at village level, in parentheses. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively. The number of observations varies slightly across columns because some respondents don't answer all questions, which also drops one small village.

treatment effect on self-reported and enumerator-observed measures of these assets are almost identically zero (Table A.5, columns 3-4). The psychological intervention also does not induce experimenter demand effects when combined with the cash transfer, which enables asset acquisition. The combined treatment arm raises the number of self-reported assets by 0.93 and field officer-verified assets by 1.17. Second, as discussed above, respondents in the psychological intervention group are not disproportionately likely to report investing through the specific activities shown in the video, which might occur if they were reporting in a way they thought would please enumerators.

Spillovers: Our village-level randomisation reduces spillover effects relative to designs that randomise within villages, schools, or even within community/social groups. We also find people talk very little about the issues raised in the psychological intervention, and even less across villages. People talk infrequently with others about their goals and challenges – fewer than two on average in the last 12 months – and this discussion is concentrated within villages (Garlick et al., 2022). We test for spillovers *within* villages in a random sample of households in psychological treatment and placebo villages who are too wealthy to receive the cash transfer on GiveDirectly's criteria. Spillover effects of the psychological intervention on the main psychological or economic outcomes are mostly small and none are statistically significant after correction for multiple hypothesis testing (Tables A.8 and A.9). The lack of within-village spillovers means that spillovers onto the placebo group, who live farther away in separate villages, unlikely. However, we view these result with some caution as untreated respondents within the same village may be affected by knowledge that others received interventions.

Group/facilitator effects: As discussed in Section 4.1, the placebo-controlled design means that interaction with outsiders or meeting with a group should not drive the psychological intervention's effects. We also see no evidence that group delivery or composition drives results: group size is balanced across treatment arms, under half of respondents report at endline that they still talk to their group members, and treatment effects do not vary by facilitators' assessment of groups' likelihood of meeting in the future. Our main findings are robust to using facilitator fixed effects.

Hawthorne or John Henry Effects: In experimental settings where the control group are aware others receive a different treatment, they could change reporting or behaviour because they are disappointed or believe the treatment group have received information that advantages them (Barrett and Carter, 2010). Alternatively, the treatment group may believe they have particular attributes which have led to their selection. Our village-level design largely rules out potential effects from knowledge that others had received a different intervention.

7.3 A Conceptual Framework Linking Psychological and Economic Outcomes

We propose a simple framework to show how the psychological intervention can activate a causal chain from aspirations, which proxy for reference points for future consumption, to investment and other economic outcomes. We also discuss how expectations can enter into this framework.

Our framework is one of reference-dependent utility, where people receive an extra utility gain, an "aspirational payoff", from achieving or exceeding their reference point for consumption. In each period t, agents derive utility from leisure l_t and consumption c_t relative to a reference point h_t, akin to aspirations, which they take as given:

$$u(c_t,l_t)=v(c_t,l_t)+z(c_t-h_t), \tag{2}$$

This follows existing models of aspirations (Dalton et al., 2016; Genicot and Ray, 2017; Lybbert and Wydick, 2018).[24] We assume v is increasing and concave in both arguments, as is standard. We also assume people gain utility at a decreasing rate from exceeding their reference point and lose utility from missing it ($z'>0,\ z''<0,\ \text{sign}\{z(c-h)\}=\text{sign}\{c-h\}$).[25] Utility is additively separable through time with discount factor δ.

People enter period t with asset holdings a_t and receive revenue y_t, which they allocate between consumption and capital investment k_t. They allocate their time endowment T between leisure and labour e_t. Their investments generate revenue $y_{t+1}=f(k_t,e_t)$ in the next period. Assets depreciate at rate γ each period. We assume revenue is increasing and concave in both arguments.

This problem has choice variables consumption and leisure, yielding first-order conditions:

$$v_{c_t}+z_{c_t}=\delta\cdot(1+f_{k_t}-\gamma)\cdot\left(v_{c_{t+1}}+z_{c_{t+1}}\right) \tag{3}$$

$$v_{l_t}=\delta\cdot f_{e_t}\cdot(v_{c_{t+1}}+z_{c_{t+1}}). \tag{4}$$

The first condition is an Euler equation: people set the marginal utility of current consumption equal to the discounted marginal utility of future consumption arising from current capital investment. The second condition is the labour-leisure trade-off: people set the marginal utility of current leisure equal to the discounted marginal utility of future consumption arising from current work. In both conditions, the marginal utility of consumption includes a term $z_{c_{t+1}}$ capturing reference dependence.

This framework shows how the psychological intervention, by raising aspirations for the future h_{t+1}, can generate the estimated treatment effects on economic outcomes. We do not formalise a model of endogenous aspirations formation. Higher h_{t+1} increases the marginal utility of future

[24] We assume that aspirations are malleable. But we follow the literature by assuming that agents cannot choose their reference point, otherwise they would set $h_t=0$ to maximise utility (Lybbert and Wydick, 2018).

[25] Other models use slightly different assumptions about the shape of z to study different behaviour, including loss aversion over gambles (DellaVigna et al., 2017; Kahneman and Tversky, 1979). These behaviours are less relevant for our study so we assume z is concave and continuous for tractability.

consumption $v_{c_{t+1}}+z_{c_{t+1}}$. Formally, this occurs because z is a concave function of c_t-h_t. Intuitively, this occurs because raising h_{t+1} shifts agents farther below their consumption reference point, making gains in future consumption more valuable. The higher marginal utility of future consumption motivates people to invest now to afford higher future consumption: they raise current capital investment to maintain condition (3) and raise current labour supply to maintain condition (4). This matches the positive treatment effects on input expenditure and on labour supply that we observe in the experiment. This causes higher future revenue and asset values through $f(k_t,e_t)$, matching the positive treatment effects on revenue and assets we observe. The framework does not explicitly model investment in or revenue from education. But, informally, the psychological intervention will increase current education expenditure for the same reasons that it increases current capital investment and labour supply, matching the positive effects on investment in younger children's education.

The effect of an increase in aspirations on consumption at the time of the endline survey is theoretically ambiguous. There is an intertemporal substitution effect: there is more incentive to invest now and consume in the future, so the time path of consumption will steepen and consumption in periods close to t will fall. But there is also a wealth effect, as the rise in current investment and labour supply increases future assets. Which of these effects dominates in any specific period depends on the parameterisation of the model, as in Deaton (1992). The positive treatment effect we estimate on consumption after 17 months suggests a large role for the wealth effect.

Expectations can enter this framework in two different ways. First, expectations may be another proxy for the reference points. In Kőszegi and Rabin (2006) and Kőszegi and Rabin (2007), expectations are themselves reference points: they both determine economic choices, and also reflect beliefs about the outcomes of economic choices, leading to complex equilibrium concepts when expectations are formed endogenously. Second, aspirations might proxy for reference points, be affected by the psychological intervention, and drive changes in economic choices and outcomes, while expectations capture beliefs about the outcomes of these changes in economic choices. In the first interpretation, expectations respond directly to the psychological intervention and then affect economic outcomes. In the second interpretation, expectations respond indirectly to the psychological intervention via effects on economic outcomes. The close empirical relationship between aspirations and expectations in our data is consistent with both interpretations, so we do not argue that either one is more plausible than the other.

In Appendix E, we show how other psychological mechanisms could enter this framework and show that increases in patience, self-beliefs, or beliefs about returns could theoretically produce the same economic effects. Treatment does not affect these other mechanisms, so we conclude it does not work through these mechanisms. We could not do this using only the economic effects.

7.4 Effects of the Cash and Combined Intervention on Psychological Mechanisms

We now turn to potential explanations for the similar economic effects of the combined cash-and-workshop intervention and the cash alone, despite the substantial effects of the psychological intervention alone. Aspirations and expectations are a very plausible mechanism for this pattern. We showed above that the psychological intervention alone seems to affect economic outcomes by increasing aspirations and expectations. However, the combined intervention and cash transfer alone produce similar effects on aspirations and expectations: the cash transfer alone increases aspirations and expectations by 0.13 and 0.18 standard deviations respectively, while the combined intervention increases aspirations and expectations by 0.18 and 0.13 standard deviations respectively (Table 4, columns 1-2).[26]

Aspirations can also explain the one clear difference between the cash and combined interventions: the combined intervention has substantially larger effects than cash on education *expenditure*. This can be explained by the combined intervention having substantially larger effects than cash on education *aspirations* (Table A.7). The two arms' effects on asset and income aspirations are more similar to each other.

Both the cash and combined interventions have similar, near-zero effects on other psychological mechanisms (columns 3-7) other than mental health. Both interventions improve the mental health index, in line with Ridley et al. (2020), although only the cash effect is statistically significant (column 8).

This pattern of results is consistent with our conceptual framework. The psychological intervention activates a causal chain from reference points, proxied by aspirations and expectations, to investment. It activates this chain by itself but does not when added to a cash transfer, potentially because the cash transfer itself activates the chain and crowds out the effect of the psychological intervention. This crowd-out highlights the importance of understanding reference point formation. We leave a full study of this to future work but evaluate two possible explanations for this crowd-out.

The first explanation is a **targeting** effect: people's reference points might simply be less sensitive to external shocks if they are wealthier. To show this, we augment our treatment effects model in equation (1) to include a proxy for baseline economic resources and interactions between this proxy and the treatment indicators. Under the targeting explanation, the psychological intervention should be less effective for households with more resources, so the interaction between resources and the psychological intervention should be negative for most economic outcomes.

[26]Multiple studies show that *conditional* cash transfers increase aspirations (Fruttero et al., 2021). But these effects may be due to conditionality-induced exposure to teachers and healthcare workers rather than effects of the cash (Chiapa et al., 2012; Macours and Vakis, 2014). Our focus on *unconditional* cash transfers avoids this ambiguity. Stutzer (2004) shows positive associations between income and aspirations for income.

Instead, this interaction is essentially zero for all economic outcomes and all three resource proxies we use: the value of all assets, assets excluding the (difficult to measure) value of land and housing, and annual consumption (Table A.12).[27] Given this result, we reject the targeting explanation.

The second possible explanation is a **windfall** effect: the windfall or unanticipated nature of the cash transfer might itself increase reference points, proxied by aspirations, more than an equivalent amount of anticipated wealth. This could reduce the psychological intervention's effects when combined with the cash transfer. The targeting and windfall explanations differ because the latter concerns unanticipated, rapid increases in wealth, while the former concerns wealth in general, which might be anticipated and gradually accumulated. The two explanations have different policy implications: targeting implies that the psychological intervention should be offered to poorer people; windfall implies that the psychological intervention should not be offered to people receiving cash transfers or other wealth shocks.

We find some evidence consistent with a windfall explanation using two tests based on responses to windfall versus other wealth. The first test compares the share of total expenditure allocated to investment between the placebo and cash groups. A standard life-cycle model predicts that this share will be higher in the cash group, as unanticipated wealth raises lifetime income and hence present consumption by more than a change in anticipated wealth would (Deaton, 1992). Instead, we find the opposite pattern. The cash treatment effect on this share is 3.1 percentage points with standard error 0.6, compared to a placebo group mean of 19.2% (Table A.13, column 1). This finding is not due to differences in total expenditure between the groups: controlling for total expenditure, the cash group invests 1.8 percentage points more of their expenditure than the placebo group (column 2); and including expenditure*treatment interactions shows that the cash group's higher investment share does not substantially vary by total expenditure (column 3).[28] This pattern is consistent with a reference-dependent utility model in which windfall wealth increases the reference point for future consumption and hence increases the share of expenditure invested.

The second test compares the relationship between aspirations and wealth for windfall versus non-windfall wealth. Figure A.4 compares the estimated effect of the cash transfer on aspirations to estimates from regressing the aspirations index on three different wealth proxies using placebo group data with different sets of controls. Under the windfall explanation, the effect of one dollar

[27]There is enough variation in baseline economic resources in the psychological group that this exercise does not rely on extrapolation outside the support of the data. The 2,237 USD PPP cash transfer corresponds to 0.18 standard deviations (SD) of baseline assets, 1.5 SDs of baseline non-land-non-housing assets, and 0.84 SDs of baseline consumption. We obtain a similar result using spline models that allow more flexible heterogeneous treatment effects. In addition, the effects of all three interventions do not vary by respondents' baseline aspirations.

[28]The regressions that control for expenditure might have endogeneity problems but the results are robust to including or excluding the prespecified baseline covariates and to trimming the tails of total expenditure to reduce sensitivity to possible outliers.

of cash transfer on aspirations (column 4) would be larger than a change in aspirations associated with a one unit increase in naturally occurring wealth (columns 1-3). This pattern clearly holds for one wealth proxy: the total value of all assets. This is the most comprehensive wealth proxy we observe but the most difficult to measure. Evidence is mixed for the other two proxies: the values of consumption and non-land assets. Figure A.5 shows similar patterns for expectations.

We conclude that the psychological intervention's substantial economic effects by itself and modest economic effects when combined with a cash transfer most likely occur because the psychological intervention and cash transfer both shift the same mechanism: reference points, proxied by aspirations or expectations. This pattern does not occur because the psychological intervention has smaller effects on wealthier households. It might occur because windfall resources raise aspirations and expectations and hence raise investment more than a standard wealth effect would predict, crowding out the aspiration-promoting effect of the psychological intervention itself. Our findings raise the possibility that cash transfers can shift economic outcomes through both conventional wealth effects and behavioural effects.

8 Conclusion

We study the role of material and psychological constraints to investment and asset accumulation for people living in poverty in Western Kenya using a four-arm randomised experiment. We provide empirical microfoundations to the idea that people living in poverty may lack opportunities to build their "capacity to aspire" (Appadurai, 2013): to set higher aspirations to improve their socioeconomic position and plan to achieve these goals. The idea is important because psychological constraints could be one channel through which adverse historical conditions of discrimination and segregation persist and perpetuate contemporary poverty (Durlauf, 1996). That these mechanisms are psychological does not imply that the disadvantaged are morally responsible for their situation. Rather, understanding these mechanisms can improve the design of interventions to target poverty.

We provide compelling evidence that aspirational capacity can be easily and cost-effectively built. We show that a short, scalable workshop targeted at psychological constraints has large effects on economic outcomes 17 months later. The workshop teaches people to set higher aspirations and medium-term goals; plan concrete, immediate steps to begin working toward those goals; and anticipate obstacles. It takes one session and does not require specially qualified facilitators. In a cluster-randomised, placebo-controlled trial in 415 villages, we find the workshop increases recipients' economic investment and downstream outcomes: they work and invest more, leading to higher revenue, asset wealth, and consumption. We show the intervention likely works by increasing participants' aspirations and expectations: the long-term goals they choose and aim

for, and their beliefs about what future is possible.

The gains in consumption and assets are achieved at least twice as cost-effectively as the gains from the large unconditional cash transfer we provide in another experimental arm. The psychological intervention might become even more cost-effective at larger scales. This comparison highlights that simple psychological interventions should be taken seriously as a tool for poverty reduction. There are potential market failures in the building of aspirational capacities. For example, there is no market where good role models can be compensated for the aspirational benefits they produce and institutions like schools or workplaces may not have incentive to develop these capacities. Further work could study how the intervention works at scale and in different contexts and whether changes persist. We also focus here on individuals, but further work could study the potential for social interactions to reinforce shifts in aspirations.

We find a complex set of results when we combine the psychological intervention and cash transfer in another experimental arm. Investment, revenue, consumption and assets are no higher in this arm than the cash transfer arm, although education expenditure is substantially higher. This result might occur because the cash transfer alone raises participants' aspirations, crowding out the aspiration- and investment-promoting effect that the psychological intervention delivers by itself. Such windfall effects may occur partly because the cash transfer is so large and delivering it in a lump sum enhances recipients' beliefs about what they can achieve. Indeed, future research could consider how transfer size structure affects aspirational as well as economic benefits. However, our work highlights that simple existing anti-poverty interventions can have important psychological benefits that might deliver long-term changes beyond their material benefits.

References

Alan, S., T. Boneva, and S. Ertac (2019): "Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit," *The Quarterly Journal of Economics*, 134, 1121–1162.

Alan, S. and S. Ertac (2018): "Fostering Patience in the Classroom: Results from a Randomized Educational Intervention," *Journal of Political Economy*, 126, 1865–1911.

Andersen, S., G. Harrison, M. Lau, and E. Rutström (2008): "Eliciting Risk and Time Preferences," *Econometrica*, 76, 583–618.

Anderson, M. (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 103, 1481–1495.

Andresen, E., J. Malmgren, W. Carter, and D. Patrick (1994): "Screening for Depression in Well Older Adults: Evaluation of a Short Form of the CES-D," *American Journal of Preventive Medicine*, 10, 77–84.

Angelucci, M. and D. Bennett (2021): "The Economic Impact of Depression Treatment in India," *Institute for the Study of Labour (IZA) Discussion Paper*, 14393.

Anzagi, S. and F. Bernard (1977): "Population Pressure in Rural Kenya," *Geoforum*, 8, 63–68.

Appadurai, A. (2013): *The Future as a Cultural Fact: Essays on the Global Condition*, London, UK: Verso Books.

Ashraf, N., G. Bryan, A. Delfino, E. Holmes, L. Iacovone, and A. Pople (2022): "Learning to See the World's Opportunities: The Impact of Imagery on Entrepreneurial Success," *London School of Economics & Political Science, Working Paper*.

Bandiera, O., R. Burgess, N. Das, S. Gulesci, I. Rasul, and M. Sulaiman (2017): "Labor Markets and Poverty in Village Economies," *The Quarterly Journal of Economics*, 132, 811–870.

Bandura, A. (1977): *Social Learning Theory*, Englewood Cliffs, New Jersey: Prentice-Hall.

Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Parienté, J. Shapiro, B. Thuysbaert, and C. Udry (2015): "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries," *Science*, 348, 12607991–126079916.

Banerjee, A., R. Hanna, G. Kreindler, and B. Olken (2017): "Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programs," *World Bank Research Observer*, 32, 155–184.

Banerjee, A., D. Karlan, D. Osei, H. Trachtman, and C. Udry (2018): "Unpacking a Multi-Faceted Programme to Build Sustainable Income for the Very Poor," *National Bureau of Economic Research Working Paper*, 24271.

Banerjee, A., D. Karlan, H. Trachtman, and C. Udry (2020): "Does Poverty Change Labour Supply? Evidence from Multiple Income Effects and 115,579 Bags," *National Bureau of Economic Research Working Paper*, 27314.

Baranov, V., S. Bhalotra, P. Biroli, and J. Maselko (2020): "Maternal Depression, Women's Empowerment, and Parental Investment: Evidence from a Large Randomised Control Trial," *American Economic Review*, 10, 824–859.

Barker, N., G. Bryan, D. Karlan, A. Ofori-Atta, and C. Udry (2022): "Cognitive Behavioral Therapy Among Ghanas Rural Poor is Effective Regardless of Baseline Mental Distress," *American Economic Review: Insights*, 4, 527–545.

Barrett, C. and M. Carter (2010): "The Power and Pitfalls of Experiments in Development Economics: Some Non-Random Reflections," *Applied Economic Perspectives and Policy*, 32, 515–548.

Beaman, L., E. Duflo, R. Pande, and P. Topalova (2012): "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India," *Science*, 335, 582–586.

Beck, A. and B. Alford (2009): *Depression: Causes and Treatment*, Philadelphia: University of Pennsylvania Press.

Benabou, R. (1996): "Equity and Efficiency in Human Capital Investment: The Local Connection," *The Review of Economic Studies*, 63, 237–264.

Benjamini, Y., A. Krieger, and D. Yekutieli (2006): "Adaptive Linear Step-up Procedures that Control the False Discovery Rate," *Biometrika*, 93, 491–507.

Bernard, T., S. Dercon, K. Orkin, and A. Seyoum Taffesse (2014): "The Future in Mind: Aspirations and Forward-looking Behaviour in Rural Ethiopia," *University of Oxford:, Working Paper*.

Bernard, T. and A. S. Taffesse (2014): "Aspirations: An Approach to Measurement with Validation Using Ethiopian Data," *Journal of African Economies*, 23, 189–224.

Bhat, B., J. de Quidt, J. Haushofer, V. Patel, G. Rao, F. Schilbach, and P. Vautrey (2022): "The Long-Run Effects of Psychotherapy on Depression, Beliefs, and Economic Outcomes," *National Bureau of Economic Research Working Paper*, 30011.

Blattman, C., S. Chaskel, J. Jamison, and M. Sheridan (2022): "Cognitive Behavior Therapy Reduces Crime and Violence over 10 Years: Experimental Evidence," *University of Chicago, Working Paper*.

Blattman, C., J. Jamison, and M. Sheridan (2017): "Reducing Crime and Violence: Experimental Evidence from Cognitive Behavioural Therapy in Liberia," *American Economic Review*, 107, 1165–1206.

Bong, M. and E. Skaalvik (2003): "Academic Self-Concept and Self-Efficacy: How Different Are They Really?" *Educational Psychology Review*, 15, 1–40.

Bossuroy, T., M. Goldstein, B. Karimou, D. Karlan, H. Kazianga, W. Parienté, P. Premand, C. Thomas, C. Udry, J. Vaillant, and K. Wright (2022): "Tackling Psychosocial and Capital Constraints to Alleviate Poverty," *Nature*, 605, 291–297.

Boswell-Dean, E., F. Schilbach, and H. Schofield (2018): "Poverty and Cognitive Function," in *The Economics of Poverty Traps*, ed. by C. Barrett, M. Carter, and J.-P. Chavas, Chicago, IL: University of Chicago Press, 57–118.

Caliendo, M., D. Cobb-Clark, and A. Uhlendorff (2015): "Locus of Control and Job Search Strategies," *The Review of Economics and Statistics*, 97, 88–103.

Campos, F., M. Frese, M. Goldstein, L. Iacovone, H. Johnson, D. McKenzie, and M. Mensmann (2017): "Teaching Personal Initiative Beats Traditional Training in Boosting Small Business in West Africa," *Science*, 357, 1287–1290.

Carlana, M., E. La Ferrara, and P. Pinotti (2022): "Goals and Gaps: Educational Careers of Immigrant Children," *Econometrica*, 90, 1–29.

Cecchi, F., A. Garcia, R. Lensink, and B. Wydick (2022): "Aspirational Hope, Dairy Farming Practices, and Milk Production: Evidence from a Randomized Controlled Trial in Bolivia," *World Development*, 160, 106087.

Chiapa, C., J. Garrido, and S. Prina (2012): "The Effect of Social Programmes and Exposure to Professionals on the Educational Aspirations of the Poor," *Economics of Education Review*, 31, 778 – 798.

Conroy, D. and M. Hagger (2018): "Imagery Interventions in Health Behaviour: A Meta-Analysis," *Health Psychology*, 37, 668–679.

Cross, A. and D. Sheffield (2019): "Mental Contrasting for Health Hehaviour Change: A Systematic Review and Meta-analysis of Effects and Moderator Variables," *Health Psychology Review*, 13, 209–225.

Cuijpers, P., M. Sijbrandij, S. Koole, G. Andersson, A. Beekman, and C. Reynolds III (2013): "The Efficacy of Psychotherapy and Pharmacotherapy in Treating Depressive and Anxiety Disorders: A Meta-Analysis of Direct Comparisons," *World Psychiatry*, 12, 137–148.

Dalton, P., S. Ghosal, and A. Mani (2016): "Poverty and Aspirations Failure," *Economic Journal*, 126, 165–188.

Dalton, P., J. Ruschenpohler, and B. Zia (2018): "Determinants and Dynamics of Business Aspirations: Evidence from Small-Scale Entrepreneurs in an Emerging Market," *World Bank Policy Research Working Paper*, 8400.

De Quidt, J. and J. Haushofer (2018): "Depression Through the Lens of Economics: A Research Agenda," in *The Economics of Poverty Traps*, ed. by C. Barrett, M. Carter, and J. Chavas, Chicago, IL: University of Chicago Press, 127–152.

Deaton, A. (1992): *Understanding Consumption*, Oxford UK: Oxford University Press.

——— (1997): *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Baltimore: Johns Hopkins University Press.

Deaton, A. and S. Zaidi (2002): *Guidelines for Constructing Consumption Aggregates for Welfare Analysis*, vol. 135, World Bank Publications.

Delavande, A. (forthcoming): "Expectations in Development Economics," in *Handbook of Economic Expectations*, Elsevier.

Delavande, A., X. Giné, and D. McKenzie (2011): "Eliciting Probabilistic Expectations with Visual Aids in Developing Countries: How Sensitive are Answers to Variations in Elicitation Design?" *Journal of Applied Econometrics*, 26, 479–497.

DellaVigna, S., A. Lindner, B. Reizer, and J. F. Schmieder (2017): "Reference-Dependent Job Search: Evidence from Hungary," *The Quarterly Journal of Economics*, 132, 1969–2018.

Dominitz, J. and C. Manski (1997): "Using Expectations Data to Study Subjective Income Expectations," *Journal of the American Statistical Association*, 92, 855–867.

Duckworth, A., T. Kirby, A. Gollwitzer, and G. Oettingen (2013): "From Fantasy to Action Mental Contrasting With Implementation Intentions (MCII) Improves Academic Performance in Children," *Social Psychological and Personality Science*, 4, 745–753.

Durlauf, S. (1996): "A Theory of Persistent Income Inequality," *Journal of Economic Growth*, 1, 75–93.

Dweck, C. (2012): "Mindsets and Human Nature: Promoting Change in the Middle East, the Schoolyard, the Racial Divide, and Willpower," *American Psychologist*, 67, 614–622.

Eble, A. and M. Escueta (2022): "When Bootstraps Aren't Enough: Demand, Supply, and Learning in a Very Low-Income Context," *Annenberg Institute at Brown University, Working Paper*.

Eckel, C. and P. Grossman (2002): "Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk," *Evolution and Human Behaviour*, 23, 281–295.

EEF (2021a): "Aspirations Interventions. Teaching and Learning Toolkit," https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/aspiration-interventions, online; accessed 7 September 2022.

——— (2021b): "Metacognition and Self-regulation. Teaching and Learning Toolkit," https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/metacognition-and-self-regulation, online; accessed 7 September 2022.

Egger, D., J. Haushofer, E. Miguel, P. Niehaus, and M. Walker (2022): "General Equilibrium Effects of Cash Transfers: Experimental Evidence from Kenya," *Econometrica*, 90, 2603–2643.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018): "Global Evidence on Economic Preferences," *The Quarterly Journal of Economics*, 133, 1645–1692.

FAO (1996): *A System of Economic Accounts for Food and Agriculture*, Rome: Food and Agriculture Organization.

Figlio, D., P. Giuliano, U. zek, and P. Sapienza (2019): "Long-Term Orientation and Educational Performance," *American Economic Journal: Economic Policy*, 11, 272–309.

Fruttero, A., N. Muller, and O. Calvo-Gonzalez (2021): "The Power and Roots of Aspirations: A Survey of the Empirical Evidence," *World Bank Policy Research Working Paper*, 9729.

Galor, O. and Ö. Özak (2014): "The Agricultural Origins of Time Preference," *American Economic Review*, 106, 3064–3103.

Garlick, R., G. Nyeki, and K. Orkin (2022): "Evaluating Multidimensional Programs in The Presence of Endogenous, Multidimensional Networks," *Duke University, Working Paper*.

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

——— (2020): "Aspirations and Economic Behavior," *Annual Review of Economics*, 12, 715–746.

Ghosal, S., S. Jana, A. Mani, S. Mitra, and S. Roy (2020): "Sex Workers, Self-Image and Stigma: Evidence from Kolkata Brothels," *The Review of Economics and Statistics*, 104, 431–448.

Gollwitzer, P. M. and P. Sheeran (2006): "Implementation Intentions and Goal Achievement: A Metaanalysis of Effects and Processes," *Advances in Experimental Social Psychology*, 38, 69–119.

Goux, D., M. Gurgand, and E. Maurin (2017): "Adjusting Your Dreams? High School Plans and Dropout Behavior," *The Economic Journal*, 127, 1025–1046.

Grosh, M. and P. Glewwe (2000): *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study*, Washington, D.C.: World Bank.

Gugerty, M. (2007): "You Can't Save Alone: Commitment in Rotating Savings and Credit Associations in Kenya," *Economic Development and Cultural Change*, 55, 251–282.

Guiso, L., P. Sapienza, and L. Zingales (2006): "Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives*, 20, 23–48.

Guyon, N. and E. Huillery (2021): "Biased Aspirations and Social Inequality at School: Evidence from French Teenagers," *The Economic Journal*, 131, 745–796.

Hanna, R., S. Mullainathan, and J. Schwartzstein (2014): "Learning Through Noticing: Theory and Evidence from a Field Experiment," *The Quarterly Journal of Economics*, 129, 1311–1353.

Haushofer, J. and J. Shapiro (2016): "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya," *The Quarterly Journal of Economics*, 131, 1973–2042.

Heckman, J. and T. Kautz (2012): "Hard Evidence on Soft Skills," *Labour Economics*, 19, 451–464.

Heckman, J., J. Stixrud, and S. Urzua (2006): "The Effects of Cognitive and Non-Cognitive Abilities on Labour Market Outcomes and Social Behaviour," *Journal of Labour Economics*, 24, 411–482.

Heller, S., A. Shah, J. Guryan, J. Ludwig, S. Mullainathan, and H. Pollack (2017): "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago," *The Quarterly Journal of Economics*, 132, 1–54.

J-PAL (2016): "J-PAL Costing Guidelines," https://www.povertyactionlab.org/sites/default/files/research-resources/costing-guidelines.pdf, online; accessed 7 September 2022.

Jack, W. and J. Habyarimana (2018): "High Hopes: Experimental Evidence on Saving and the Transition to High School in Kenya," *Georgetown University Initiative on Innovation, Development and Evaluation, Working Paper*.

Janzen, S., N. Magnan, S. Sharma, and W. Thompson (2017): "Aspirations Failure and Formation in Rural Nepal," *Journal of Economic Behavior and Organization*, 139, 1–25.

Jensen, R. and E. Oster (2009): "The Power of TV: Cable Television and Women's Status in India," *The Quarterly Journal of Economics*, 124, 1057–1094.

John, A. and K. Orkin (2022): "Can Simple Psychological Interventions Increase Preventive Health Investment?" *Journal of the European Economic Association*, 20, 1001–1047.

Kahneman, D. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–291.

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

Kaur, S., S. Mullainathan, S. Oh, and F. Schilbach (2021): "Do Financial Concerns Make Workers Less Productive?" *National Bureau of Economic Research Working Paper*, 28338.

Kenya National Social Protection Secretariat (2022): "National Safety Net Program," https://www.socialprotection.or.ke/social-protection-components/social-assistance/national-safety-net-program, online; accessed 7 September 2022.

KICD (2017): "Basic Education Curriculum Framework," https://kicd.ac.ke/wp-content/uploads/2017/10/CURRICULUMFRAMEWORK.pdf, online; accessed 7 September 2022.

King, L. (2001): "The Health Benefits of Writing About Life Goals," *Personality and Social Psychology Bulletin*, 27, 798–807.

Kipchumba, E. K., C. Porter, D. Serra, M. Sulaiman, et al. (2021): "Infuencing Youths' Aspirations and Gender Attitudes Through Role Models: Evidence from Somali Schools," *Texas A&M University, Working Paper*.

Kőszegi, B. and M. Rabin (2006): "A Model of Reference-Dependent Preferences," *The Quarterly Journal of Economics*, 121, 1133–1165.

——— (2007): "Reference-Dependent Risk Attitudes," *American Economic Review*, 97, 1047–1073.

Kwasnicka, D., J. Presseau, M. White, and F. Sniehotta (2013): "Does Planning How to Cope with Anticipated Barriers Facilitate Health-Related Behaviour Change? A Systematic Review," *Health Psychology Review*, 7, 129–145.

La Ferrara, E. (2016): "Mass Media and Social Change: Can We Use Television to Fight Poverty?" *Journal of the European Economic Association*, 14, 791–827.

——— (2019): "Presidential Address: Aspirations, Social Norms, and Development," *Journal of the European Economic Association*, 17, 1687–1722.

La Ferrara, E., A. Chong, and S. Duryea (2012): "Soap Operas and Fertility: Evidence from Brazil," *American Economic Journal: Applied Economics*, 4, 1–31.

Laajaj, R. and K. Macours (2021): "Measuring Skills in Developing Countries," *Journal of Human Resources*, 56, 1254–1295.

Lindqvist, E. and R. Vestman (2011): "The Labour Market Returns to Cognitive and Non-Cognitive Ability: Evidence from the Swedish Enlistment," *American Economic Journal: Applied Economics*, 3, 101–128.

Liu, E. (2013): "Time to Change What to Sow: Risk Preferences and Technology Adoption Decisions of Cotton Farmers in China," *The Review of Economics and Statistics*, 95, 1386–1403.

Locke, E. A. and G. P. Latham (2002): "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35 Year Odyssey," *American Psychologist*, 57, 705–717.

Lockwood, P. and Z. Kunda (1997): "Superstars and Me: Predicting the Impact of Role Models on the Self," *Journal of Personality and Social Psychology*, 73, 91–103.

Loury, G. (2009): *The Anatomy of Racial Inequality*, Cambridge, MA: Harvard University Press.

Loveday, P., G. Lovell, and C. Jones (2018): "The Best Possible Selves Intervention: A Review of the Literature to Evaluate Efficacy and Guide Future Research," *Journal of Happiness Studies*, 19, 607–628.

Lucas, A. and I. Mbiti (2012): "Access, Sorting and Achievement: The Short-Run Effects of Free Primary Education in Kenya," *American Economic Journal: Applied Economics*, 4, 226–253.

Lund, C., K. Orkin, M. Witte, T. Davies, J. Walker, J. Haushofer, S. Murray, J. Bass, L. Murray, and V. Patel (2022): "Treating Mental Health Conditions Improves Labor Market and Other Economic Outcomes in Low and Middle-Income Countries," *University of Oxford, Working Paper*.

Luttmer, E. and M. Singhal (2011): "Culture, Context, and the Taste for Redistribution," *American Economic Journal: Economic Policy*, 3, 157–79.

Lybbert, T. and B. Wydick (2018): "Poverty, Aspirations and the Economics of Hope," *Economic Development and Cultural Change*, 66, 709–753.

Macours, K. and R. Vakis (2014): "Changing Households' Investment Behaviour Through Social Interactions with Local Leaders: Evidence from a Randomised Transfer Programme," *Economic Journal*, 124, 607–633.

Mahmud, M., K. Orkin, and E. Riley (2022): "Economic and Psychological Constraints to Women's Empowerment," *University of Oxford, Working Paper*.

Marx, D. M. and S. J. Ko (2012): "Superstars Like Me: The Effect of Role Model Similarity on Performance Under Threat," *European Journal of Social Psychology*, 42, 807–812.

McIntosh, C. and A. Zeitlin (2021): "Cash Versus Kind: Benchmarking a Child Nutrition Program Against Unconditional Cash Transfers in Rwanda," *arXiv Preprint arXiv:2106.00213*.

——— (2022): "Using Household Grants to Benchmark the Cost Effectiveness of a USAID Workforce Readiness Program," *Journal of Development Economics*, 157, 102875.

McKelway, M. (2021): "Women's Employment in India: Intra-Household and Intra-Personal Constraints," *Dartmouth, Working Paper*.

McKenzie, D., A. Mohpal, and D. Yang (2022): "Aspirations and Financial Decisions: Experimental Evidence from the Philippines," *Journal of Development Economics*, 156, 102846.

Morisano, D., J. Hirsh, J. Peterson, R. Pihl, and B. Shore (2010): "Setting, Elaborating, and Reflecting on Personal Goals Improves Academic Performance," *Journal of Applied Psychology*, 95, 255–264.

Murphy, S., L. Frank, J. Chatterjee, and L. Baezconde-Garbanati (2014): "Narrative Versus Non-Narrative: The Role of Identification, Transportation and Emotion in Reducing Health Disparities," *Journal of Communication*, 63, 1–19.

Oettingen, G. and P. Gollwitzer (2010): "Strategies of Setting and Implementing Goals," in *Social Psychological Foundations of Clinical Psychology*, ed. by J. Maddux and J. Tangney, New York: The Guilford Press, 114–135.

Ozier, O. (2018): "The Impact of Secondary Schooling in Kenya: A Regression Discontinuity Analysis," *Journal of Human Resources*, 53, 157–188.

Porter, C. and D. Serra (2020): "Gender Differences in the Choice of Major: The Importance of Female Role Models," *American Economic Journal: Applied Economics*, 12, 226–54.

Rapport, N. (2016): "Aspiration as Universal Human Capacity," in *Voice and Matter: Communication, Development and the Cultural Return*, ed. by O. H. Oscar and T. Tufte, Gothenburg: Nordicom, 217–225.

Ridley, M., G. Rao, F. Schilbach, and V. Patel (2020): "Poverty, Depression, and Anxiety: Causal Evidence and Mechanisms," *Science*, 370, eaay0214.

Riley, E. (2019): "Role Models in Movies: the Impact of Queen of Katwe on Students' Educational Attainment," *The Review of Economics and Statistics*, 1–48.

Rojas Valdes, R., B. Wydick, and T. Lybbert (2022): "Can Hope Elevate Microfinance? Evidence from Oaxaca, Mexico," *Oxford Economic Papers*, 74, 236–264.

Rosenberg, M. and H. B. Kapland (1982): *Social Psychology of the Self-Concept*, Arlington Heights, Illinois: Harlan Davidson.

Ross, P. (2019): "Occupation Aspirations, Education Investment, and Cognitive Outcomes: Evidence from Indian Adolescents," *World Development*, 123, 104613.

Ross, P., P. Glewwe, D. Prudencio, and B. Wydick (2021): "Developing Educational and Vocational Aspirations Through International Child Sponsorship: Evidence from Kenya, Indonesia, and Mexico," *World Development*, 140, 105336.

Rotter, J. (1966): "Generalised Expectancies for Internal Versus External Control of Reinforcement," *Psychological Monographs: General and Applied*, 80, 1–28.

Schennach, S. (2022): "Measurement Systems," *Journal of Economic Literature*, 60, 1223–63.

Sedlmayr, R., A. Shah, and M. Sulaiman (2020): "Cash-plus: Poverty Impacts of Alternative Transfer-based Approaches," *Journal of Development Economics*, 144, 102418.

Spolaore, E. and R. Wacziarg (2013): "How Deep are the Roots of Economic Development?" *Journal of Economic Literature*, 51, 325–69.

Streufert, P. (2000): "The Effect of Underclass Social Isolation on Schooling Choice," *Journal of Public Economic Theory*, 2, 461–482.

Stutzer, A. (2004): "The Role of Income Aspirations in Individual Happiness," *Journal of Economic Behaviour and Organisation*, 54, 89–109.

US Department of Education (2007): "Encouraging Girls in Math and Science," https://ies.ed.gov/ncee/wwc/PracticeGuide/5, online; accessed 7 September 2022.

Wager, S. and S. Athey (2018): "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests," *Journal of the American Statistical Association*, 113, 1228–1242.

Willis, P. (1977): *Learning to Labour. How Working Class Kids Get Working Class Jobs*, New York: Columbia University Press.

Willis, W. J. (1987): *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*, Chicago: University of Chicago Press.

Yeager, D. and C. Dweck (2020): "What Can be Learned from Growth Mindset Controversies?" *American Psychologist*, 75, 1269–1284.

Appendices

A Additional Results

As we pre-specified, we adjust for multiple hypothesis testing over tests that use components of the same prespecified aggregate or concept. This applies to the two components of age-group-specific education expenditure in Table A.1; input expenditure and hired labour in Table A.2, the four components of non-land assets in Table A.3, the beliefs about returns to specific inputs in Table A.4, the three self-beliefs in Table A.6, the two depression indicators in Table A.6, and the three aspirations measures in Table A.7. We also adjust for multiple hypothesis testing across the six economic aggregates in each of Tables A.8, A.9, and A.12 for consistency with the approach we use in the paper. We do not adjust across all measures in a table if they are not components of the same prespecified aggregate.

Table A.1: Treatment Effects on Individual-level Education Expenditure

	(1)	(2)	(3)	(4)	(5)	(6)
	Education Expenditure		Fee Expenditure		Non-fee Expenditure	
Household members aged:	6-13	14-20	6-13	14-20	6-13	14-20
Psych	9.03*	2.3	7.67*	2.5	1.01	-1.55
	(5.01)	(14.8)	(3.95)	(13.2)	(1.56)	(3.75)
	[.]	[.]	[0.118]	[1.000]	[0.350]	[1.000]
Cash	9.56**	37.6**	7.50*	29.4*	2.15*	7.15*
	(4.66)	(18.8)	(3.86)	(16.6)	(1.29)	(4.17)
	[.]	[.]	[0.107]	[0.096]	[0.107]	[0.096]
Combined	13.29**	69.6***	11.18**	62.7***	1.74	7.76**
	(5.56)	(17.7)	(4.51)	(15.8)	(1.44)	(3.81)
	[.]	[.]	[0.029]	[0.001]	[0.130]	[0.022]
P: cash = psych	0.923	0.063	0.968	0.105	0.483	0.054
	[.]	[.]	[1.000]	[0.118]	[1.000]	[0.118]
P: cash = combined	0.502	0.130	0.411	0.075	0.785	0.888
	[.]	[.]	[1.000]	[0.177]	[1.000]	[0.799]
P: psych = combined	0.467	0.000	0.446	0.000	0.660	0.021
	[.]	[.]	[1.000]	[0.001]	[1.000]	[0.011]
P: cash + psych	0.458	0.236	0.485	0.168	0.498	0.700
= combined	[.]	[.]	[0.993]	[0.507]	[0.993]	[0.539]
Placebo mean	85.7	342	61.2	278	24.6	63.0
# clusters	411	408	410	408	411	408
# obs	12,003	8,528	11,936	8,447	11,936	8,437

Notes: All variables are at the individual level, with one observation for each household member aged 6-20. All currency values are measured in constant 2018 USD PPP. Education expenditure is the total expenditure on school-related fees and non-fee expenses during the current and preceding school years for each child in the relevant age group. Non-fee expenditure includes school related supplies (e.g. books) and uniforms. The number of clusters varies across columns because some small villages have no sampled households with members aged 6-13 or 14-20. The sample size is higher for total expenditure than either of fee or non-fee expenditure. If only one type of expenditure is missing for a household, we calculate that household's total education expenditure as the non-missing component times the sample mean ratio of total education expenditure over the non-missing expenditure component. Coefficients are from an OLS regression of each outcome on a vector of treatment assignments, the baseline outcome, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent, age, gender and prespecified baseline covariates. For each outcome variable, we report the coefficients of interest and heteroskedasticity-robust standard errors, clustered at village level, in parentheses. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively. Sharpened q-values controlling for the false discovery rate across outcomes within each family, namely each of the education measures for each age group, are shown in brackets.

Table A.2: Treatment Effects on Inputs to and Outputs from Farm and Non-farm Activities

	(1)	(2)	(3)	(4)	(5)	(6)
	Revenue	Input Expenditure	Hired Labour	Technology Adoption	Labour Supply (Days)	Returns to Factors
Panel A: Farm (Agriculture and Livestock)						
Psych	16.2	9.8	-0.6	0.088	15.7	7.0
	(47.6)	(9.3)	(15.1)	(0.103)	(9.8)	(18.2)
	[.]	[1.000]	[1.000]	[.]	[.]	
Cash	63.1	83.3***	27.8	0.209**	19.9*	-17.8
	(48.3)	(10.3)	(19.4)	(0.095)	(10.7)	(19.3)
	[.]	[0.001]	[0.083]	[.]	[.]	
Combined	20.0	93.4***	63.3***	0.152	2.9	-16.8
	(44.9)	(10.5)	(21.3)	(0.107)	(10.7)	(20.0)
	[.]	[0.001]	[0.002]	[.]	[.]	
P: cash	0.355	0.000	0.115	0.206	0.663	0.144
= psych	[.]	[0.001]	[0.062]	[.]	[.]	
P: cash	0.363	0.429	0.127	0.526	0.113	0.959
= combined	[.]	[0.341]	[0.341]	[.]	[.]	
P: psych	0.935	0.000	0.002	0.549	0.193	0.199
= combined	[.]	[0.001]	[0.002]	[.]	[.]	
P: cash	0.378	0.985	0.202	0.288	0.025	0.822
+ psych	[.]	[0.971]	[0.676]	[.]	[.]	
= combined						
Placebo mean	733	163	99.7	2.90	341	146
# clusters	413	413	413	413	413	413
# obs	7,242	7,242	7,243	7,235	7,240	7,232
	Revenue	Input Expenditure	Hired Labour	Technology Adoption	Labour Supply (Days)	Profits
Panel B: Non-farm						
Psych	284**	174**	16.8	0.084**	9.0*	109
	(136)	(87)	(14.3)	(0.035)	(5.3)	(70)
	[.]	[0.100]	[0.138]	[.]	[.]	
Cash	452***	305***	15.2	0.082**	18.2***	213***
	(139)	(92)	(16.9)	(0.036)	(6.0)	(72)
	[.]	[0.002]	[0.226]	[.]	[.]	
Combined	557**	442**	17.1	0.116***	13.4**	167
	(275)	(187)	(17.2)	(0.042)	(6.6)	(114)
	[.]	[0.039]	[0.190]	[.]	[.]	
P: cash	0.269	0.166	0.915	0.954	0.095	0.167
= psych	[.]	[0.498]	[0.843]	[.]	[.]	
P: cash	0.723	0.481	0.908	0.388	0.479	0.706
= combined	[.]	[1.000]	[1.000]	[.]	[.]	
P: psych	0.191	0.063	0.979	0.421	0.453	0.544
= combined	[.]	[0.144]	[0.958]	[.]	[.]	
P: cash	0.500	0.832	0.495	0.339	0.100	0.209
+ psych	[.]	[1.000]	[1.000]	[.]	[.]	
= combined						
Placebo mean	815	478	43.2	0.350	106	313
# clusters	413	413	413	413	413	413
# obs	7,241	7,082	7,243	7,243	7,240	7,031

Notes: All variables are at the household level and scaled to annual figures. All currency values are measured in constant 2018 USD PPP. Farm activities capture agricultural and livestock related production. Non-farm activities capture production in non-farm enterprises owned or operated by household members. Revenue, input expenditure, hired labour expenditure and labour supply are measured as in Table 1. Technology adoption for farm activities is a dummy equal to one if the household used one of 14 modern agricultural practices since intervention. For non-farm enterprises, it is one if during the last 12 months, an enterprise introduced new or improved products or services or went into a new market or accessed new customers. For farm enterprises, returns to factors of production is revenue minus expenditure on intermediate inputs and costs of renting in assets minus costs of hired labour. For non-farm enterprises, profits are revenue minus input expenditure and hired labour. Model specification is as in Table 2. For each outcome variable, we report the coefficients of interest and heteroskedasticity-robust standard errors, clustered at village level, in parentheses. Sharpened q-values controlling for the false discovery rate across outcomes within the family 'Inputs and Hired Labour' are shown in brackets. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Table A.3: Treatment Effects on Assets

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Non-land Asset Components					
	Non-land Assets	Durables	Livestock	Savings	Maize	Land Purchases	Housing Expenditures
Psych	98**	43*	25.4	26.4**	2.97	12.8	18
	(46)	(25)	(26.9)	(11.7)	(4.70)	(8.0)	(26)
	[.]	[0.144]	[0.299]	[0.113]	[0.359]	[.]	[.]
Cash	406***	236***	132.4***	29.8***	9.35*	34.6**	487***
	(50)	(31)	(27.7)	(10.1)	(5.26)	(14.0)	(34)
	[.]	[0.001]	[0.001]	[0.003]	[0.020]	[.]	[.]
Combined	352***	224***	77.0***	46.2***	7.92	43.8***	465***
	(47)	(27)	(26.5)	(16.2)	(5.51)	(10.6)	(33)
	[.]	[0.001]	[0.005]	[0.005]	[0.040]	[.]	[.]
P: cash = psych	0.000	0.000	0.000	0.787	0.267	0.062	0.000
	[.]	[0.001]	[0.001]	[0.553]	[0.217]	[.]	[.]
P: cash = combined	0.292	0.689	0.053	0.341	0.810	0.566	0.591
	[.]	[1.000]	[0.269]	[1.000]	[1.000]	[.]	[.]
P: psych = combined	0.000	0.000	0.067	0.204	0.390	0.006	0.000
	[.]	[0.001]	[0.112]	[0.158]	[0.243]	[.]	[.]
P: cash + psych	0.025	0.167	0.040	0.597	0.552	0.854	0.430
= combined	[.]	[0.335]	[0.192]	[0.502]	[0.502]	[.]	[.]
Placebo mean	1,529	765	576	122	65.6	9.22	196
# clusters	413	413	413	413	412	413	413
# obs	7,242	7,242	7,243	7,241	7,170	7,239	7,168

Notes: All variables are at the household level and scaled to annual figures. All currency values are measured in constant 2018 USD PPP. Non-land assets are made up of durable assets, livestock, savings and stocks of dried maize. Respondents estimate the value of household holdings of each asset if they were to sell them today in their current condition. For cash savings, we include savings in multiple places, as well as the payout households receive from ROSCAs to which they belong. Land purchases are the value of any compound and non-compound land that was purchased since the intervention. Housing expenditures includes the costs of any housing repairs, maintenance and construction since intervention. Coefficients are from an OLS regression of each outcome on a vector of treatment assignments, the baseline outcome, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent and prespecified baseline covariates. For each outcome variable, we report the coefficients of interest and heteroskedasticity-robust standard errors, clustered at village level, in parentheses. Sharpened q-values controlling for the false discovery rate across outcomes within the family of non-land non-housing assets are shown in brackets. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Table A.4: Treatment Effects on Beliefs About Returns to Specific Investments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Maize Farming					Education		
	Yields			Returns		Income		Returns
	Usual	w/ Fertiliser	w/ Labour	Fertiliser	Labour	Secondary	Degree	Degree
Psych	2.95	-6.65	-4.34	0.009	-0.036	-193.8	-392	0.041
	(6.91)	(12.57)	(9.81)	(0.077)	(0.031)	(140.9)	(341)	(0.139)
	[.]	[.]	[.]	[1.000]	[1.000]	[.]	[.]	[1.000]
Cash	3.50	15.40	-8.75	0.152*	-0.064**	-78.5	-200	0.126
	(7.15)	(13.85)	(9.41)	(0.078)	(0.030)	(138.8)	(315)	(0.146)
	[.]	[.]	[.]	[0.086]	[0.086]	[.]	[.]	[0.149]
Combined	1.59	-4.16	-10.96	0.025	-0.034	-67.9	-98	0.053
	(6.99)	(12.75)	(10.08)	(0.074)	(0.034)	(125.2)	(330)	(0.145)
	[.]	[.]	[.]	[1.000]	[1.000]	[.]	[.]	[1.000]
P: cash = psych	0.947	0.130	0.648	0.100	0.364	0.461	0.570	0.550
	[.]	[.]	[.]	[0.430]	[0.574]	[.]	[.]	[0.579]
P: cash = combined	0.798	0.159	0.808	0.093	0.338	0.933	0.745	0.581
	[.]	[.]	[.]	[0.387]	[0.510]	[.]	[.]	[0.633]
P: psych = combined	0.859	0.849	0.522	0.835	0.950	0.342	0.381	0.931
	[.]	[.]	[.]	[1.000]	[1.000]	[.]	[.]	[1.000]
P: cash + psych	0.617	0.483	0.874	0.190	0.118	0.255	0.273	0.542
= combined	[.]	[.]	[.]	[0.398]	[0.398]	[.]	[.]	[0.398]
Placebo mean	191	459	258	2.02	0.442	3,684	11,695	3.32
# clusters	413	411	410	411	410	410	410	410
# obs	7,115	6,420	6,147	6,420	6,147	5,137	5,137	5,137

Notes: Columns (1), (2), and (3) show the expected yield of maize with respectively no fertiliser and their current labour input, 50kg of fertiliser per acre and their current labour input, and no fertiliser and 12 hours of extra labour per week. The yields are for a one acre plot in the next long rains season, measured in gorogoro (a local unit approximately equal to 2 kilograms) per acre. The questions ask about DAP fertiliser, the most commonly used fertiliser in the region. Columns (4) and (5) show the implied returns from respectively extra fertiliser and extra labour, with a scale where 1 = 100% return. Columns (6) and (7) show the expected annual income for the respondent's child in two scenarios: complete secondary schooling with a KCSE certificate and a university degree. Column (8) shows the implied return to a university degree, with a scale where 1 = 100% return. The questions ask about the same child as the education aspirations and expectations questions, at age 30, with earnings expressed in 2018 USD PPP. Coefficients are from an OLS regression of each outcome on a vector of treatment assignments, the baseline outcome, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent and prespecified baseline covariates. For each outcome variable, we report the coefficients of interest and heteroskedasticity-robust standard errors, clustered at village level, in parentheses. Sharpened q-values controlling for the false discovery rate across the three returns measures are shown in brackets. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Table A.5: Treatment Effects on Information, Mimicry and Experimenter Demand

	(1)	(2)	(3)	(4)
			Asset Ownership	
	Information Recall	Mimicry of Videos	Self-reported	Field Officer Verified
Psych	-0.040	0.005	-0.027	0.06
	(0.025)	(0.042)	(0.352)	(0.41)
Cash	-0.016	-0.022	1.758***	1.58***
	(0.026)	(0.039)	(0.361)	(0.40)
Combined	-0.024	0.052	0.933***	1.17***
	(0.024)	(0.045)	(0.318)	(0.40)
P: cash = psych	0.401	0.529	0.000	0.000
P: cash = combined	0.759	0.098	0.012	0.322
P: psych = combined	0.531	0.311	0.004	0.008
P: cash + psych = combined	0.353	0.255	0.106	0.427
Placebo mean	0.274	-0.000	20.6	17.3
# clusters	409	413	413	413
# obs	6,536	7,235	7,243	7,243

Notes: Information recall is an indicator equal to one if respondents correctly recall information about the returns to education for Kenyan men from Ozier (2018), which appears in both the psychological and placebo video. This question is asked the same day that respondents watch the videos. Mimicry of videos is the standardised sum of indicator variables coded to one if the respondent engaged in any of the following activities at endline, all of which are featured in the videos: (a) weaved baskets; (b) kept savings in a jar; (c) attended a sewing class; (d) trained as a teacher; (e) grew vegetables to sell on the market. Field officer verified assets are objective measures of the number of seven durable assets – cooking pots and pans, jerry cans, chairs/sofa, tables, radios, TVs and poultry houses. The field officer asked to count these assets after completing the endline survey. The self-reported assets are the number of those specific assets reported in the assets module earlier in the survey. Coefficients are from an OLS regression of each outcome on a vector of treatment assignments, the baseline outcome, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent and prespecified baseline covariates. There is no multiple test correction as measures are not from families capturing the same concept. For each outcome variable, we report the coefficients of interest and heteroskedasticity-robust standard errors, clustered at village level, in parentheses. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Table A.6: Additional Treatment Effects on Psychological Mechanisms

	(1)	(2)	(3)	(4)	(5)	(6)
		Index Components				
	Self-beliefs Index	Self-Efficacy	Growth Mindset	Locus of Control	Depression Score >10	Depression Score >13
Psych	0.006	-0.138	0.305	-0.049	0.010	0.004
	(0.046)	(0.143)	(0.224)	(0.096)	(0.018)	(0.019)
	[.]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Cash	-0.053	-0.206	-0.032	-0.110	-0.031**	-0.038**
	(0.042)	(0.138)	(0.214)	(0.094)	(0.016)	(0.019)
	[.]	[0.568]	[0.568]	[0.568]	[0.050]	[0.050]
Combined	0.025	0.033	0.119	0.048	-0.030*	-0.024
	(0.044)	(0.145)	(0.206)	(0.097)	(0.016)	(0.018)
	[.]	[1.000]	[1.000]	[1.000]	[0.120]	[0.120]
P: cash = psych	0.198	0.628	0.131	0.552	0.021	0.023
	[.]	[0.720]	[0.646]	[0.720]	[0.024]	[0.024]
P: cash = combined	0.043	0.057	0.420	0.097	0.933	0.355
	[.]	[0.170]	[0.170]	[0.170]	[1.000]	[1.000]
P: psych = combined	0.687	0.221	0.391	0.355	0.020	0.110
	[.]	[0.642]	[0.642]	[0.642]	[0.043]	[0.059]
P: cash + psych	0.224	0.041	0.601	0.125	0.701	0.659
= combined	[.]	[0.720]	[0.646]	[0.720]	[0.024]	[0.024]
Placebo mean	0.000	24.1	23.1	17.4	0.658	0.462
# clusters	413	413	413	413	413	413
# obs	7,221	7,211	7,209	7,213	7,213	7,213

Notes: Notes: Column (1) is an Anderson (2008) index consisting of variables in columns (2), (3) and (4). The index definition includes all observations that have non-missing values for at least one component, using only the non-missing components in the averaging. The self-beliefs index is made up of growth mindset, self-efficacy, and locus of control scales. Self-efficacy is measured with the Schwarzer and Jerusalem (1995) scale. Growth mindset is measured with an adapted version of the 6-item Implicit Theories of Intelligence scale (Blackwell et al., 2007). Locus of control is measured using the Internal subscale from the Internal, Powerful Others and Chance (IPC) scale (Levenson, 1981). In Table 3, mental health is the 10-item CES-D depression scale from Andresen et al. (1994), multiplied by minus one. Here, we generate binary variables where individuals with scores at or above a threshold are identified as at high risk of depression or as experiencing psychological distress. This is how the score is used if it is used for clinical screening. Different studies in sub-Saharan Africa use thresholds of 13 (Baron et al., 2017) or 10 (Kilburn et al., 2016). Coefficients are from an OLS regression of each outcome on a vector of treatment assignments, the baseline outcome, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent and prespecified baseline covariates. For each outcome variable, we report the coefficients of interest and heteroskedasticity-robust standard errors, clustered at village level, in parentheses. Sharpened q-values controlling for the false discovery rate across outcomes within each family, namely self-beliefs and depression, are shown in brackets. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Table A.7: Treatment Effects on Aspirations

	(1)	(2)	(3)	(4)
		Index Components		
	Aspirations Index	Assets	Income	Education
Psych	0.092***	265	127	0.226***
	(0.035)	(565)	(216)	(0.067)
	[.]	[0.742]	[0.742]	[0.003]
Cash	0.130***	2,008***	611***	0.057
	(0.036)	(624)	(205)	(0.076)
	[.]	[0.005]	[0.005]	[0.178]
Combined	0.178***	2,511***	570**	0.130*
	(0.040)	(606)	(242)	(0.072)
	[.]	[0.001]	[0.020]	[0.030]
P: cash = psych	0.324	0.007	0.030	0.019
	[.]	[0.022]	[0.022]	[0.022]
P: cash = combined	0.258	0.457	0.874	0.292
	[.]	[1.000]	[1.000]	[1.000]
P: psych = combined	0.027	0.000	0.062	0.138
	[.]	[0.002]	[0.066]	[0.101]
P: cash + psych	0.398	0.788	0.602	0.105
= combined	[.]	[1.000]	[1.000]	[0.463]
Placebo mean	0.000	8,499	5,357	15.5
# clusters	413	413	413	410
# obs	7,232	7,204	7,185	6,102

Notes: Column (1) is an Anderson (2008) index consisting of variables in columns (2), (3) and (4). The index definition includes all observations that have non-missing values for at least one component, using only the non-missing components in the averaging. Income aspirations are the level of annual income that a household would like to reach at the end of the next ten years. Income is defined as all sources of cash income for the household, including earnings from production and transfers from any NGO or government programmes. Asset aspirations are the level of assets that the household would like to reach at the end of the next ten years, including their house, furniture, consumer goods and a transport vehicles. Income and assets are measured in constant 2018 USD PPP. Education aspirations are the aspirations for years of education attained by a randomly selected child, set to missing for households without children. Coefficients are from an OLS regression of each outcome on a vector of treatment assignments, the baseline outcome, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent and prespecified baseline covariates. For each outcome variable, we report the coefficients of interest and heteroskedasticity-robust standard errors, clustered at village level, in parentheses. Sharpened q-values controlling for the false discovery rate across outcomes within each family are shown in brackets. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Table A.8: Spillover Effects of Psychological Intervention on Economic Outcomes of Ineligible Households

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Index Components					
	Economic Index	Labour Supplied (Days)	Inputs & Hired Labour	Education Expenditure	Revenue	Non-Land Assets	Consumption Expenditure
Psych	-0.007	-8.20	62	-3.49	210	-31	-25.1
	(0.049)	(27.44)	(140)	(64.51)	(258)	(108)	(123.0)
	[.]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Placebo mean	0.000	549	1,053	851	2,573	2,264	3,839
# clusters	306	306	306	303	306	306	306
# obs	2,792	2,789	2,792	2,154	2,792	2,792	2,783

Notes: This table shows treatment effects of the psychological intervention on economic outcomes for ineligible households living in the study villages. This sample was only surveyed in a subset of study villages. Each of the economic variables is defined in Table 1. Coefficients are from an OLS regression of each outcome on a vector of treatment assignments, the baseline outcome, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent and prespecified baseline covariates. For each outcome variable, we report the coefficients of interest and heteroskedasticity-robust standard errors, clustered at village level, in parentheses. Sharpened q-values controlling for the false discovery rate across outcomes within each family are shown in brackets. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Table A.9: Spillover Effects of Psychological Intervention on Psychological Mechanisms of Ineligible Households

	(1)	(2)	(3)	(4)
	Self-belief Index	Aspirations Index	Expectations Index	Mental health Z-score
Psych	0.047	-0.010	-0.010	0.099*
	(0.068)	(0.051)	(0.065)	(0.055)
	[1.000]	[1.000]	[1.000]	[0.390]
Placebo mean	0.000	0.000	0.000	-0.000
# clusters	306	306	306	306
# obs	2,785	2,784	2,782	2,776

Notes: This table shows treatment effects of the psychological intervention on psychological outcomes for ineligible households living in the study villages. Each of the psychological variables are defined in Section 3.1. We did not measure time or risk preferences or beliefs about returns to investments in this sample. Coefficients are from an OLS regression of each outcome on a vector of treatment assignments, the baseline outcome, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent and prespecified baseline covariates. For each outcome variable, we report the coefficients of interest and heteroskedasticity-robust standard errors, clustered at village level, in parentheses. Sharpened q-values controlling for the false discovery rate across outcomes within each family are shown in brackets. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Table A.10: Pairwise Correlations of Psychological Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	AI	DF	NPB	RTI	S-bI	RI	MHZ	EI
Aspirations Index	1.00	-0.05	0.02	0.02	0.07	0.04	0.10	0.60
Discount Factor	-0.05	1.00	0.26	-0.00	-0.02	-0.02	-0.04	-0.05
No Present bias	0.02	0.26	1.00	0.01	-0.01	0.01	-0.00	0.01
Risk Taking Index	0.02	-0.00	0.01	1.00	0.02	0.03	-0.00	0.03
Self-belief Index	0.07	-0.02	-0.01	0.02	1.00	-0.07	0.09	0.04
Returns Index	0.04	-0.02	0.01	0.03	-0.07	1.00	0.02	0.01
Mental Health Z-score	0.10	-0.04	-0.00	-0.00	0.09	0.02	1.00	0.14
Expectations Index	0.60	-0.05	0.01	0.03	0.04	0.01	0.14	1.00

Notes: This table shows pairwise correlations between psychological outcomes. All measures are defined in Section 3.1.

Table A.11: Relationship Between Aspirations, Expectations, and Other Psychological Outcomes

	(1)	(2)	(3)	(4)
	Asp index	Asp index	Exp index	Exp index
	OLS	LASSO	OLS	LASSO
Discount factor	-0.125		-0.121	
	(0.12)		(0.12)	
No present bias	0.052		0.030	
	(0.34)		(0.57)	
Risk-taking	0.012		0.022	
	(0.63)		(0.36)	
Self-beliefs index	0.062***		0.017	
	(0.01)		(0.45)	
Beliefs about returns	0.036		0.007	
	(0.20)		(0.79)	
Mental health Z-score	0.086***	0.013	0.141***	0.063
	(0.00)		(0.00)	
Observations	1716	1716	1717	1717
R^2	0.02		0.02	

This table shows the relationship between the aspirations and expectations indices and other psychological measures defined in Figure 1 in the placebo group at endline. Odd columns are the estimates from OLS regressions of the Aspirations Index (1) and Expectations Index (3) on the other psychological mechanisms. Even columns are the estimates from LASSO regressions of the Aspirations Index (2) and Expectations Index (4) on the other psychological mechanisms, with the lamba parameter chosen using cross validation. Heteroskedasticity-robust standard errors shown in parentheses for the OLS estimates. No standard errors shown for the LASSO estimates because post-model-selection inference is not valid. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Table A.12: Treatment Effects of the Psychological Intervention at Different Levels of Baseline Economic Resources

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Index Components					
	Economic Index	Labour Supplied (Days)	Inputs & Hired Labour	Education Expenditure	Revenue	Non-Land Assets	Consumption Expenditure
Panel A: Consumption as wealth proxy							
Psych	0.112***	27.1**	230**	268*	23.0	99.8**	143*
	(0.035)	(11.7)	(99)	(154)	(27.1)	(45.8)	(74)
	[.]	[0.064]	[0.064]	[0.064]	[0.109]	[0.064]	[0.064]
Interaction	0.000	-0.002	0.088*	0.062	-0.003	0.035	-0.027
* wealth	(0.000)	(0.005)	(0.047)	(0.069)	(0.011)	(0.021)	(0.031)
	[.]	[1.000]	[0.439]	[0.600]	[1.000]	[0.439]	[0.600]
Placebo mean	0.000	525	857	2,101	640	1,529	3,796
# clusters	413	413	413	413	412	413	412
# obs	7,243	7,240	7,243	7,243	6,273	7,242	7,224
Panel B: Non-land assets as wealth proxy							
Psych	0.110***	26.2**	232**	257*	22.7	96.1**	139*
	(0.035)	(11.6)	(102)	(156)	(26.8)	(45.6)	(75)
	[.]	[0.077]	[0.077]	[0.077]	[0.136]	[0.077]	[0.077]
Interaction	0.000	-0.006	0.109	-0.059	-0.024	0.031	0.033
* wealth	(0.000)	(0.008)	(0.080)	(0.112)	(0.020)	(0.056)	(0.058)
	[.]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Placebo mean	0.000	525	857	2,101	640	1,529	3,796
# clusters	413	413	413	413	412	413	412
# obs	7,243	7,240	7,243	7,243	6,273	7,242	7,224
Panel C: Assets including land as wealth proxy							
Psych	0.111***	26.8**	229**	261*	21.2	96.0**	139*
	(0.036)	(11.6)	(101)	(156)	(27.3)	(45.8)	(75)
	[.]	[0.076]	[0.076]	[0.079]	[0.130]	[0.076]	[0.076]
Interaction	-0.000	-0.001	0.006	0.001	-0.001	-0.001	0.006
* wealth	(0.000)	(0.001)	(0.006)	(0.010)	(0.003)	(0.005)	(0.005)
	[.]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Placebo mean	0.000	525	857	2,101	640	1,529	3,796
# clusters	413	413	413	413	412	413	412
# obs	7,243	7,240	7,243	7,243	6,273	7,242	7,224

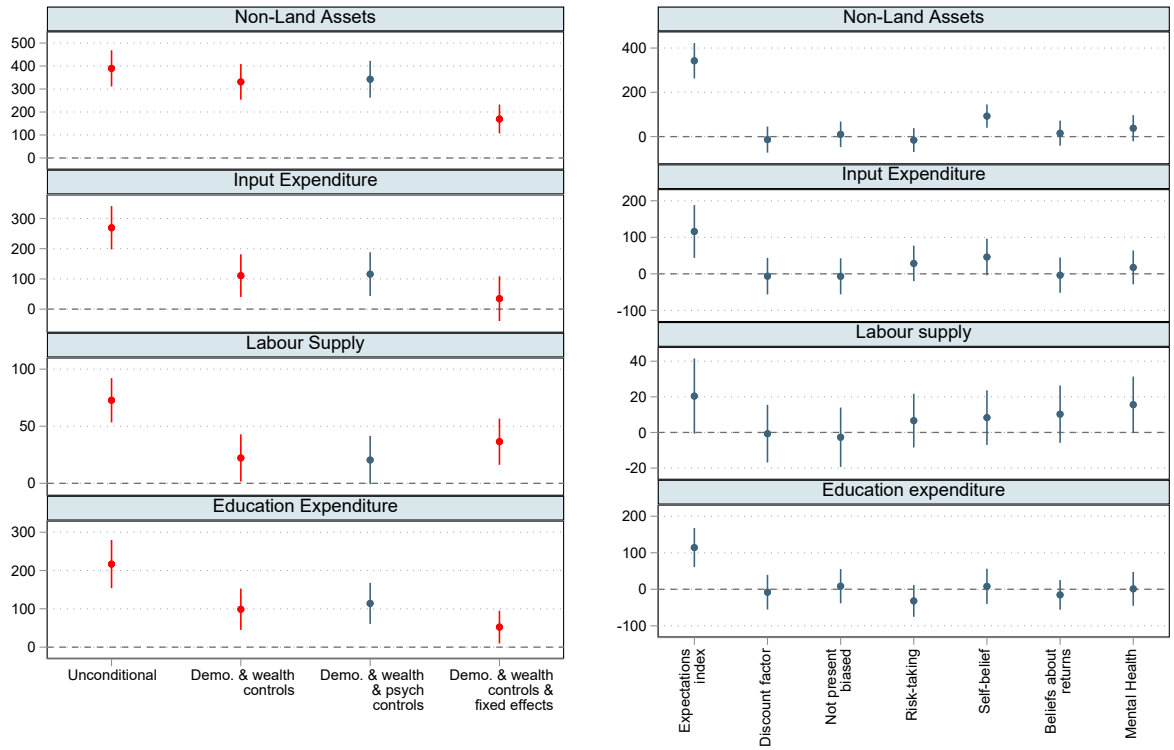
Notes: This table shows heterogeneous treatment effects of the psychological intervention on economic outcomes. Coefficients are from an OLS regression of each outcome on a vector of treatment assignments, a wealth proxy, the interaction of treatment assignments and the wealth proxy, the baseline outcome, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent and prespecified baseline covariates. Each panel shows results from using a separate wealth proxy. The wealth proxy in panel C is the value of all non-loan assets including the respondents' assessment of the value of their land and housing. All other proxies and all outcomes are defined in Table 1. We report heteroskedasticity-robust standard errors, clustered at village level, in parentheses. Sharpened q-values controlling for the false discovery rate across outcomes within each family are shown in brackets. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Table A.13: Treatment Effects on the Investment Share of Total Expenditure

	(1)	(2)	(3)
	Inv. share	Inv. share	Inv. share
Cash	0.0307***	0.0175***	0.0176**
	(0.0061)	(0.0053)	(0.0053)
Combined	0.0395***	0.0256***	0.0254***
	(0.0068)	(0.0054)	(0.0054)
Total expenditure (1000s)		0.0198***	0.0211***
		(0.0007)	(0.0010)
Cash X total expenditure (1000s)			-0.0020
			(0.0014)
Combined X total expenditure (1000s)			-0.0014
			(0.0015)
Placebo mean outcome	0.192	0.192	0.192
Placebo mean regressor		5.388	5.388
Placebo std dev. regressor		3.139	3.139
# clusters	412.000	412.000	412.000
# obs	7156.000	7116.000	7116.000

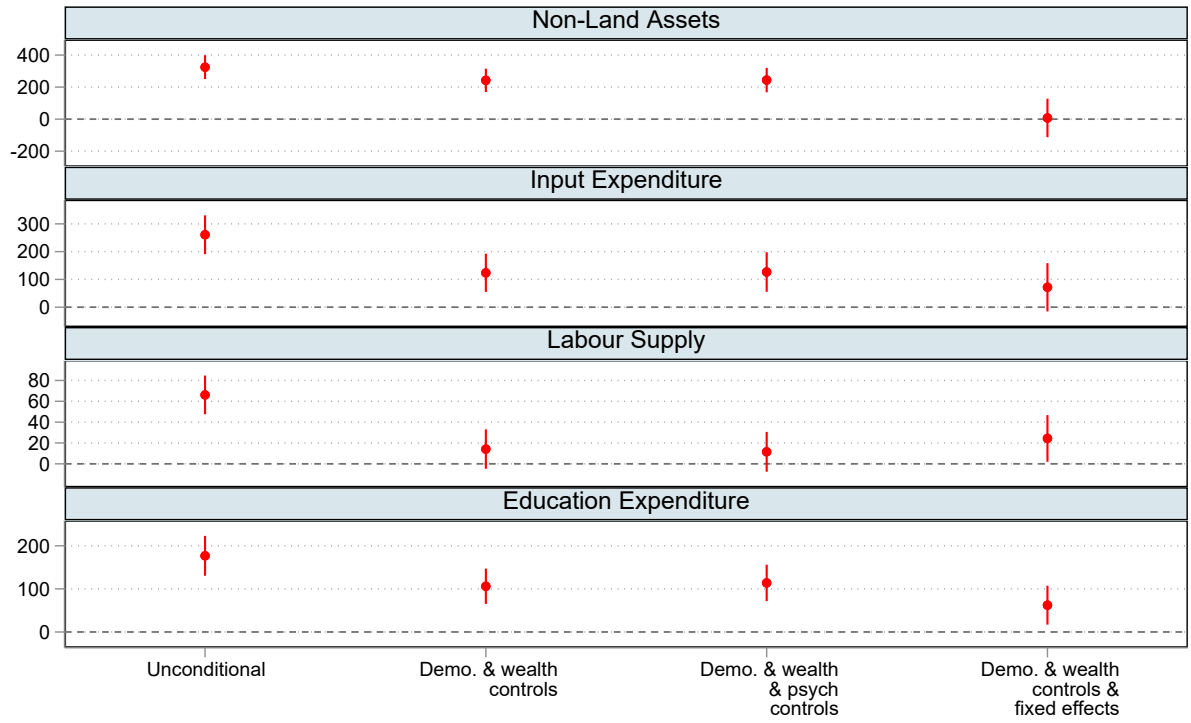
Notes: This table shows treatment effects of the cash and combined interventions on the share of total expenditure that is invested. Coefficients in column (1) are from an OLS regression on a vector of treatment assignments, sublocation fixed effects, endline month fixed effects, an indicator for the endline being answered by a proxy respondent and prespecified baseline covariates. Coefficients in column (2) are from a regression that adds total expenditure and coefficients in column (3) are from a regression that also adds interactions between treatment assignments and total expenditure. All expenditure and investment share measures have the top percentile trimmed to reduce sensitivity to outliers. Total expenditure is reported in 1000s of USD PPP and is the sum of expenditure on consumption, education, productive inputs, and hired labour; total investment excludes expenditure on consumption. We report heteroskedasticity-robust standard errors, clustered at village level, in parentheses. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Figure A.1: Asset- and Investment-Expectations Relationship Using Non-experimental Variation



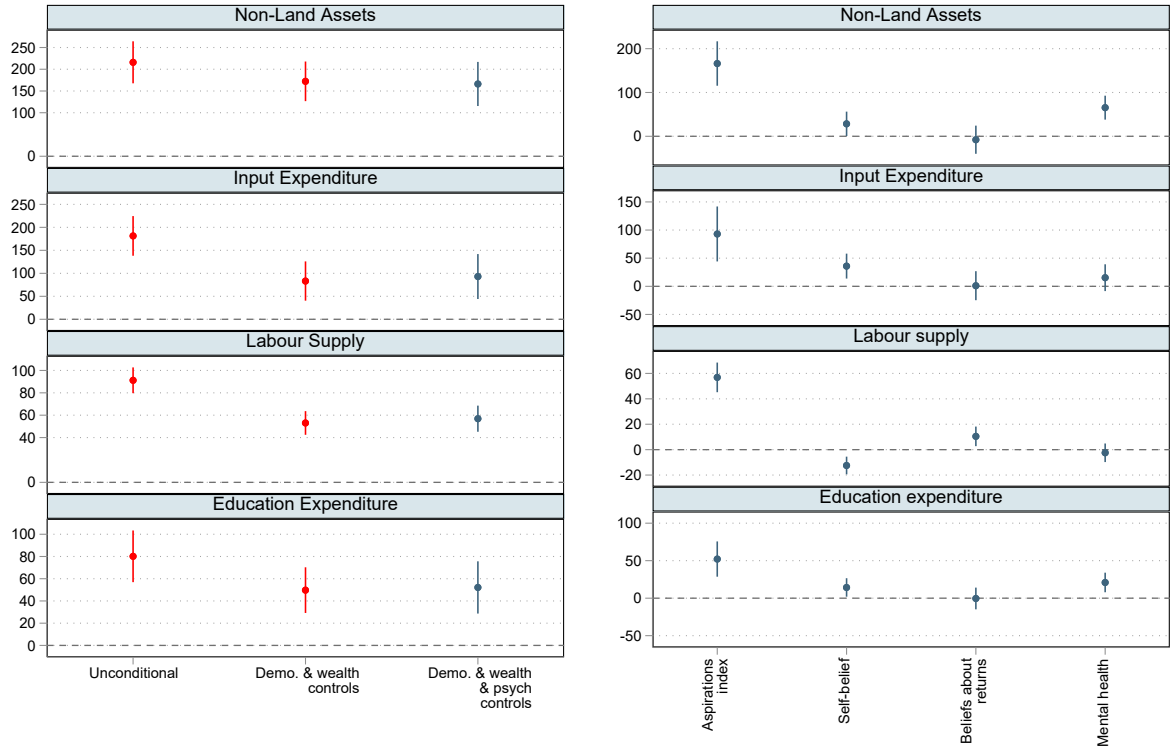
Notes: This figure shows coefficients and 95% confidence intervals from regressing different wealth and investment measures on psychological characteristics. The left hand panel shows coefficients and 95% confidence intervals from regressing each of non-land assets, input expenditure, labour supply and education expenditure, respectively, on the standardised expectations index and other variables. The first column in the left-side panel shows the coefficients on the expectations index from bivariate regressions. The second column shows the coefficients on the expectations index controlling for respondent age, education, marital status, household size, number of school-aged members, county fixed effects and (except for the top row) asset value and consumption. The third column shows the coefficients on the expectations index controlling for the same variables and the psychological characteristics shown in the right hand panel. The fourth column shows the coefficients on the expectations index controlling for the same variables as in the second column, plus respondent fixed effects. All asset and investment measures are defined in the footnote below Table 1. The expectations index and psychological measures are defined in Section 3.1. Here, each is standardised to allow for coefficient comparison. The right-hand panel of this figure shows coefficients and 95% confidence intervals on all psychological variables from the same regression that generates the third column in the left-hand panel. All regressions use the endline placebo group data with the top percentiles of expectations, investment, assets, and consumption trimmed. Sample size is 1376 to 1747 depending on the choice of controls and investment measure. The smaller sample sizes are for education expenditure, as this is set to missing for households with no school-aged children. The confidence intervals are estimated using heteroskedasticity-robust standard errors.

Figure A.2: Asset- and Investment-Aspirations Relationship Using Non-experimental Variation



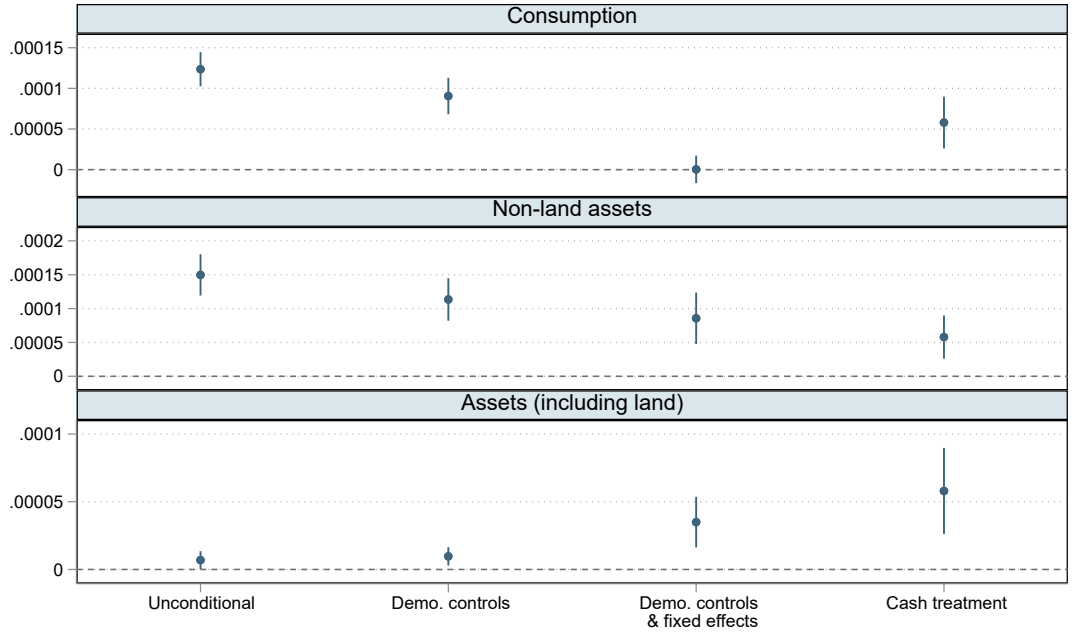
Notes: This figure shows coefficients and 95% confidence intervals from regressing different wealth and investment measures on psychological characteristics. It shows coefficients and 95% confidence intervals from regressing each of non-land assets, input expenditure, labour supply and education expenditure, respectively, on the standardised aspirations index and other variables. The first column shows the coefficients on the aspirations index from bivariate regressions. The second column shows the coefficients on the aspirations index controlling for respondent age, education, marital status, household size, number of school-aged members, county fixed effects and (except for the top row) asset value and consumption. The third column shows the coefficients on the aspirations index controlling for the same variables and the psychological characteristics shown in Figure 1. The fourth column shows the coefficients on the aspirations index controlling for the same variables as in the second column, plus respondent fixed effects. All asset and investment measures are defined in the footnote below Table 1. The aspirations index and psychological measures are defined in Section 3.1. All regressions use the endline placebo group data with the top percentiles of expectations, investment, assets, and consumption trimmed. Sample size is 1376 to 1747 depending on the choice of controls and investment measure. The smaller sample sizes are for education expenditure, as this is set to missing for households with no school-aged children. The confidence intervals are estimated using heteroskedasticity-robust standard errors.

Figure A.3: Asset- and Investment-Aspirations Relationship Using Non-experimental Variation at Baseline



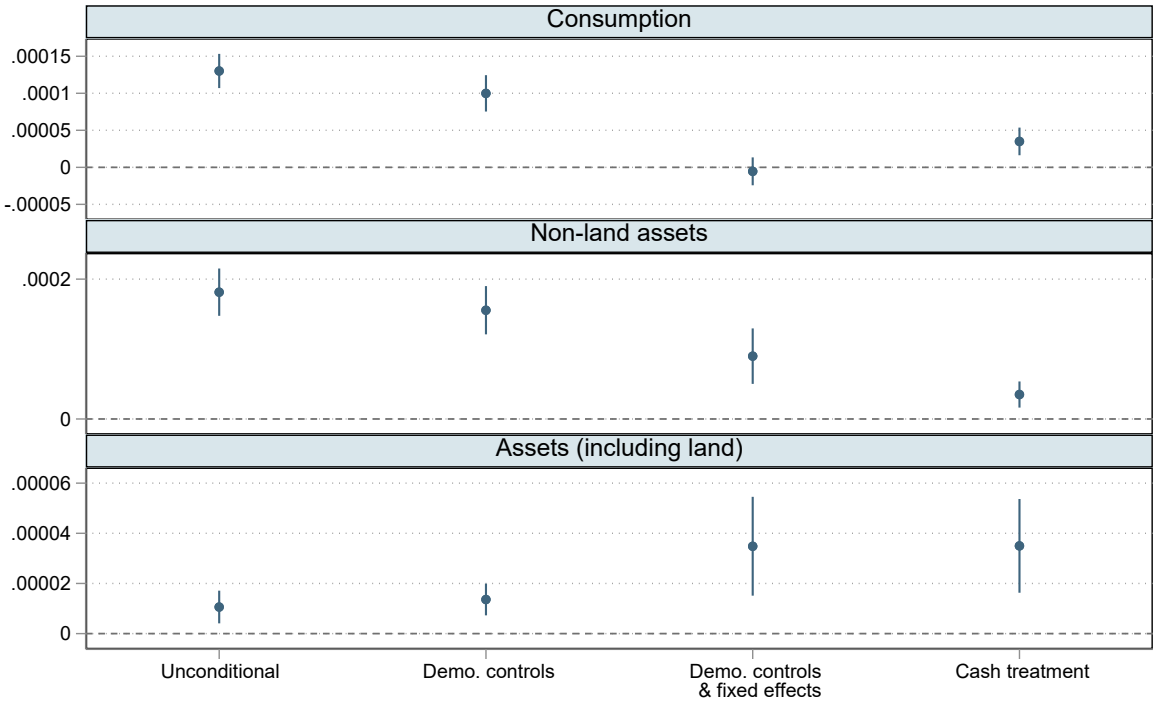
Notes: This figure shows coefficients and 95% confidence intervals from regressing different wealth and investment measures on psychological characteristics at baseline. The left hand panel shows coefficients and 95% confidence intervals from regressing each of non-land assets, input expenditure, labour supply and education expenditure, respectively, on the standardised aspirations index and other variables. The first column in the left-hand panel shows the coefficients on the aspirations index from bivariate regressions. The second column shows the coefficients on the aspirations index controlling for respondent age, education, marital status, household size, number of school-aged members, county fixed effects and (except for the top row) asset value and consumption. The third column shows the coefficients on the aspirations index controlling for the same variables and the psychological characteristics shown in the right hand panel. Compared to Figure A.2, no fixed effects specification is reported as the relationship reports on cross-sectional data at baseline and time and risk preferences are not reported because they were not measured at baseline. All regressions use the full baseline sample with the top percentiles of aspirations, investment, assets, and consumption trimmed. All asset and investment measures are defined in the footnote below Table 1. The psychological measures are defined in Section 3.1. Here, each is standardised to allow for coefficient comparison. Beliefs about returns is the standardised measure of beliefs about returns to fertiliser. Sample size is 5731 to 8175 depending on the choice of controls and investment measure. The smaller sample sizes are for education expenditure, as this is set to missing for households with no school-aged children. The confidence intervals are estimated using heteroskedasticity-robust standard errors.

Figure A.4: Aspirations-Wealth Relationship Using Non-experimental Variation



This figure shows estimates and 95% confidence intervals of the aspirations-wealth relationship estimated in different ways. The first, second, and third rows measure wealth using respectively the value of annual consumption, the value of non-land and non-housing assets, and the value of total assets. The wealth proxy measures are defined in the footnote below Table 1. The first column is estimated using a bivariate regression of the aspirations index on each wealth proxy. The second column is from a regression including respondent age, education, marital status, household size, number of school-aged members and county fixed effects. The third column is from a regression including the same controls and respondent fixed effects using the panel data. The fourth column is the treatment effect of the cash transfer on aspirations, divided by the 2,237 USD PPP value of the cash transfer so it has the same scale as the first three columns. This is included as a benchmark to help interpret the magnitudes of the non-experimental estimates. All regressions use the endline placebo group data with the top percentiles of aspirations and wealth trimmed. Sample size is 1716 to 1743 depending on the choice of controls and wealth proxy. The confidence intervals are estimated using heteroskedasticity-robust standard errors.

Figure A.5: Expectations-Wealth Relationship Using Non-experimental Variation



This figure shows estimates and 95% confidence intervals of the expectations-wealth relationship estimated in different ways. The first, second, and third rows measure wealth using respectively the value of annual consumption, the value of non-land and non-housing assets, and the value of total assets. The wealth proxy measures are defined in the footnote below Table 1. The first column is estimated using a bivariate regression of the expectations index on each wealth proxy. The second column is from a regression including respondent age, education, marital status, household size, number of school-aged members and county fixed effects. The third column is from a regression including the same controls and respondent fixed effects using the panel data. The fourth column is the treatment effect of the cash transfer on expectations, divided by the 2,237 USD PPP value of the cash transfer so it has the same scale as the first three columns. This is included as a benchmark to help interpret the magnitudes of the non-experimental estimates. All regressions use the endline placebo group data with the top percentiles of expectations and wealth trimmed. Sample size is 1717 to 1743 depending on the choice of controls and wealth proxy. The confidence intervals are estimated using heteroskedasticity-robust standard errors.

A.1 Heterogeneous Treatment Effects

This appendix describes our tests for treatment effect heterogeneity and reports the results. We focus on two outcomes: the economic index and the aspirations index. We estimate treatment effect heterogeneity across the economic index and eight prespecified baseline characteristics: widowed or unmarried (versus married), aspirations index, expectations index, self-beliefs index, age, education, non-land asset value, and household size.

We estimate heterogeneous treatment effects in two ways. First, we estimate heterogeneous treatment effects using treatment-interacted regressions:

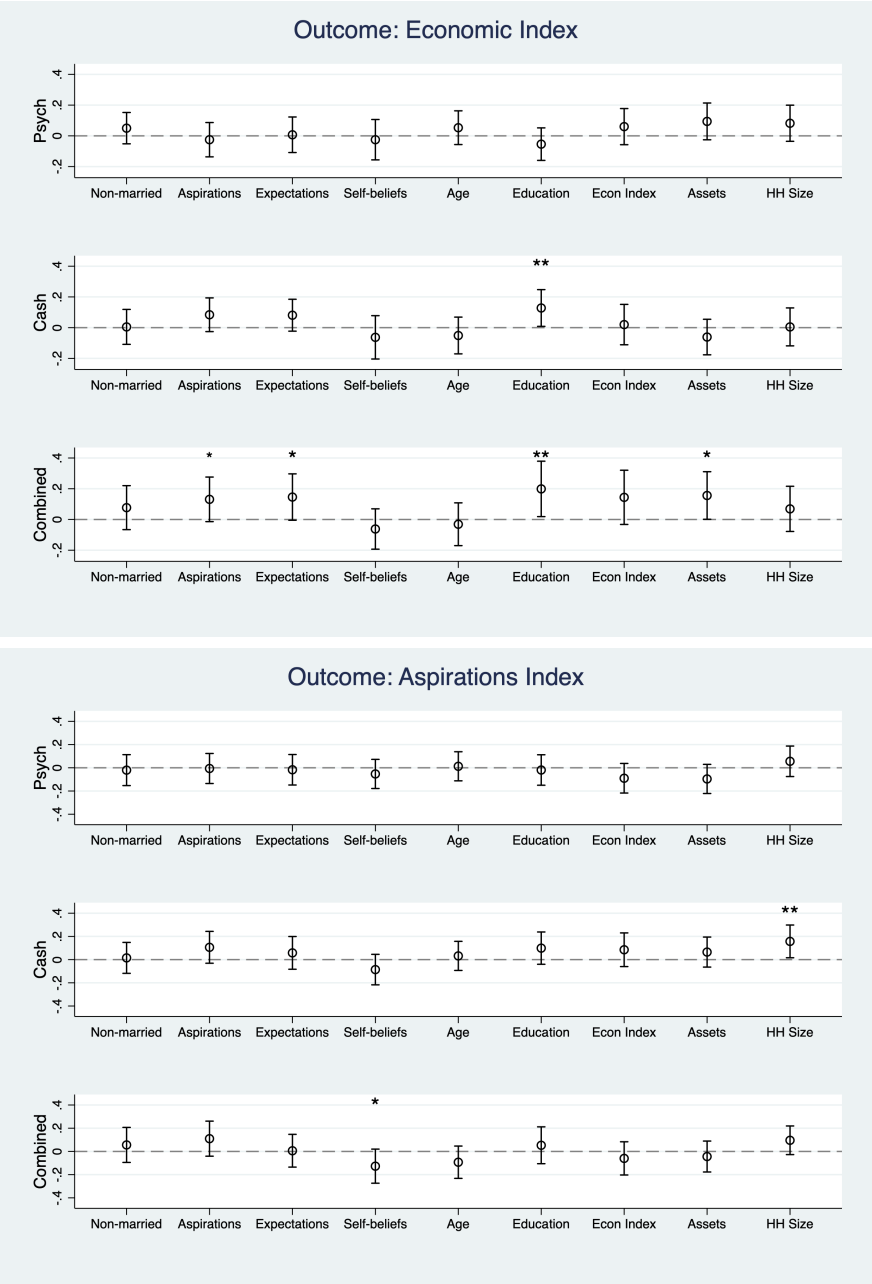
$$\begin{aligned} Y_{iv} = &\text{Cash}_v \cdot \beta_C + \text{Psych}_v \cdot \beta_P + \text{Combined}_v \cdot \beta_{CP} + \mathbf{X_{iv}} \cdot \mathbf{\Gamma} \\ &+ \text{Cash}_v \cdot W_{iv} \cdot \alpha_C + \text{Psych}_v \cdot W_{iv} \cdot \alpha_P + \text{Combined}_v \cdot W_{iv} \cdot \alpha_{CP} + \epsilon_{iv} \end{aligned} \tag{5}$$

where W_{iv} is the relevant baseline characteristic. We convert all continuous measures into indicators equal to one for values above the sample median. We report the estimated interaction effects (α_C,α_P,α_B) in Figure A.6. These are seldom large and the fraction of statistically significant estimates is no larger than would arise by chance. These results provide no support for heterogeneity in treatment effects once we adjust for multiple hypothesis testing.

Second, we estimate heterogeneous effects using a causal forest (Wager and Athey, 2018). We first residualise the outcomes with respect to covariates using a standard regression forest. (Causal forests also require values of the treatment propensity score; we know these exactly from the randomization and hence do not need to estimate them.) We then run a causal forest on these residuals to generate the causal forest estimator. Estimation proceeds as follows. We randomly partition the dataset into training and testing samples in a 80/20 split. In the training dataset, we construct a set of 1001 trees, repeatedly split the data into cells based on values of the nine baseline characteristics and estimate treatment effects within these cells. Each tree is "honestly" fit: the data is used to estimate only the within-leaf treatment effect or to decide on split placement, but not both. We then generate the forest estimate by averaging these prediction rules across trees. The use of separate training and testing datasets prevents overfitting. We then apply the causal forest ensemble decision rule to the testing data to estimate heterogeneous treatment effects across the cells and run tests on these effects.

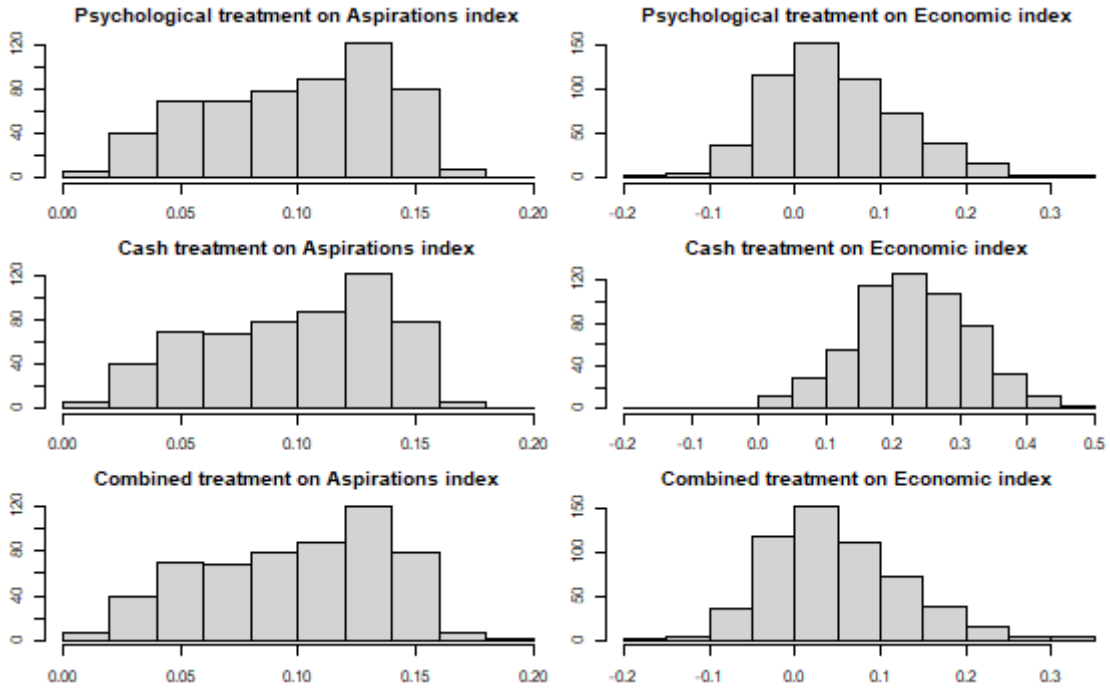
Figure A.7 shows the distribution of heterogeneous treatment effects over the cells for the psychological, cash, and combined interventions. There is limited evidence of heterogeneous treatment effects. The effects of all three treatments on the aspirations index range over the cells from 0 to 0.18 standard deviations, with almost all estimates falling between 0.05 and 0.15. The psychological intervention and combined effects on the economic index are almost all between

Figure A.6: Heterogeneous Treatment Effects Estimated Using Interacted Regression Models



Notes: This figure shows heterogeneous treatment effects estimated using the treatment*covariate interacted regression model in equation 5. The outcome in the first panel is the summary economic index, defined in the footnote to Table 2. The outcome in the second panel is the aspirations index, defined in the footnote to Figure 1. All variables interacted with treatment are measured at baseline and continuous measures are converted into indicators equal to one for values above the sample median. Confidence intervals are estimated using heteroskedasticity-robust standard errors, clustered by village. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Figure A.7: Distribution of Heterogeneous Treatment Effects Estimated Using Random Causal Forests



Notes: This figure presents histograms of heterogeneous treatment effects estimated using the random causal forest. The outcome in the left-hand column is the summary economic index, defined in the footnote to Table 2. The outcome in the right-hand column is the aspirations index, defined in the footnote to Figure 1. The x-axis represents the treatment effects.

-0.05 and 0.2 standard deviations, while the cash effects are slightly higher.

We use these estimates to conduct an omnibus test for the presence of heterogeneous treatment effects, following Tibshirani et al. (2022). The test's null hypothesis is that the treatment effects in all cells estimated by the forest are equal. Table A.14 shows that we fail to reject this null hypothesis for all interventions (row 2 of each panel) and fail to reject a goodness-of-fit test for the forest itself (row 1).

Table A.14: Tests for Treatment Effect Heterogeneity Using Random Causal Forests

	Psych	Cash	Combined
Panel A: Economic index			
p-value for null hypothesis: forest fits data	0.98	0.98	0.98
p-value for null hypothesis: forest detects no HTEs	0.82	0.15	0.81
Panel B: Aspirations index			
p-value for null hypothesis: forest fits data	0.96	0.96	0.85
p-value for null hypothesis: forest detects no HTEs	0.64	0.99	0.64

Notes: This table shows results of tests for heterogeneous treatment effects based on random causal forests. All cells display p-values for testing the listed hypotheses. Each column corresponds to a different treatment arm. Each row corresponds to a different null hypothesis. The first row in each panel reports a goodness of fit test; we fail to reject the null hypothesis that the forest estimated using the training data fits testing data. In the second row, we fail to reject the null hypothesis that the effects over all forest cells are jointly equal. The forest uses the same nine baseline dimensions of heterogeneity shown in Figure A.6. The outcome in the first panel is the summary economic index, defined in the footnote to Table 2. The outcome in the second panel is the aspirations index, defined in the footnote to Figure 1. Inference uses a village-clustered bootstrap following Semenova and Chernozhukov (2021).

A.2 Participant Goals

The psychological intervention encourages participants to set goals for their future. Our aspirations measures provide quantitative proxies for these goals in relatively general domains: assets, income, and education. In this section, we describe the relationship between our survey measures of aspirations and the actual goals that participants describe during the facilitated exercises, to provide more information about the proxy relationship. In the psychological and combined groups, one fieldworker observed the group while respondents presented their goals for five years' time and recorded three goals for each respondent. We code the open-ended responses using STATA's regular expression function, regexm, to filter for a given string in the open-ended goals variable. We define a dummy = 1 if a participant has a goal which contains the string. We do not quantify goals in this analysis, as most participants expressed their goals in qualitative terms. We record goals only for participants in the psychological and combined interventions. The placebo exercises did not include any goal-setting discussion that we could have used to record goals.

Few respondents set goals for the same activities shown in the videos, reinforcing the evidence from Section 7 that the treatment effects of the psychological intervention are not driven by mimicry. To show this, we examine the proportion of respondents that report a goal including each string in Table A.15. Only 3% of respondents set goals in tailoring, Judy's core activity in the video. Only 1% set a savings goal and 7% a goal related to farming infrastructure, although savings and building a greenhouse were described in the videos. Many (48-52%) set goals related to chickens,

Table A.15: Proportion of Respondents Reporting Goals by Domain

	(1)	(2)
Keywords	% Reporting goal psychological group	% Reporting goal combined group
Productive assets	0.96	0.93
Chicken, chick, hen, poultry	0.52	0.48
Goat	0.03	0.03
Sheep	0.01	0.02
Cow, cattle	0.33	0.32
Farming infrastructure[(a)]	0.07	0.07
Income		
Small business goals[(b)]	0.36	0.37
Income-generating farming goals	0.77	0.80
Tailoring goal	0.03	0.04
Savings and credit	0.01	0.01
Savings, save, ROSCA	0.01	0.01
Credit, loan	0	0
Housing[(c)]	0.87	0.85
Education	0.51	0.50
Primary and secondary schooling[(d)]	0.50	0.49
Tertiary education[(e)]	0.06	0.06

Notes: Bolded row titles represent indicators for any subset of groupings of the reference categories beneath them. Proportions do not necessarily sum within reference categories, as often respondents will nominate more than one of each of the members of a category. Proportions do not sum to 100 as participants set more than one goal. [(a)] is made up of the keywords "hoe, irrigation, fence, mill, tractor, greenhouse, tank, plot". [(b)] is made up of "retail, trade, business, tailor, tailoring, shop, sell, fish, rent, boda, motorbike cereal, omena, mala, commercial". [(c)] is made up of "house, home, roof, walls, floor, tin, simba, kitchen, room". [(d)] is made up of "school, educat, educate, education, schooling and any keywords in [(e)]. [(e)] is made up of "university, college, diploma, tertiary".

but this is unlikely to reflect mimicry as 80% of households in the psychological group at baseline have a chicken, as do 78% in the combined group. There is little evidence of differential goal-setting between the psychological and combined groups, with respondents setting very similar goals.

We find positive, statistically significant relationships between a dummy for having a goal in a domain and an individuals' level of aspirations in that domain. We report bivariate regressions of having a goal in a particular domain at the time of intervention on aspirations in that domain at endline in Table A.16. The regression coefficients are larger than the treatment effects of the psychological intervention. This suggests our measures of aspirations are informative proxies for having a specific goal in that domain, consistent with psychological literature suggesting aspirations and long-term goals are similar conceptually (Locke and Latham, 2002).

Table A.16: Relationship between Domain-Specific Aspirations and Goals

	(1)	(2)	(3)
	Aspirations measures		
Dummy if has goal	Education (years)	Income (USD PPP)	Assets (USD PPP)
Education goal	0.25		
P-value	0.00		
Income goal		739	
P-value		0.02	
Asset goal			1376
P-value			0.42

Notes: This table shows the relationship between endline aspirations measures and the presence of at least one elicited goal in the same domain. The goal definitions are specified in Table A.15. We group some of the goals into being related to education, income or assets to correspond to dimensions of the aspirations measure. The reported values are coefficients and p-values from OLS regressions with no covariates and heteroskedasticity-robust standard errors. The regressions use data from the psychological and combined intervention groups.

B Additional Information about Data Collection, Sampling & Eligibility, Treatment Assignment & Receipt

B.1 Data Collection

In this appendix we describe each stage of data collection and show that endline attrition does not differ between treatment groups. We collected data in four stages: (1) a **village census** of elders in 415 villages; (2) a **household census** of 32,921 households in these villages; (3) a **baseline survey** of 8,309 sampled households; and (4) an **endline survey**, aimed at the same households.

We use the **village and household censuses** to determine study eligibility, create the baseline sample, and measure some pre-treatment characteristics (e.g. household size). Any adult household member could answer the household census questions.

We use the **baseline household surveys** as our main sources of pre-treatment household-level economic measures and respondent-level psychological measures. The baseline respondent in each household is the "primary adult woman." This is typically the female spouse in a couple-headed household or the widowed head of household. Baseline surveys ran from April 2016 to March 2017.

The interventions ran from November 2016 to July 2017. GiveDirectly did not operate in any geographic area until we had completed baseline surveys for all villages in that area. So respondents' baseline responses should not reflect anticipation of treatment.

We use the endline surveys as our main sources of post-treatment household-level economic measures and respondent-level psychological measures. Endline surveys ran from May 2018 to

February 2019. We used a **three-stage endline tracking protocol**. First, we surveyed village elders about where each household and primary female respondent were living, before endline started. If the household or respondent had migrated, we got information on their new location. Second, we ran the endline survey, including with respondents who migrated to another village in the study area. Third, we surveyed 94 respondents who moved outside the study area, in-person where possible (e.g., in the nearest city, Kisumu) or otherwise by phone.

We surveyed 84.7% of baseline respondents at endline. In another 2.4% of households we could not survey the baseline respondent but did survey a proxy household representative. This gives a total response rate of 87.1%. Of the endline sample, 1.2% are phone surveys with migrants and 0.4% are households that split between baseline and endline. For split households, we surveyed the original respondent and a representative from the other part of the split household. For economic measures, we use proxy responses if the respondent cannot be found and we average both responses for split households. For psychological measures, we only use responses from the baseline respondent. Our main findings are robust to alternative ways of handling proxy responses and split households, partly because these cases are rare.

Table B.1 shows that **attrition does not differ by treatment assignment** (column 1). Attrition is slightly lower for larger households and respondents with higher values of the self-beliefs index (column 3). However, attrition does not differ by treatment $\times$ baseline household characteristics (column 5), suggesting that the composition of the sample remains similar between groups.

We also conducted **baseline and endline price surveys** of 55 markets in the study area. We use the price survey data to construct some economic measures. At baseline, we created a list of all markets in and adjacent to the study area and identified 55 markets that each opened multiple days per week. We collected price information for the most commonly purchased goods and services at each market at baseline and endline including food products, livestock prices, livestock input prices, non-food items, services, and wages for different types of labour. Baseline price surveys ran from August to November 2016. Endline price surveys ran from May to September 2018.

B.2 Eligibility Criteria and Sampling

Eligibility criteria: Households in the census were classified as study-eligible if they (1) included an adult female member and were either (2a) widow-headed or (2b) classified as poor by a multidimensional poverty index. The first criterion was chosen because the psychological intervention was aimed at adult women and the second criterion was chosen to align with GiveDirectly's eligibility criteria for cash transfers.

GiveDirectly used a multidimensional poverty index that defined households as eligible if

they met ANY ONE of these criteria: (1) household's per capita housing space was less than 62,000cm^2; (2) household had no telephone and had a mud floor; (3) household head was a widow and has a mud floor; (4) household had an orphan child or (5) household was homeless. GiveDirectly chose these criteria based on prior research in Western Kenya showing that these quick-to-collect criteria were strongly associated with more complex poverty measures based on consumption or multidimensional poverty indices. Approximately 40% of households in the study villages met the eligibility criteria. GiveDirectly's eligibility criteria have changed since the study period and they now give transfers to most or all households in a village.

Our research team excluded households if they were polygamous (due to difficulties associated with household definition); if the household head was a child (due to difficulties with informed consent from minors); if the household was homeless (due to difficulty finding them); or if the household did not contain an adult female (since the chosen psychological intervention is aimed at adult females).

Sampling process for baseline: In each village, we randomly drew two samples of eligible households from the census: the "target" and "reserve" households. Field officers were instructed to find each target household for the baseline survey. If a target household refused to participate or could not be located (e.g. due to migration), the field officers included one household on the reserve list as a replacement. We define the study sample as all households that completed the baseline survey. The idea of reserve households was used only for the baseline survey. In latter rounds of data collection, households that refused to participate or could not be located are treated as attriters. We sampled up to 18 target and 6 reserve households per village in Homa Bay county. In Siaya county, where the villages are substantially larger, we generally did not impose an upper limit on the number of target households per village. Sampling probabilities vary by village because we don't sample exactly the same share of treatment-eligible households in each village. None of the descriptive or causal results in the paper change substantively when we reweight the data to account for variation in sampling probabilities.

B.3 Treatment Assignment and Treatment Receipt

Randomisation protocol: We conducted the randomisation using data from the village and household censuses. Randomisation took place after the censuses but before the baseline survey. GiveDirectly did not operate in any area until we had completed baseline surveys for all villages in that area. So respondents' baseline responses are measured before they learn their treatment status and potentially before they ever hear about the cash transfer program. We stratify randomisation on location (an administrative division in Kenya containing roughly 10-50 villages); village size

(a dummy for if village size exceeded the sample median); and a measure of village assets. For the village asset measure, we calculated the first principal component of household-level indicators equal to one if the household owns a solar panel, television, fridge, iron, radio, watch or clock, telephone, bicycle, motorbike, truck, or car. We then calculated village-level averages of this index and created an indicator variable equal to one if the village asset index exceeded the sample median. Within each block, we randomly assigned villages to the four experimental arms. Some block sizes were not multiples of four. We took the residual villages from these blocks, combined them with other villages in the same location and same asset category, and randomised within these units.

Balance tests: Table B.2 reports the results of balance tests in the baselined sample. All tests compare the mean values of baseline village- and household-level covariates between the placebo group and each treatment arm (columns 3, 5, and 7) or jointly across all four arms (column 8). The test results are consistent with random assignment: we reject equality across all four arms at the 10% level for only 1/25 tests and at the 5% level for 0/25 tests.

Treatment receipt: Table B.3 reports the relationship between treatment assignment and treatment receipt. Columns 1-4 show statistics for all baselined households and columns 5-8 show statistics for only endlined households. We focus discussion on the endlined households because this is the sample used to estimate treatment effects. The first row shows the number of baselined and endlined households assigned to each treatment group.

Panel A shows the share of households in each treatment group (in columns) that received each possible combination of treatments. We do not show separate rows for receiving the psychological and placebo interventions because, by design, households could only receive whichever one of these interventions was assigned to their village. Receipt of the psychological and placebo interventions is similar across all experimental arms. In the placebo and psychological groups, respectively 89 and 90% of endlined households receive the placebo/psychological intervention that they are offered (row 3, columns 5 and 6).

The cash and combined groups have four different treatment receipt measures. They can receive neither the cash transfer nor the placebo/psychological intervention (row 1 of Panel A); cash but not the placebo/psychological intervention (row 2); the placebo/psychological intervention but not cash (row 3) or both interventions (row 4). Receipt of the cash transfers is similar between the cash and combined experimental arms: respectively 79 and 81% of endlined households receive the cash transfer (sum of rows 2 and 4, columns 7 and 8). Receipt of the workshops is also similar between the cash and combined experimental arms: respectively 90 and 91% of endlined households receive the workshop (sum of rows 3 and 4, columns 7 and 8).

Table B.1: Relationship between Attrition, Treatment Assignment, and Baseline Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
Psych	0.020	(0.014)			0.055	(0.056)
Cash	0.009	(0.013)			0.027	(0.050)
Combined	-0.001	(0.012)			0.064	(0.049)
Household Size			-0.011***	(0.002)	-0.007*	(0.004)
Age			-0.001*	(0.000)	-0.008	(0.005)
Self-beliefs Index			-0.013***	(0.004)	-0.006	(0.005)
Non-land Assets			0.004	(0.003)	-0.002	(0.005)
Consumption			0.001	(0.001)	-0.000	(0.001)
At Least Primary Education			0.005	(0.008)	-0.000	(0.001)
At Least Secondary Education			0.019	(0.017)	-0.000	(0.001)
Married			0.018*	(0.010)	-0.001	(0.001)
Household Size * Psych					-0.009	(0.008)
Household Size * Cash					0.000	(0.010)
Household Size * Combined					-0.016	(0.011)
Age * Psych					0.002	(0.010)
Age * Cash					0.000	(0.006)
Age * Combined					0.009	(0.008)
Self-beliefs Index * Psych					0.001	(0.008)
Self-beliefs Index * Cash					0.004	(0.008)
Self-beliefs Index * Combined					0.001	(0.003)
Non-land Assets * Psych					-0.001	(0.004)
Non-land Assets * Cash					0.002	(0.004)
Non-land Assets * Combined					-0.001	(0.004)
Consumption * Psych					0.003	(0.015)
Consumption * Cash					-0.013	(0.021)
Consumption * Combined					0.012	(0.024)
At Least Primary Education * Psych					0.004	(0.023)
At Least Primary Education * Cash					0.044	(0.033)
At Least Primary Education * Combined					-0.025	(0.046)
At Least Secondary Education * Psych					-0.022	(0.048)
At Least Secondary Education * Cash					-0.042	(0.046)
At Least Secondary Education * Combined					0.008	(0.017)
Married * Psych					0.039	(0.028)
Married * Cash					0.001	(0.026)
Married * Combined					-0.002	(0.024)
P: All arms = 0	0.362					
P: cash = psych	0.411					
P: cash = combined	0.433					
P: psych = combined	0.114					
Placebo mean	0.122		0.122		0.122	
# obs	8,309		8,309		8,309	

Notes: This table shows the relationship between attrition, treatment assignment, and prespecified baseline covariates. All columns show regressions of a household-level indicator for not being surveyed at endline on treatment arm indicators (cols 1-2); baseline covariates (cols 3-4); and treatment arm indicators, baseline covariates, and their interactions (cols 5-6). All regressions include sublocation fixed effects. The consumption and asset aggregates are measured in constant 2018 USD PPP (000s). The self-beliefs index consists of growth mindset, self-efficacy and internal locus of control scales. If a baseline covariate is missing, we replace the missing values with the sample mean and include a missing data indicator. Heteroskedasticity-robust standard errors, clustered at village level, are reported in parentheses. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively.

Table B.2: Baseline Summary Statistics and Tests of Balance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Placebo	Pysch		Cash		Combined		F-test	
	Mean	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	p-value	# obs
Panel A: Village-level Characteristics from Census									
Has Primary School	0.490	-0.010	0.890	0.057	0.427	0.028	0.702	0.802	415
Has Market	0.288	-0.075	0.232	-0.040	0.528	-0.028	0.669	0.685	415
Has Clinic	0.163	-0.075	0.129	-0.048	0.323	0.002	0.964	0.294	415
Number of Households	96.3	-2.73	0.679	-5.35	0.349	0.062	0.993	0.756	415
Mean Household Asset Score	0.030	-0.032	0.515	-0.017	0.728	-0.022	0.649	0.929	415
Floor Material is Mud or Organic†	0.666	0.023	0.158	-0.003	0.872	-0.009	0.569	0.187	415
Roof Material is Grass, Leaves or Other†	0.054	-0.006	0.297	0.000	0.976	0.003	0.631	0.420	415
Walls Material is Unburnt Bricks or Mud†	0.846	0.005	0.672	0.003	0.821	0.001	0.954	0.973	415
Drinking Water is Piped/Well†	0.385	0.047	0.089	-0.012	0.705	0.020	0.526	0.206	415
Lighting is Electricity†	0.284	0.004	0.749	0.008	0.492	-0.003	0.787	0.796	415
Panel B: Eligible Respondent Characteristics from Census									
Married	0.584	0.039	0.027	0.018	0.310	0.001	0.961	0.093	8,309
Age	40.8	-0.603	0.283	-0.866	0.153	-0.808	0.197	0.485	8,302
At Least Primary Education	0.423	0.018	0.253	0.026	0.109	0.024	0.123	0.345	8,274
Household Owns a Mobile Phone	0.741	-0.008	0.628	0.017	0.265	0.011	0.490	0.441	7,743
Panel C: Eligible Household Characteristics from the Baseline									
Household Size	5.31	0.075	0.388	0.031	0.720	0.021	0.818	0.845	8,309
Dependency Ratio	1.35	-0.004	0.923	0.023	0.537	0.026	0.500	0.802	8,308
Number of Household Members Under 16	2.85	0.020	0.751	0.052	0.438	0.032	0.649	0.890	8,309
Revenue Aggregate	1,834	-71.0	0.381	32.0	0.693	38.6	0.674	0.572	8,309
Consumption Aggregate	4,331	-78.1	0.470	-83.2	0.420	-98.7	0.359	0.778	8,295
Investment Aggregate	699	-125	0.067	-48.6	0.507	-40.2	0.661	0.228	8,309
Non-land Asset Aggregate	1,230	-7.57	0.887	-0.049	0.999	-3.80	0.946	0.999	8,309
Total Household Labour Supply (Days)	431	-3.50	0.808	-6.72	0.646	-0.218	0.988	0.965	8,283
Education Expenditure	439	7.00	0.779	-10.2	0.717	-13.0	0.589	0.853	6,958
Index of Self-beliefs	-0.000	0.013	0.789	0.051	0.246	-0.031	0.520	0.395	8,270
Index of Aspirations for Future Outcomes	0.000	0.025	0.534	0.012	0.742	0.063	0.104	0.357	8,283

Notes: The table reports balance tests for characteristics measured in the village census, household census, and baseline surveys. Panel A reports regressions at the village level. Panel B reports characteristics of eligible respondents who are the primary women in eligible households. Panel C reports household-level characteristics. All balance tests are implemented by regressing the characteristic on a vector of treatment assignments and sublocation fixed effects. The regressions use one observation per village for the village-level characteristics and one observation per household for the household- and respondent-level characteristics. Inference is performed using heteroskedasticity-robust standard errors, clustered by village for regressions with household- or respondent-level characteristics. Column (1) reports the placebo mean for each characteristic. Columns (2)-(7) report treatment arm-specific coefficients and p-values. Column (8) reports the p-value from test of joint equality of means for all four treatment arms. Column (9) reports the number of observations. The average number of households in each village that completes the census is 75. The household asset score is constructed using principal component analysis on indicators for household ownership of a telephone, bicycle, solar panel, TV, fridge, radio, watch/clock, motorbike, truck and iron box (charcoal or electric). Economic variables in Panel C are measured in constant 2018 USD PPP annually. The dependency ratio is the number of household members under 16 divided by the number of members 16 or above. Outcomes with a † sign denote village-level proportions constructed from household-level data.

The only comparison where there are differences in treatment receipt rate between arms is that there is a higher receipt rate for the psychological/placebo intervention than the cash transfer. This is relevant only for comparing outcomes between the psychological and cash groups. But this has a limited effect on our comparison of the two interventions' benefit-cost ratios because lower cash take-up lowers both the numerator – because the average benefits are intention-to-treat effects – and the denominator – because the average costs are per person offered each intervention – relative to full take-up.

Panel B shows that differences in treatment receipt rates between relevant pairs of experimental arms are not statistically significant. There is only one measure of treatment receipt for the psychological and placebo groups: the share who completed the intervention they were offered. This differs by only 0.6 percentage points with $p=0.66$ (row 3, columns 5/6). The differences in these receipt measures range from 0.2 to 1.7 percentage points and no difference is significant at conventional levels with $0.54<p<0.84$ (rows 1-4, columns 7/8).

Treatment receipt is not clearly related to households' baseline characteristics (analysis not shown). The only clear relationship is that larger households are more likely to receive cash transfers, potentially because they are more likely to have an adult at home and available to meet GiveDirectly's census team.

We use data from GiveDirectly's census data and payment records to construct detailed reasons why some households did not receive cash transfers. GiveDirectly conducted a separate census in most villages to determine eligibility after we conducted our baseline survey, as in Egger et al. (2022). This had two advantages for the research. First, it allowed us to complete the baseline surveys in each location before GiveDirectly entered the location and households in cash-assigned villages learned about GiveDirectly. Second, it avoided households treating the research team as the source of cash transfers. However, it meant that a few households classified as transfer-eligible using our census data and hence included in our baseline sample might not be classified as transfer-eligible by GiveDirectly.

Panel C shows the reasons why some households assigned to receive cash transfers did not receive them. The portion of the sample who do not receive transfers is similar in the cash and combined groups, suggesting few differences in sample composition between these groups. Amongst our endlined households, GiveDirectly classified 3% as ineligible based on their census data (row 2). Another 2% of households were not found by GiveDirectly's census team and hence could not receive transfers, typically because no adult was home for any of the multiple visits the census team made to each house (row 3). 1% of households moved out of their village between our baseline survey and GiveDirectly's census (row 4).

Table B.3: Treatment Receipt Statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline Sample by Treatment Group				Endline Sample by Treatment Group			
	Placebo	Psych	Cash	Combined	Placebo	Psych	Cash	Combined
# households	2,012	2,057	2,085	2,157	1,767	1,766	1,814	1,896
Panel A: Cash and Psychological/Placebo Intervention Receipt Rates (%)								
None	16.20	16.29	14.92	11.54	10.64	9.97	7.22	5.80
Cash only	0.05	0.00	2.59	2.83	0.06	0.00	2.65	2.80
Psych/plcb only	83.75	83.62	14.15	14.79	89.30	89.92	13.45	13.03
Cash and psych/plcb	0.00	0.10	68.35	70.84	0.00	0.11	76.68	78.38
Panel B: Test for Equal Treatment Receipt Rates between Groups (p-values)								
Cash only	-			0.730	-			0.844
Psych/plcb only	0.946			0.764	0.664			0.832
Cash and psych/plcb	-			0.485	-			0.538
All	-			0.231	-			0.722
Panel C: Cash Transfer Receipt Rates and Reasons for Non-receipt (%)								
Received cash	-	-	70.94	73.67	-	-	79.33	81.17
Ineligible	-	-	3.07	2.83	-	-	2.87	2.85
Not found/home	-	-	2.30	2.32	-	-	1.71	1.58
Left village	-	-	0.96	1.21	-	-	0.66	0.84
Refused	-	-	18.13	15.25	-	-	10.64	9.12
Unknown/Other	-	-	4.60	4.73	-	-	4.80	4.43

Notes: This table shows receipt of treatment for households in each treatment arm. Columns (1), (2), (3) and (4) are statistics for the baseline sample while columns (5), (6), (7) and (8) are for the endline sample. Panel A shows the percentage of each treatment group that received each of the four possible treatment combinations: no treatment (row 1), only cash (row 2), only the psychological/placebo intervention (row 3), or both cash and the psychological/placebo intervention (row 4). 'Psych/plcb' in row 3 and 4 refer to watching the assigned video and completing the assigned exercises as part of the psychological intervention or placebo. Panel B shows p-values from tests across pairs of treatment groups for equal rates of treatment receipt: only cash (row 1), only the psychological/placebo intervention (row 2), and both cash and the psychological/placebo intervention (row 3). The fourth row shows the p-value from a joint test of equality of all four treatment receipt rates. Each p-value is centred between the columns corresponding to the two groups on which the test is run. Panel C shows the percentage of the cash-assigned treatment groups that received cash (row 1 of the panel) or did not receive cash for each reason (rows 2-6). See the in-text discussion for detailed explanations of each reason for not receiving cash. Inference is performed using heteroskedasticity-robust standard errors, clustered at village level.

Another 10% of households in our endline sample refused the cash transfers at the time GiveDirectly offered them (row 5). Substantial non-receipt is not unusual in studies of asset transfers. For example, the multi-country study by Banerjee et al. (2015) reports that a bundled antipoverty programme had a refusal rate of 48% in India and close to zero in other countries. Refusal of unconditional or nearly unconditional cash benefits is common in high-income economies, as documented in a review by Ribar (2014). It is difficult to compare this to other studies of unconditional cash transfers because few studies report transfer receipt rates. For example, two

published studies of GiveDirectly cash transfers do not report receipt rates (Egger et al., 2022; Haushofer and Shapiro, 2016), nor do Baird et al. (2011).

C Additional Information about Interventions

C.1 Cash Transfers

Marking Village Boundaries: Some study villages are contiguous. Our data collection team mapped village boundaries with village elders. We shared satellite maps showing the village boundaries with GiveDirectly to create a mutually agreed definition of villages. We also shared a list of households in our baseline sample assigned to the placebo and psychological arms with GiveDirectly, to minimise the risk that any would accidentally receive cash transfers. Only 3 out of 4,069 households assigned to the placebo and psychological arms received cash transfers.

Information about GiveDirectly operations: In cash treatment villages, GiveDirectly held a meeting open to all households in the village to explain their programme and inform village residents that GiveDirectly would be working in their village. The eligibility criteria were not disclosed, although households were told that poorer households would be targeted. GiveDirectly emphasised that the transfers were from an independent non-governmental organisation, and not the result of any government programme. These meetings occurred after the study baseline was finished in all nearby villages.

Information about transfers: Transfers were sent using the mobile money system M-Pesa. Households were asked to register for M-Pesa if they had not done so, which could be done at a network of agents in most small stores. If an eligible household did not have a mobile phone to access M-Pesa, GiveDirectly provided one and subtracted its cost from the value of the final transfer. Each household nominated one member whose M-Pesa account would receive the transfer. Households were told if there was any intra-household conflict about the transfer, they could be disqualified.

C.2 Psychological and Placebo Interventions

C.2.1 Summaries of Video Content

The psychological and placebo videos are posted at https://mbrg.bsg.ox.ac.uk/aspirations-and-goal-setting-video-intervention. The psychologically active videos tell stories about the lives of Judy and Josefine. Judy and and her partner Oyoo are smallholder farmers with few assets, and are expecting a child. They discuss struggling to get by and what they want for the future. They decide that within five years, they want to put an iron roof on their house, for their children to complete school and for Judy to start a business. They set intermediate goals to save 100

shillings each week, grow more vegetables to sell at market, purchase a plot within a year. They put money in a small container to save. Judy learns to sew, overcoming some obstacles to do so, and starts a successful business making clothes. They succeed in buying an iron roof. She adjusts her business plan to deal with competition from cheap imported clothing and her business prospers. Eventually, they send their child to university.

Josefine is a teacher and farmer. She tells her life story of how she came to be successful. She begins by remembering that she used to beg for money as a child and work as a casual day labourer. Her teacher describes how she dropped out of school several times. Another woman describes teaching Josefine to weave baskets to sell at market. Her husband describes how she saved money from this to go to high school. She explains how she learned conservation farming to improve the productivity of her plot. Her husband describes how she eventually started her chicken-rearing business, despite five failed attempts, including when she overfed her chickens and they did not lay eggs. She outlines plans to build a greenhouse. She also describes her struggles to succeed at teacher training college, when she was much older than others and struggled to learn, but persisted and achieved good grades. She encourages viewers to continue learning throughout their lives.

C.2.2 Exercise Script for Psychological Intervention

Field officer (FO): Show "Aspirations female" video. After the participants have watched both the videos: "You have now seen two stories. The first story was about Judy, who one day decided that instead of accepting her circumstances as being fixed and difficult to alter, she was determined to improve her economic situation for herself and her family. The second story was about Josefine, who decided that it was never too late to learn new skills and that success depends on one's willingness to learn. She continued seeking out opportunities instead of waiting for them to come to her. She found learning and improving to be fulfilling and exciting. You saw that both Judy and Josefine were very determined to make plans for their lives and to never give up on these plans. They believed that even people who are from poor backgrounds can make real changes in their own lives using resources that are widely available. Remember, as stated at the start of the video, these stories are parables. The women in the videos are not real people. Their stories are fictional and have been written to teach us lessons. Their stories may not reflect the lives of everyone in Western Kenya."

Best Possible Selves: "Now let's do an exercise like Judy did. Please take a moment to imagine your "best possible self" in 5 years, in 2021/2. This means that you imagine in 5 years, in the year 2021/2, after everything has gone as well as it possibly could. You have worked hard and succeeded at accomplishing everything you wanted. Think of this as the realization of your dreams. You have reached your full potential. Imagine the best possible way that your life could look in 5 years, in the year 2021/2. You can think of what your house will look like, what your

farm will look like, what your business will look like, which new abilities you will have developed, or what your family will look like, what your children or grandchildren will be doing. Try to draw what you just pictured. You can draw anything. It does not have to be a beautiful drawing; the only thing that is important is that you know what it shows. The drawing can be as simple as you want, but please try to think of as many details as possible. And please let me know if you need help with the drawing. This is also an exercise that you do on your own, so while you draw, I will ask you to only focus on your own drawing."

Field officer (FO): While the participants are drawing ask them to explain what they are drawing to keep them engaged. Encourage them and ask for details. E.g. if they draw a house, ask them "How large? What is it made of? Where is it located?" If they draw livestock, ask "which kind, how many." If they draw a plot ask "how large," what crop do you want to grow". Encourage them to draw more than one goal (preferably no less than three). "Imagine that what you drew is actually a picture of your life in 5 years, in the year 2021/2. Please, explain to the group what your future looks like in as much detail as possible."

Field officer (FO): For each question below, try to let the participants speak without interruption. "I will now ask you to share your drawings and the futures they represent with me."

Field officer (FO): Take notes of specific goals. If they give very short answers ask clarifying questions as you did while they drew.

Personal Goal-setting: "To achieve your ideal future 5 years from now, in 2021/2, it sounds like you should be working towards achieving the following goals: *Field officer (FO): List goals from drawing and make sure to present them in a very clear and specific way, e.g. "get a tin roof," "send children to university", "get a plot of land". Exclude things they say that are more "fluffy" such as "know who I am" and "be happy". List only goals that are concrete and specific: they have to summarizable in approximately 5 words.*

"You mentioned several goals: [repeat goals from previous round]. Now, I'd like you to think of which of these goals are most important to you and which goals you are most certain that you can achieve by 2021/2. Which is the goal that is the most important for you to achieve that you also think is possible? [Wait for answer.] Which is the second? [Wait for answer.] Which is the third? [Continue until all goals are ranked.]

Field officer (FO): Write down ranked goals for participant and repeat the list to the participant. For the following questions: ask each question separately and wait for an answer before moving on to the next question.

"How long do you think, realistically, that it will take you to achieve your goal of [insert highest ranked goal] if you actively work towards achieving it already from today?"

Implementation Intentions: "In order to achieve your goal of [insert highest ranked goal] you need to have a plan. For this plan to work it has to be as specific and detailed as possible. What can you do or who can you go talk to today or over the next week in order to work towards achieving your goal of [insert highest ranked goal]?"

Anticipation of Obstacles: "In order for your plan to work, you need to have a strategy for what to do if something unexpected happens that challenges your plan. Such strategy could be, for example, to agree with your family members that you will sit down together and discuss what to do if something unexpected happens that challenges your plan. Another strategy could be to identify specific people you will consult with if something happens that challenges your goal. For example, in the first movie Judy saved money in two different places, such that she had one saving that she would only use for achieving her goal and a backup saving she can use if something unexpected happens. What works as a good strategy can be very different from person to person, so what can you do to avoid giving up on your goal if you encounter obstacles?"

Field officer (FO): The following questions should be answered only for the chosen goal for each participant.

Mental Contrasting: "How would your future be better than your life now if you achieved your goal of [insert highest ranked goal]?" "How would it make you feel to achieve this goal?"

Great work [name of person]! You now know how to set goals and have determined plans and strategies that will help you achieve your most important goals. Remind me again, what are you going to do today or over the next week in order to achieve your most important goal?

Field officer (FO): Pause for response and provide support if needed.

"Keep the drawing as a reminder of your goals. This can be your life in 2021/2 if you stay focused on achieving your goals and never give up. You can do whatever you want with this drawing but we hope you will put it to good use. For example, you can put it on the wall as a reminder of your goals or to show visitors which goals you are working toward. You can also show it to your spouse or parents to engage them in working towards your goals. Or you could store it in a box or keep it somewhere safe, so that one day you can show it to your children when you have achieved your goal to inspire them to set goals like you have done. As a thank you for your participation today, I'd like to give a wall calendar with the pictures of Judy and Josefine. On this calendar, there is a blank space. First, I'd like to ask you whether you would like me to write your most important goal in this blank space?"

Field officer (FO): Ask if they would like to write their most important goal written in this blank space. If they want to, help them write it in the space (e.g. purchase a tin roof, learn to sew).

Remember how [Judy and Josefine] overcame obstacles to achieve their goals. Thank you

for your participation today and I hope you have found this time we have spent to be useful! I have also brought a bunch of stickers and if you want, you can choose a sticker or multiple stickers that remind you of our conversation, and I would encourage you to place it on your calendar."

C.2.3 Administration of Psychological Intervention

The psychological and placebo intervention administration varied through time for logistical reasons. In villages treated early, individuals watched the video and completed the exercises with a single facilitator. In villages treated later, people watched the video in pairs on a tablet. Groups of 2-4 people, randomly assigned to a group, completed the exercise together, with two facilitators. Psychological and placebo intervention villages within a sublocation all receive treatment either in groups or individually, so treatment implementation is balanced within sublocation. We control for sublocation fixed effects, so psychological interventions are always administered in the same format as the placebo interventions to which they are compared. As we outline on page 16, group size and composition have no meaningful effect on treatment effects or interpretation of our results.[29]

D Details on Costing

This appendix provides additional information about our cost calculations and benefit-cost comparisons. For comparison purposes, we convert all costs to USD PPP in the analysis month-year (October 2018) to account for inflation and the exchange rates in the same way as the benefit estimates.

Table D.1 shows the components of the two aggregated cost categories shown in Table 3, the average variable cost of delivering each intervention and average total cost. Panel 1 gives the components of average variable cost (Costs II in Table 3). For the cash transfers, the intervention delivery cost is the value of the cash transfer itself. The transfers were delivered by mobile money, so field visits were not required for delivery. The number differs in the cash and combined arms because rates of take-up differed slightly. For the psychological intervention, the intervention delivery costs cover the field operation to deliver the videos and exercises incurred by the implementing partner IPA: field officer training and wages; purchases of tablets, power banks, and projectors (all fully depreciated); and purchases of calendars and stickers for participant exercises. This may be a conservative estimate, as IPA surveyors are better paid than lay counsellors (e.g., community health workers) who deliver other psychological interventions.

Panels 2 and 3 give the other components of average total cost (Costs I in Table 3). Panel 2 gives programme costs. For cash, the programme costs cover GiveDirectly's field operation to recruit participants, conduct a census, determine eligibility and register eligible households. We

[29]The group assignment protocol is described in our pre-analysis plan for data collected straight after the psychological/placebo intervention, posted at https://www.socialscienceregistry.org/trials/996.

also include non-field staff and office costs: a portion of administrative costs to run GiveDirectly in-country and globally, allocated based on this study's proportion of the overall GiveDirectly cash transfer budget. For the psychological intervention, we assume the same costs (i.e. we assume GiveDirectly were running the same census, registration and operation, but delivering the psychological intervention instead of cash). However, GiveDirectly's operational model is unlikely to be the most cost-effective way to deliver the psychological intervention. For example, it might not be cost-effective to target this intervention at poor households, as the costs of a census and then revisiting poor households for registration might be more than the costs of visiting a village once and including all households in the psychological intervention.

Table D.1: Components of Cost Estimates

	Psych	Cash	Combined
Panel 1: Average variable cost of intervention delivery			
Cash transfer costs[(a)]	0.00	1,930.76	2,005.11
Field personnel and travel	42.29	0.00	42.29
Intervention materials	5.31	0.00	5.31
Intervention hardware	6.71	0.00	6.71
Panel 2: Average variable cost of programming			
Personnel	40.39	40.39	40.39
Field costs and travel	39.84	39.84	39.84
Mobile money and banking	1.41	1.41	1.41
Office	2.73	2.73	2.73
Indirect costs	134.06	134.06	134.06
Panel 3: Fixed costs of intervention development			
Field trip	5.78	0.00	5.78
Production and editing	74.39	0.00	74.39
I: Average total cost = (1) + (2) + (3)	352.90	2,149.18	2,358.02
II: Average variable cost (intervention delivery only)(b) = (1)	1,930.76	54.31	2,059.42

Notes: All figures are in USD PPP from October 2018. See in-text notes for explanations of each type of cost.

Panel 3 gives the costs of intervention development. We set the intervention development cost of the cash transfer to zero because we cannot reliably estimate the cost of creating GiveDirectly. For the psychological intervention, the intervention development cost covers all payments made to Khanga Rue Media, the company that filmed and edited the videos, divided by the number of participants in the psychological and combined arm. This does not include the researchers' time on formative and preparatory work. We exclude the cost of developing the placebo interventions from this calculation.

E Conceptual Framework for Alternative Psychological Mechanisms

This appendix describes how psychological mechanisms other than aspirations and expectations could enter the conceptual framework from Section 7.3. We begin with a simplified framework that omits the reference-dependent utility term. In this framework, utility over current consumption and leisure is $u(c_t,l_t)$. Revenue from current capital investment and labour supply is $y_{t+1}=f(k_t,e_t;\theta)$ with productivity parameter θ. We continue to assume that utility is additively separable through time with discount factor δ and that utility and revenue are both increasing and concave functions of their arguments.

This formulation yields standard first-order conditions. People set marginal utility of current consumption equal to the discounted marginal utility of future consumption arising from current capital investment and the marginal utility of current leisure equal to the discounted marginal utility of future consumption arising from current labour:

$$u_{c_t}=\delta\cdot(1+f_{k_t}-\gamma)\cdot u_{c_{t+1}} \tag{6}$$

$$u_{l_t}=\delta\cdot f_{e_t}\cdot u_{c_{t+1}} \tag{7}$$

Before introducing the psychological mechanisms, we note that this framework can explain the treatment effects of the cash transfer. The cash transfer raises current assets a_t, leading to higher consumption c_t and higher capital investment k_t. The positive cash transfer effect on labour supply e_t is possible in this framework in two ways. First, it may occur if the cash transfer influences labour supply through a behavioural channel, as we discuss in Section 7.4. Second, it may occur if capital and labour are technical complements in the production function ($f_{kl}>0$), in which case the cash transfer will increase the right-hand side of condition (7) and people will reallocate time from leisure to work to increase the left-hand side and maintain equality. This second mechanism also requires that own and hired labour are imperfect substitutes, as in LaFave and Thomas (2016). The positive effects on future revenue y_{t+1} and asset value a_{t+1} follow naturally from higher capital investment and labour supply into the revenue function $f(k_t,e_t;\theta)$.

A higher **belief about the return to investment** or higher **beliefs about one's ability to achieve desired outcomes** enter the conceptual framework as a higher perceived value of the productivity parameter θ, as in Ghosal et al. (2020) and McKelway (2021). If θ enhances the marginal productivity of labour supply and capital investment ($f_{k\theta},f_{e\theta}>0$), then the psychological intervention will raise the right-hand sides of conditions (6) and (7). To maintain the first-order conditions, people will reallocate from consumption to capital investment and from leisure to labour supply. This will increase future revenue and asset values through the revenue function. This will steepen the time path of consumption but have an ambiguous effect on the level of consumption

at endline, for the same reasons as in the reference-dependence framework from Section 7.3.

A change in **time preferences** raises the discount factor δ, increasing the utility weight on future relative to current consumption. In the more general quasi-hyperbolic discounting framework from Laibson (1997), agents discount utility τ periods in the future by $\beta \cdot \delta^{\tau}$ rather than δ^{τ}. The psychological intervention could change preferences in this framework by raising either β or δ. Predictions are the same as for beliefs about return to investment and ability.

Risk aversion could be introduced by making future revenue stochastic conditional on current investment, generalising the first-order condition (6) to $u_{c_t} = \delta E_t\left[(1+f_{k_t}-\gamma)u_{c_{t+1}}\right]$. The psychological intervention enters this framework by reducing risk aversion. This reduces the value of certain consumption and leisure today relative to uncertain future outcomes and *can* lead to higher current labour supply and capital investment and hence higher future revenue and asset value. But the exact prediction will depend on the extent to which increased absolute or relative risk aversion affects the covariance across states of the world between f_{k_t} and $u_{c_{t+1}}$.

Improved **mental health**, which we measure through a depression scale, may enter the framework in multiple ways due to the diversity of depression's symptoms, making predictions difficult. Depression may lead to pessimistic beliefs (MacLeod and Salaminiou, 2001), reduced cognitive functioning (Prado et al., 2018; Keller et al., 2019) or altered time and risk preferences (Yaakov et al., 2019; Cobb-Clark et al., 2021). These could enter the framework respectively as low expected returns to investment, lower productivity or altered preferences, with predictions as outlined above.

F Departures from Preanalysis Plan

Our estimation and inference methods and our outcome measurement are mostly prespecified. Our preanalysis plan (PAP) is posted at https://www.socialscienceregistry.org/trials/996. This appendix describes some departures from the preanalysis plan, none of which substantially changes the estimated treatment effects.

F.1 Measurement

1. We report variables using annual scaling instead of the prespecified monthly scaling to improve comparability to other papers in the cash transfers literature. This multiplies all treatment effects and standard errors by 12, leaving hypothesis test results unchanged.
2. The PAP specifies that the asset aggregate would include net financial liabilities. We omit this component as it is noisily measured.
3. The PAP specifies that household labour earnings (in the revenue aggregate) would include remittances received from household members who are migrants; that the "expenditure on

inputs and hired labour" aggregate would include migration expenditures and that the labour supply aggregate would include migrant labour supply. We leave remittances, expenditure and labour of migrants living away from home out of these aggregates as households were uncertain of these values. Migration is relatively rare, with only 4% of labour supply coming from migrants.

4. We use a recall period of the last 30 days for revenue from non-farm enterprises, not the prespecified recall period of the last 30 days that the enterprise was in operation. This increases comparability with all other revenue and investment measures.
5. We rename "investment into economic activity" from the PAP to "expenditure on inputs and hired labour" in order to convey more clearly what is included in this outcome.
6. To measure investment into human capital, the PAP specified an index consisting of education expenditure (total expenditure on education divided by the number of children in the household) and participation (total number of school days attended over the last five days school was in session divided by the number of children in the household). There are two deviations from the PAP. First, we use education expenditure as the main measure of human capital investment rather than combining it in an index with participation. Second, we do not divide expenditure by the number of children. These two changes give us a household-level pecuniary measure that is more comparable to the other economic aggregates and can be used in cost-benefit analysis. There are no treatment effects of any treatment on education participation.
7. In the labour supply aggregate, we make two deviations. First, the PAP specified that total labour supply be an index of labour supplied to household agricultural activities, household livestock activities, household non-farm enterprises, and activities outside of the household (casual work, salaried employment, and labour supply by migrants). Instead of the index, we construct a labour supply aggregate measured in days to allow more meaningful interpretation of the magnitude. We scale labour supplied to each activity to annual figures, as outlined in Appendix G. Second, we omit labour supply by household members aged 15 or younger to make the labour supply measure consistent with the prespecified revenue aggregate, which only includes labour earnings from adults.
8. We omit education expenditure from the consumption aggregate because we report education expenditure separately.
9. We prespecified that we would estimate treatment effects on measures of cognition but dropped these from the endline survey and cannot estimate treatment effects on them.

F.2 Econometric Specification

We replace prespecified randomisation block fixed effects with sublocation fixed effects. Sublocations are administrative units in Kenya each containing roughly 10 contiguous villages. The sublocation fixed effects absorb slightly more outcome variation and marginally improve baseline balance. The two sets of fixed effects are highly correlated, with Cramér's V = 0.73.

We prespecified defining the treatment indicators as assigned to receive a cash transfer, assigned to receive a psychological intervention, and assigned to receive both interventions. In the paper, we instead use assigned to receive *only* a cash transfer, assigned to receive *only* a psychological intervention, and assigned to receive both interventions. The respective estimating equations yield identical information from the experiment.

F.3 Summary Indices and Adjustments for Multiple Hypothesis Testing

We depart from the PAP in two ways. First, we adjust for multiple testing by reporting sharpened q-values for the effect of each treatment and for tests of equality of each pair of treatment effects. This approach controls the false discovery rate for tests across different outcomes (Benjamini et al., 2006). We implement this adjustment across all the main economic aggregates and across all the main psychological mechanisms, as well as across the components of each aggregate or index. This is more conservative than the preanalysis plan, which only specified adjusting for multiple testing across components of each economic aggregate. Second, we construct a non-prespecified index, following Anderson (2008), to summarise all the treatment effects on economic aggregates.

F.4 Data Collected Directly After Intervention

We intended to collect data on aspirations, expectations, self-efficacy and growth mindset immediately after the psychological/placebo interventions. We do not include this data in the final paper as there were errors in data collection. We did not collect data on measures of asset or income expectations or asset aspirations and are missing measures of education aspirations and expectations for half the sample. These measures are also missing for respondents who miss the psychological or placebo intervention. This means that there is perfect compliance in the sample for whom we collect data after intervention, but partial compliance in the endlined sample. This complicates interpretation of magnitudes of the intention-to-treat effects that we report. For readers interested in dynamics despite these caveats, we note that all three interventions have positive, significant effects on the self-beliefs index of similar magnitude, in the sample of women who attend the psychological or placebo intervention. Effects are 0.17 standard deviations (SD) for cash, 0.24 SD for the psychological intervention and 0.35 SD for the combined intervention, with no significant differences between groups.

G Details on Measurement

G.1 Economic Measures

This subsection describes how we construct each of the six main economic aggregates from measures in our survey data. All aggregates except education spending are annualised. Any components of aggregates not measured with a 12 month recall are scaled to be annual measures. We convert all component measures into annual scales before combining them into aggregates. Recall periods vary across component measures according to standard measurement practice (Grosh and Glewwe, 2000).

G.1.1 Labour Supply

Total household labour supply is an aggregate of labour supply in the following activities for all household members over 15:

1. **Crop agriculture in last long rains:** Total days worked in the most recent long rains season, multiplied by 1.5 to approximate the annual agricultural labour supply over the short and long rains seasons. The 1.5 scaling factor is based on the baseline survey, where we measured both short and long rains labour supply.
2. **Livestock:** Total days worked in the last 4 weeks for all activities related to livestock raised or owned by the household.
3. **Non-farm enterprises:** Total days worked in the last 4 weeks in household enterprise(s).
4. **Labour outside the household:** In the last 4 weeks, total days worked in casual work or in any salaried employment for someone outside the household.

G.1.2 Expenditure on Inputs and Hired Labour

Expenditure on inputs and hired labour is an aggregate of the following:

1. **Agricultural input expenditures:** An aggregate of total expenditure in the last two seasons on fertiliser, seeds and seedlings, insecticide, storage, farm implements, irrigation and pumping equipment, farm machinery and fuel, and fees or interest for farm-related financial services; total rent paid for plots rented by the household in the last 12 months; and expenditure on hired labour for agriculture. The last is the total number of days of labour hired in for agriculture in the most recent long rains season multiplied by average wage rates for the village for the activity from the village price surveys, multiplied by 1.5 to approximate the annual agricultural labour demand over the short and long rains seasons.
2. **Livestock input expenditures:** An aggregate of total expenditure on purchasing livestock in the last 12 months; on all livestock inputs used in the last 12 months (feed, veterinary services, medicines and vaccines, equipment, transportation and construction of enclosures);

and on hired labour for livestock production. The last is the total number of days of labour hired in for livestock in the last 4 weeks multiplied by wages from the village price surveys. For animals, we only consider production costs in the year of measurement following FAO (1996).

3. **Non-agricultural expenditures:** An aggregate of total expenditure by any non-farming, non-livestock enterprises owned or operated by household members in the last 30 days, plus contributions by the household to any community group economic activities, excluding ROSCAs. Non-farm enterprise expenditures excludes durable goods/assets but includes maintenance (of land, buildings, or equipment), stocks and inventory, wages, fuel and rent costs.

G.1.3 Education expenditure

Total expenditure on education is expenditure on school and activity fees, other school related supplies and uniform cost in the January to December 2017 school year and since the start of the 2018 school year for all children aged 6-20. The recall period does not perfectly coincide with the period from treatment to endline because some households were treated during the 2017 school year. However, treatment timing is balanced across arms, so this issue will not bias the treatment effect estimates.

G.1.4 Revenue

Total revenue is an aggregate of the following::

1. **Revenue from agriculture:** An aggregate of revenue from household crop production and rental income from renting out land across economic activities. Revenue from household crop production is the total value of production for each crop the household grew in each of the last two rainy seasons, including both production sold and production kept and consumed in-kind. We measure units produced by the household. To value the units produced, the calculation uses a hierarchy of valuation approaches, relying on the next best approach whenever the required data for the preferred approach is not available for a crop.[30] Where costs have been incurred but output has not been harvested, we include the households estimate of work-in-progress production in output (Grosh and Glewwe, 2000). Rental income is rent received from renting out land or any buildings on it in the last 12 months.
2. **Revenue from livestock rearing and produce:** An aggregate of revenue from livestock

[30]The hierarchy of approaches to value production is: (a) the respondent's direct assessment of crop value; (b) the unit price from sales by the same household of the same crop-unit; (c) the crop-unit price obtained from the market price survey; (d) the sub-location median of other households' direct assessments of the value of the same crop-unit; (e) the sub-location median of other households' direct assessments of the value of the same crop, converted using a universal unit conversion ratio; (f) the sub-location median of other households' sales prices of the same crop-unit; (g) the sub-location median of other households' sales prices of the same crop, converted using the universal unit conversion ratio.

sales and livestock production. Livestock sales is the total value of animals the household sold in the last 12 months.[31] Revenue from livestock produce is the total value of livestock produce, such as meat, milk or hides, produced by the household in the last 30 days, including produce that is kept and consumed. We use the reported quantity produced and the price per unit produced to estimate the value of production.[32]

3. **Revenue from non-agricultural activities:** An aggregate of total sales, including value of in-kind income, from any non-farming, non-livestock enterprises owned or operated by household members in the last 30 days; total earnings received by any household member from any community group business activity, such as renting out of group assets, in the last 12 months; and earnings in cash or in kind from renting out any assets owned by the household in the last 12 months.
4. **Total household labour earnings:** For all household members in the last 4 weeks, an aggregate of total earnings in cash and in-kind from casual work and total earnings in cash and in-kind from salaried work for someone outside the household.

G.1.5 Non-land Assets

Total household non-land assets is an aggregate of:

1. **Durable assets**: Sum of the respondent's estimate of the value of household holdings of each asset of different types, if they were to sell them today in their current condition[33] and the respondent's estimate of the value of any shares household members have in non-financial assets owned by a group.
2. **Livestock:** Total value of the respondent's estimate of how much money they would get if they sold all mature and immature livestock of different types that they own today.
3. **Savings:** Total value of savings of all household members held at home, with friends and neighbours, with shopkeepers, with microcredit groups, in mobile money accounts, in bank accounts and in ROSCAs.
4. **Stocks of dried maize:** Value of stocks of dried maize currently owned by the household,

[31] We unfortunately omitted the value of livestock slaughtered for own consumption at endline.

[32] For livestock products sold by the household, we multiply the quantity sold by the most common price per unit of these sales, which the household reported. For livestock products that are not sold by the household, we multiply the reported quantity with the most common price the household estimated it would receive if it sold its products. If household price data is not available, we value production using local prices, using first prices from the price surveys and then median prices in the sub-location.

[33] This variable sums over the values of two classes of assets in our survey. First, we asked all households to report quantities and values of "core" assets that they own. Second, we asked households to report the value of a randomly selected asset from a list of 8 "randomised" assets, though households reported quantities for all of these assets. For the "randomised" assets without a value, we impute the value based on the reported quantity and the mean value of the asset for the sample (from observations for which we have both variables). The set of "randomised" assets was chosen based on which assets appear least frequently and had low variance in values in the baseline survey.

valued using unit prices of dried maize obtained from the market price survey.

G.1.6 Consumption

Total household consumption is an aggregate of:[34]

1. **Food consumption:** In the last 7 days, value of household consumption of 18 core food items, scaled to capture total consumption following the approach in Beegle et al. (2012), plus household food consumption outside the home. The PAP details the scaling approach.
2. **Non-food non-durable consumption:** In the last 30 days, value of household consumption of nine core non-food non-durable items, including household goods, fuel, hairdressing and transport fares, scaled to capture total non-food non-durable consumption following the approach in Beegle et al. (2012). The PAP details the scaling approach.
3. **Expenditure on durable goods:** In the last 12 months, value of household expenditure on durable items and the maintenance of durable items.
4. **Social expenditure:** Value of household expenditure on charitable donations, worship contributions, social and entertainment expenditures, weddings and bride price. Regular worship contributions are 30 days recall, wedding expenditures including bride price are since-intervention recall, and the other items all have 12-month recall periods.

G.2 Psychological Measures

Any indices are constructed as an inverse covariance-weighted average of several components, following Anderson (2008). In the next subsection, we describe how we adapt and validate psychological scales for measures of self-beliefs and depression.

G.2.1 Aspirations and Expectations for Future Outcomes

We measure aspirations and expectations in different domains of life – income, assets and children's education. We define assets as "the worth of your house, your furniture, consumer goods like a TV and fridge and any transport vehicles." We define income as "all sources of cash income for your household, including what you earn from all agricultural and non-agricultural activities, and money that you have received from any NGO or government programmes." We measure aspirations and expectations for educational attainment and current educational attainment for one named child, the child aged closest to 14 from the roster of resident household members, following the Indonesian Family Life Survey Round 4. In each of these domains, we measure, in order, the household's current position (for example, "What is the total value of all the assets owned by your household?"), the respondent's aspirations, the level the respondent would like their household to

[34] We do not include actual or imputed value of housing rental expenditure.

reach, and the respondent's expectations, the level the respondent thinks their household will reach.

The **aspirations index** is an Anderson (2008) index of responses to three questions: aspirations for children's education from the question "What level of schooling would you like *child name* to achieve?" where *child name* is the name of the child selected from the household roster; aspirations for future assets from the question "What is the level of assets that you would like your household to reach at the end of the next 10 years?"; and aspirations for future income from the question "What is the level of annual income that you would like your household to reach at the end of the next 10 years?".

The **expectations index** is an Anderson (2008) index of three measures: expectations for children's education from the question "What level of education do you think *child name* will achieve?" for the same child in the aspirations question; mean expected assets at the end of the next ten years, constructed using the procedure described below; and mean expected annual income at the end of the next ten years, constructed using the procedure described below.[35]

To calculate expectations for household assets and annual income, we elicit probabilistic subjective expectations for assets and income following Dominitz and Manski (1997) and McKenzie et al. (2013) to construct the full distribution of expected assets and income for each respondent. This approach requires four steps, run separately for each of assets and income. First, the enumerator asks respondents for their estimate of the minimum and maximum value of each variable. Second, based on the minimum and maximum values entered, the survey calculates three respondent-specific thresholds that split the range into four intervals of equal length denoted k_1,k_2,k_3. Third, the enumerator draws the values of each threshold on a visual aid. They gave each respondent 10 buttons and asked them to allocate buttons to the four intervals, with the number of buttons representing how likely the respondent thinks it is that the realisation of the variable will fall into a given interval. The respondent was required to use all 10 buttons. Finally, these thresholds and interval probabilities allow us to construct $F_{i,k}=P(y_i<Y_{i,k})$ for each respondent i and threshold $k=k_1,k_2,k_3,MAX$. We then fit a respondent-level lognormal distribution and estimate the mean of this distribution.

G.2.2 Expectations about Returns to Specific Investments

The **returns index** is made up of three separate measures of beliefs about returns to investment in different activities. First, we measure beliefs about **the percentage change in yields that will result from use of fertiliser**. Respondents are asked the amount of dry maize they would

[35] We also measure asset and income expectations directly from the questions "What is the level of assets that you think your household will reach at the end of the next ten years?" and "What is the level of annual income that you think your household will reach at the end of the next ten years?" These responses are highly correlated with the estimated means of the distributions: $\rho=0.64$ for assets and 0.80 for income. All results in the paper are similar for the directly elicited and estimated measures.

harvest in the next long rains season from a one acre plot like most other plots in their area, not using fertiliser, and the amount they would harvest when applying 50 kg of DAP (the most commonly used fertiliser locally) per acre while planting. They are told the DAP is free and it is the best, official quality of DAP. Second, we measure beliefs about **the percentage change in yields from more agricultural labour**. Respondents are asked the amount of dry maize they would harvest in the next long rains season from a one acre plot like most other plots in their area, not using fertiliser, and the amount they would earn if working 12 hours more per week. Third, we measure beliefs about the **percentage change in income from investment in university education**. Respondents estimate of the monthly income (at age 30) that their child closest to 14 (as in the aspirations question) would earn if they finished a university degree and the income they would earn if leaving school at the end of secondary schooling (form 4) with a KCSE certificate.

As a rough sense check, we compare respondents' beliefs about returns to econometric estimates from published research on comparable populations. The perceived returns are on average higher but not massively so. We interpret the comparison with caution because of the standard concerns about endogeneity and heterogeneity in returns estimates. The return to fertiliser use belief in the placebo group has median and mean of respectively 114 and 202%, while Duflo et al. (2008) estimate an experimental gross rate of return of 91% to a smaller increase in fertiliser use in maize farming in a nearby part of Kenya. The return to education belief in the placebo group has median and mean of respectively 200 and 322%, while Montenegro and Patrinos (2014) estimate that Kenyans with tertiary education earn on average 122% more than those with secondary education. However, the latter estimate does not take into account the potentially large earnings gains from education-facilitated rural-to-urban migration, which is a common strategy in this context.

G.2.3 Self-beliefs

The **self-beliefs index** is made of three scales.

1. A Schwarzer and Jerusalem (1995) scale of **self-efficacy**, with high scores indicating high general self-efficacy. Respondents rate the extent to which each of seven statements is true for them from 1 (not at all true) to 4 (completely true). The score is the sum of seven responses, ranging from 7 to 28. Items include "It is easy for me to stick to my aims and accomplish my goals" and "When I am confronted with a problem, I can usually find several solutions".
2. An adapted version of the Implicit Theories of Intelligence scale (Blackwell et al., 2007) to measure **growth mindset**. The higher the score, the more the respondent has a growth mindset. Respondents are asked to state the extent to which they disagree with each statement on a scale of 1 (agree strongly) to 6 (disagree strongly). The final score is the sum of six responses, ranging from 6 to 36. There are 3 fixed mindset and 3 growth mindset

items. We reverse code the 3 growth mindset items. We replace intelligence with skillset as the focus of the measure to make the scale applicable in our context of a rural, adult population. A sample item is "You can always greatly change what skills you have".

3. A shortened version of the Internal sub-scale from the Internal, Powerful Others and Chance (IPC) scale (Levenson, 1981) to measure **internal locus of control**.[36] The respondents state the extent to which they agree with each statement on a scale from 1 (disagree strongly) to 4 (agree strongly). The final score is the sum of five responses and ranges from 5 to 20. Items include "When I make plans, I am almost certain to make them work" and "My life is determined by my own actions".

G.2.4 Preferences

We do two incentivised tasks with the respondents to elicit time and risk preferences. To construct measures of the **discount factor** and an indicator for **present bias**, we use responses from a standard Multiple Price List (MPL) to measure time preferences over money (Coller and Williams, 1999; Andersen et al., 2008). We ask respondents to make a choice seven times between two amounts offered early or later. We do this twice: in the near time frame, we offer money tomorrow or in 15 days and in the future time frame, money in 15 days versus in 29 days. The amount offered at the earlier date is always equal to KSh 400, while the amount offered at the later date increases from KSh 360 to 1600. Among the time questions, respondents are told there is an equal chance for each choice to be randomly selected for payment. We randomize whether the respondents make the decisions in the near or future time frame first and control for the order in which questions appear.

Our measure of the **discount factor** assumes a linear utility function in money and measures the discount factor using the switch from receiving money soon to later.[37] For example, we ask respondents if they would prefer to receive KSh 400 in 15 days or KSh 440 in 29 days. If they choose KSh 440, then they have a two-week discount factor between 0.91 and 1 and we assign them the mid-point of 0.96. If they choose KSh 400, then we check their choice between KSh 400 in 15 days and KSh 700 in 29 days and so on until they switch. **Present bias** is an indicator for if respondent switches to the (higher) future amount later in the near time frame (tomorrow vs. 15 days), than in the future time frame (15 vs. 29 days). We recode this to a "not present biased" indicator to simplify comparison with other mechanisms.[38]

For **risk preferences**, we use the method developed by Eckel and Grossman (2002).

[36] We use 5 of the 8 original items and exclude items not relevant to our population, such as one about driving a car.

[37] Following Andersen et al. (2008) we use the future time frame for the discount factor to account for any transaction costs or additional risk of future income. We use the first switch point for respondents who switch multiple times.

[38] 25% of the people in the placebo group are present-biased over money, consistent with the proportion in other low-income settings: 28% in Ashraf et al. (2006), 17% in Kaur et al. (2015), 17% in John (2020) and 23% in Cassidy (2018).

Respondents are asked to make one choice from 6 gambles that are presented to them. Each gamble has two choices: A and B. There is 50% chance of receiving a low payoff (A) and 50% chance of receiving a high payoff (B). The amount the respondent will get for the option they choose will depend on whether A or B is randomly chosen by the computer. One gamble gives a certain return and the other choices increase linearly in risk (as measured by the standard deviation). Under the assumption of constant relative risk aversion (CRRA), the gamble chosen by the respondent corresponds to a coefficient of relative risk aversion and wealth level. This allows for the identification of varying levels of risk aversion. We use the rank of the choices from 1-6, increasing in level of risk. We standardise this measure to have mean zero and standard deviation one in the placebo group, then switch its sign to simplify comparison with other mechanisms. We randomised the order of the certain choice and other choices and control for the order in which questions appear.

Incentivisation works as follows. Before doing the tasks, respondents are told that the computer will draw a lottery across respondents and there is a 1 in 20 chance that they could actually be paid for these tasks. They are told that there is an equal chance for either the time question or risk question to be randomly selected by the computer for payment. Respondents are informed whether they have won anything, the amount won and when they should expect to receive the amount at the end of the survey. All payments are made via M-Pesa. Half the time the time preference task appears first and the other half the risk elicitation task appears first. We control for the order in which the questions appear.

G.2.5 Mental Health

Mental health is captured with a z-score from a depression severity score calculated from the 10-item Centre for Epidemiological Studies Depression Scale (CES-D) scale (Andresen et al., 1994). Respondents are asked to give the frequency with which they experience symptoms described from 1 (rarely or none of the time, or up to 1 days a week) to 4 (all of the time, or 5-7 days a week). The total score is a sum of all 10 items, with scoring on questions 5 and 8 reversed. We then reverse-code the scale, such that higher scores reflect lower depressive symptoms, to aid comparison with other tests of mechanisms.

G.2.6 Other Measures

1. **Mimicry of videos**: This is the sum of dummy variables coded to one if the respondent engaged in any of the following activities at endline, all of which are featured in the videos: (a) weaved baskets; (b) kept savings in a jar; (c) attended a sewing class; (d) trained as a teacher; (e) grew vegetables to sell on the market.
2. **Field officer-verified assets**: This is the total quantity of the following asset types ob-

served by the field officer: cooking pots and pans, jerry cans, chairs/sofa, tables, radio, TV, poultry house.

3. **Information recall**: An indicator variable if the respondent correctly recalled specific information contained in both videos, about the returns to education for Kenyan men, straight after watching the videos.
4. **Cognitive ability**: An inverse covariance-weighted index of three measures of working memory. These are a digit span measure, where the score is length of the longest sequence that respondents can correctly recall minus two points (Wechsler, 1958); a fluid intelligence measure, the number of correct responses to six Raven's matrices (Raven, 1990); and cognitive control, measured as the number of correct responses to 3 numerical Stroop tasks comprised of 25 number sequences (Stroop, 1992).

G.3 Development and Validation of Psychological Scales

We translated and backtranslated the depression and self-belief scales and validated the translations according to standard psychometric practice. Three enumerators did English-Luo forward translation, while another three did blind Luo-English back translation. Differences were reconciled in a group discussion. We then conducted cognitive debriefings with 22 Luo-speaking respondents to capture concepts respondents struggled to understand and identified more context-appropriate concepts to use (Pan et al., 2008; Tourangeau, 2003). We then assess each measure against several standard psychometric evaluation criteria to show scales are valid and reliable measures according to these criteria (Table G.1). We show results for the endline sample in placebo villages. Results are similar for the measures collected in the baseline sample and for the full endline sample in all villages. First, Cronbach's α values are close to 0.6 or higher for all scales (Column 3 of Table G.1). Cronbach's α measures the extent to which all the items in each scale measure the same construct, or the inter-relatedness of the items (Cronbach, 1951). Psychologists recommend that scales have α values above 0.7 (Streiner, 2003), although lower values are common for fieldworker-administered questionnaires with participants with limited literacy (Laajaj and Macours, 2021). Second, all scales meet or almost meet the recommended thresholds for three model fit criteria (Column 4-6 of Table G.1). These measure whether the empirical factor structure in our data match the structure in the original papers.[39] We use theoretical factor structures from Scholz et al. (2002) for self-efficacy, Levenson (1981) for locus of control, Abd-El-Fattah and Yates (2006) for growth mindset, and Andresen et al. (1994) for depression. Third, we also collected pilot data with 115

[39]Psychometricians recommend that scales have values above 0.95 for the comparative fit index (CFI), above 0.9 for the Tucker-Lewis index (TLI), and below 0.08 for the root mean squared error approximation (RMSEA) (Cheung and Rensvold, 2002).

respondents from non-study villages in Siaya county before the main baseline survey to validate measures. Correlations between the same measures two weeks apart from this dataset range from 0.34 to 0.58. There are no standard criteria for test-retest correlations but these are fairly high. Finally, no item had non-response rates exceeding 20%, indicating items were well-understood.

Table G.1: Results of psychometric validation

			Model fit			Two-week
Scale	# obs	Cronbachs α	CFI	TLI	RMSEA	Test-retest
Self-efficacy	1753	0.662	0.929	0.894	0.0611	0.565
Internal Locus of Control	1756	0.571	0.955	0.909	0.0649	0.377
Growth Mindset	1754	0.593	0.955	0.915	0.0651	.
Depression	1657	0.748	0.946	0.928	0.0529	0.343

Notes: This table shows the results of psychometric validation tests. Number of observations varies because we drop observations for which more than 20% of items are missing in a scale. The test-retest coefficients are reported for the pilot sample.

Appendix References

Abd-El-Fattah, S. and G. Yates (2006): "Implicit Theory of Intelligence Scale: Testing for Factorial Invariance and Mean Structure," in *Australian Association for Research in Education Conference, Adelaide, South Australia*, 1–14.

Andersen, S., G. Harrison, M. Lau, and E. Rutström (2008): "Eliciting Risk and Time Preferences," *Econometrica*, 76, 583–618.

Anderson, M. (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 103, 1481–1495.

Andresen, E., J. Malmgren, W. Carter, and D. Patrick (1994): "Screening for Depression in Well Older Adults: Evaluation of a Short Form of the CES-D," *American Journal of Preventive Medicine*, 10, 77–84.

Ashraf, N., D. Karlan, and W. Yin (2006): "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *The Quarterly Journal of Economics*, 121, 635–672.

Baird, S., C. McIntosh, and B. Ozler (2011): "Cash or Condition? Evidence from a Cash Transfer Experiment," *The Quarterly Journal of Economics*, 126, 1709–1753.

Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Parienté, J. Shapiro, B. Thuysbaert, and C. Udry (2015): "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries," *Science*, 348, 12607991–126079916.

Baron, E., T. Davies, and C. Lund (2017): "Validation of the 10-Item Centre for Epidemiological Studies Depression Scale (CES-D-10) in Zulu, Xhosa and Afrikaans Populations in South Africa," *BMC Psychiatry*, 17, 1–14.

Beegle, K., J. De Weerdt, J. Friedman, and J. Gibson (2012): "Methods of Household Consumption Measurement Through Surveys: Experimental Results from Tanzania," *Journal of Development Economics*, 98, 3–18.

Benjamini, Y., A. Krieger, and D. Yekutieli (2006): "Adaptive Linear Step-up Procedures that Control the False Discovery Rate," *Biometrika*, 93, 491–507.

Blackwell, L., K. Trzesniewski, and C. Dweck (2007): "Theories of Intelligence and Achievement Across the Junior High School Transition: A Longitudinal Study and an Intervention," *Child Development*, 78, 246–63.

Cassidy, R. (2018): "Are the Poor So Present-Biased?" *Institute for Fiscal Studies*, *Working Paper No. 18/24*.

Cheung, G. and R. Rensvold (2002): "Evaluating Goodness-Of-Fit Indexes for Testing Measurement Invariance," *Structural Equation Modelling*, 9, 233–255.

Cobb-Clark, D., S. Dahmann, and N. Kettlewell (2021): "Depression, Risk Preferences and Risk-Taking Behavior," *Journal of Human Resources*, 57, 1566–1604.

Coller, M. and M. Williams (1999): "Eliciting Individual Discount Rates," *Experimental Economics*, 2, 107–127.

Cronbach, L. (1951): "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika*, 16, 297–334.

Dominitz, J. and C. Manski (1997): "Using Expectations Data to Study Subjective Income Expectations," *Journal of the American Statistical Association*, 92, 855–867.

Duflo, E., M. Kremer, and J. Robinson (2008): "How High Are Rates of Return to Fertiliser? Evidence from Field Experiments in Kenya," *American Economic Review Papers and Proceeding*, 98, 482–488.

Eckel, C. and P. Grossman (2002): "Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk," *Evolution and Human Behaviour*, 23, 281–295.

Egger, D., J. Haushofer, E. Miguel, P. Niehaus, and M. Walker (2022): "General Equilibrium Effects of Cash Transfers: Experimental Evidence from Kenya," *Econometrica*, 90, 2603–2643.

FAO (1996): *A System of Economic Accounts for Food and Agriculture*, Rome: Food and Agriculture Organization.

Ghosal, S., S. Jana, A. Mani, S. Mitra, and S. Roy (2020): "Sex Workers, Self-Image and Stigma: Evidence from Kolkata Brothels," *The Review of Economics and Statistics*, 104, 431–448.

Grosh, M. and P. Glewwe (2000): *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study*, Washington, D.C.: World Bank.

Haushofer, J. and J. Shapiro (2016): "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya," *The Quarterly Journal of Economics*, 131, 1973–2042.

John, A. (2020): "When Commitment Fails: Evidence From a Field Experiment," *Management Science*, 66, 503–529.

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

Keller, A., J. Leikauf, B. Holt-Gosselin, B. Staveland, and L. Williams (2019): "Paying Attention to Attention in Depression," *Translational Psychiatry*, 9, 1–12.

Kilburn, K., H. Thirumurthy, C. Halpern, A. Pettifor, and S. Handa (2016): "Effects of a Large-Scale Unconditional Cash Transfer Program on Mental Health Outcomes of Young People in Kenya," *Journal of Adolescent Health*, 58, 223–229.

Laajaj, R. and K. Macours (2021): "Measuring Skills in Developing Countries," *Journal of Human Resources*, 56, 1254–1295.

LaFave, D. and D. Thomas (2016): "Farms, Families, and Markets: New Evidence on The Completeness of Markets in Agricultural Settings," *Econometrica*, 84, 1917–1960.

Laibson, D. (1997): "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 112, 443–478.

Levenson, H. (1981): "Differentiating Among Internality, Powerful Others, and Chance," in *Research with the Locus of Control Construct Vol.1*, ed. by H. M. Lefcourt, New York: Academic Press, 15–63.

MacLeod, A. and E. Salaminiou (2001): "Reduced Positive Future-Thinking in Depression: Cognitive and Affective Factors," *Cognition & Emotion*, 15, 99–107.

McKelway, M. (2021): "Women's Employment in India: Intra-Household and Intra-Personal Constraints," *Dartmouth, Working Paper*.

McKenzie, D., J. Gibson, and S. Stillman (2013): "A Land of Milk and Honey With Streets Paved with Gold: Do Emigrants Have Over-Optimistic Expectations About Incomes Abroad?" *Journal of Development Economics*, 102, 116 – 127.

Montenegro, C. E. and H. A. Patrinos (2014): "Comparable Estimates of Returns to Schooling Around the World," *World Bank Policy Research Working Paper*, 7020, 1–41.

Ozier, O. (2018): "The Impact of Secondary Schooling in Kenya: A Regression Discontinuity Analysis," *Journal of Human Resources*, 53, 157–188.

Pan, Y., A. Landreth, M. Hinsdale, H. Park, and A. Schoua-Glusberg (2008): "Methodology for Cognitive Testing of Translations in Multiple Languages," in *Survey Methodology*, U.S. Census Bureau, 2008-2, 1–8.

Prado, C., S. Watt, and S. Crowe (2018): "A Meta-Analysis of the Effects of Antidepressants on Cognitive Functioning in Depressed and Non-Depressed Samples," *Neuropsychology Review*, 28, 32–72.

Raven, J. C. (1990): *Standard Progressive Matrices: Sets A, B, C, D & E*, Oxford, UK: Oxford Psychologists Press.

Ribar, D. (2014): "How to Improve Participation in Social Assistance Programs," *IZA World of Labor*, 104.

Scholz, U., B. G. Doña, S. Sud, and R. Schwarzer (2002): "Is General Self-Efficacy a Universal Construct? Psychometric Findings from 25 Countries." *European Journal of Psychological Assessment*, 18, 242.

Schwarzer, R. and M. Jerusalem (1995): *Generalised Self-Efficacy Scale*, Windsor, UK: NFER-NELSON.

Semenova, V. and V. Chernozhukov (2021): "Debiased Machine Learning of Conditional Average Treatment Effects and Other Causal Functions," *Econometrics Journal*, 24, 264–289.

Streiner, D. (2003): "Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency," *Journal of Personality Assessment*, 80, 99–103.

Stroop, J. R. (1992): "Studies of Interference in Serial Verbal Reactions," *Journal of Experimental Psychology: General*, 121, 15.

Tibshirani, J., S. Athey, E. Sverdrup, and S. Wager (2022): "Generalized Random Forests Package," Accessed from https://grf-labs.github.io/grf/.

Tourangeau, R. (2003): "Cognitive Aspects of Survey Measurement and Mismeasurement," *International Journal of Public Opinion Research*, 15, 3–7.

Wager, S. and S. Athey (2018): "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests," *Journal of the American Statistical Association*, 113, 1228–1242.

Wechsler, D. (1958): *The Measurement and Appraisal of Adult Intelligence*, Williams & Wilkins Co.

Yaakov, M., Z. Shtudiner, O. Suhorukov, and N. Grisaru (2019): "Time and Risk Preferences and Consumption Decisions of Patients with Clinical Depression," *Journal of Behavioral and Experimental Economics*, 78, 138–145.