The Future in Mind: Aspirations and Long-term Outcomes in Rural Ethiopia*

Tanguy Bernard, Stefan Dercon, Kate Orkin
Giulio Schinaia, Alemayehu Seyoum Taffesse

19th September 2023

Abstract

Aspirations have been posited to condition the future-oriented choices of individuals and thus can play a role in the persistence of poverty or the effort to break out of it. In a randomised control trial in remote, rural Ethiopia, we assess the effectiveness of an intervention seeking to change how poor people perceive their future opportunities, alter their aspirations and, through that, modify their investment decisions. A treatment group was shown video documentaries about the lives of individuals from similar communities who escaped poverty through their own efforts and, as such, can serve as role models. Five years after the screening took place, the treated households increased future-oriented investments in agriculture and in children's education. The results can be explained by the increase in aspirations in terms of lifetime goals. Overall, this research uniquely provides evidence that a light-touch behavioural intervention can have persistent economic impacts on a poor population.

Keywords: aspirations, long-term, investment, role models
JEL codes: D03; I31

*Author affiliations: Bernard - Bordeaux School of Economics, University of Bordeaux & CNRS; Dercon - University of Oxford; Orkin - University of Oxford; Schinaia - University of Oxford; Taffesse - International Food Policy Research Institute. This paper is dedicated to the memory of Ayelech Kifle, main character of the first short documentary produced for this study, which inspired the later development of other videos. We thank the subjects of the documentaries for sharing their stories and all participants for their time over the five years of the study. The authors thank research assistants Enrico Guizzardi, Geetika Nagpal, Fanaye Tadesse, and Marc Witte, fieldwork managers Tewodros Abate, Kibrom Hirfrfot and Bezabih Tesfaye; Mekamu Kedir, Emily Schmidt, IFPRI Addis Ababa staff; Doba *Woreda* Administration, Next Studio and Tadesse Fayissa for producing documentaries, and staff at International Food Policy Research Institute and at the Centre for Study of African Economies and Oxford Department for International Development. This research was funded by the UK Department for International Development (DFID) as part of the Institutions for Pro-Poor Growth Consortium (iiG), SEVEN (Social Equity Venture Fund) and USAID through a Feed the Future Ethiopia grant. Bernard and Taffesse acknowledge the support of the IFPRI Development Strategy and Governance, and Markets, Trade and Institutions divisions.

1 Introduction

The persistence of extreme poverty has long concerned economists and policymakers. Recent theoretical research has highlighted the role that social and psychological factors can play in perpetuating it.[1] A growing body of evidence shows that interventions targeting specific psycho-social attributes can lead to short-term changes in beliefs and positive economic outcomes among low-income groups.[2] However, despite a few recent studies on the long-term effects of clinical psychological interventions on depressed adults (Baranov et al., 2020; Bhat et al., 2022), it remains unclear whether any easily scalable and population-wide behavioural intervention can have lasting impacts on economic outcomes. This paper addresses this gap: we show that a light-touch behavioural intervention has persistent economic impacts over a five-year horizon on a random sample of a population mostly living in extreme poverty.

Our intervention is grounded in the hypothesis that those living in extreme poverty may find it hard to envision a better future for themselves, lowering their aspirations, which in turn may limit their incentives to invest for the future. This idea is particularly relevant for people living in high poverty areas who may have had fewer successful role models from their community to look up to for inspiration (Durlauf, 1996; Appadurai, 2004; Ray, 2006; Genicot and Ray, 2020). The intervention is designed to increase individuals' economic aspirations using visual media by exposing participants to stories of locally successful role models that could help them envision a better future for themselves (Bandura, 1977a; La Ferrara, 2019). We conceptualise aspirations as desired goals for the future, which motivate investment and effort in order to attain them (Locke and Latham, 1990; Bandura, 1977b).

We test this intervention with a randomised field experiment in a remote, poor district in Ethiopia. Some participants were randomly invited to watch video documentaries we made about individuals from similar communities who had succeeded in agriculture or business through their own efforts. In the videos, the role models describe how they improved their socio-economic position from being poor to being relatively successful, through setting goals, careful choices, perseverance and hard work. Meanwhile, another group of participants (the placebo group) were randomly invited to watch an Ethiopian entertainment programme. A control group were simply surveyed. We collected data before the intervention, straight after the screening occurred, after six months, and again five years later.

We find that this simple intervention has improved economic outcomes after five

1. See, for example, Durlauf (1996); Benabou (1996); Bisin and Verdier (2001); Appadurai (2013); Dalton, Ghosal, and Mani (2016); Besley (2016); Genicot and Ray (2017); Lybbert and Wydick (2018).

2. Recent studies include Haushofer, Mudida, and Shapiro (2020); McKelway (2021); Angelucci and Bennett (2021); Ashraf et al. (2022); Barker et al. (2022); Ghosal et al. (2022). See Kremer, Rao, and Schilbach (2019) for a review.

years by increasing effort and investment. First, five years after the experiment, treated households report higher labour supply and more use of agricultural inputs. They spend around one extra hour working on their own farms every day and invest 21 per cent ($7 PPP) more in the purchase of seeds, fertiliser and pesticides. Second, we observe persistent increased investments in human capital among treated households. At the five-year follow-up, treated households spend approximately 39 per cent more than other study participants on their children's education. Their children have attained more years of schooling: twice as many children who were of school-going age at the time of screening have completed full primary school five years later. Third, these investments have yielded meaningful changes in living standards: while ordinarily measured food consumption at the time of the five-year follow-up is not different to untreated households, treated households report fewer months of food security, improved housing quality and have accumulated 29 per cent ($22 PPP) larger holdings of durable goods, like cellphones and household equipment. They also report a somewhat higher subjective wellbeing. We find changes in economic behaviour started soon after treatment: six months after the experiment, treated individuals had increased savings and labour supply. At the time, treatment had also increased enrolment, educational expenditure and time studying.

We find evidence consistent with the economic changes induced by our intervention being the result of increases in the individuals' aspirations and expectations for the future. We use locally validated survey measures of aspirations and expectations (Bernard and Taffesse, 2014). These capture the level of income, assets, or children's education individuals hoped to achieve in their lifetime or thought they would achieve in ten years. We find that the treatment increased aspirations and expectations in the treatment group right after the video screening and still five years later.

Our findings are consistent with a reference-dependent behavioural model (Kőszegi and Rabin, 2006) where, in line with Dalton, Ghosal, and Mani (2016) and Genicot and Ray (2017), we define aspirations in economic terms as a reference point, and show how changes in aspirations might affect effort and investment. Incorporating reference point utility changes the standard results of dynamic optimisation, predicting that exogenous shifts in goals as reference points would lead to more effort and investment with future benefits, as we find empirically.

We can plausibly exclude some alternative mechanisms through which the intervention could have led to this outcome. We measure time and risk preferences, grit, information transmission, and beliefs about the returns to technology and find no change in these. We do find some effects six months after exposure to the videos on measures of locus of control — beliefs about whether individuals or fate control people's lives — and thus cannot rule out that they contribute to changes in behaviour. However, these do not persist after five years, and thus do not appear to contribute to the durability of effects. The design of the intervention also rules out further mechanisms. Unlike in other studies that rely on

variation in exposure to real-life role models such as teachers or peers, participants receive no mentorship or support other than the one time exposure to our videos (Kearney and Levine, 2020). Exposure to video screening itself or to outsiders or being selected for the intervention do not account for effects: a placebo group shown a local entertainment programme are unaffected relative to a control group. Finally, we can also exclude that our intervention gave "false hope" or "led to frustration" with lower outcomes as a result, a possibility highlighted by some existing models of aspirations (Dalton, Ghosal, and Mani, 2016; Genicot and Ray, 2017): we see persistently higher aspirations, as well as higher aspirations gaps, and small positive effects on subjective wellbeing, rather than negative ones. Our evidence suggests that aspirations condition behaviour both in the short and long term.

Lastly, we find no systematic evidence of spillovers across individuals from the same village. We find no effects on control individuals in treated villages relative to individuals in "pure control" villages who were narrowly excluded from the original intervention and surveyed for the first time in the five-year follow-up. Variation in treatment intensity at the village-level also did not lead to different treatment effects. Overall, given our sample size, we cannot fully rule out spillovers, although there is no conclusive evidence pointing towards their existence.

This study adds to the research literature in three ways. First, we provide the first experimental evidence on the long-term effects of an intervention targeting aspirations on economic investment. Several theoretical models posit that aspirations can influence investment (Genicot and Ray, 2017; Dalton, Ghosal, and Mani, 2016; Besley, 2016; Bogliacino and Ortoleva, 2013) and several papers use observational data to document aspiration-investment associations (Janzen et al., 2017; Ross, 2019; Serneels and Dercon, 2021; Eble and Escueta, 2022). A growing number of papers show a causal link between interventions trying to shift aspirations and economic outcomes in the short-run. Some experiments use light-touch interventions similar to ours targeting women living in disadvantaged circumstances (Lubega et al., 2021; Orkin et al., 2023) or entrepreneurs (Batista and Seither, 2019); others involve more intensive training to promote future-oriented behaviour (Lybbert and Wydick, 2019; Rojas Valdes, Wydick, and Lybbert, 2021; Cecchi et al., 2022; McKenzie, Mohpal, and Yang, 2022) or planning exercises (Orkin et al., 2023). Relative to this body of work, we provide the first evidence on the long-term causal effects (beyond 18 months) of an intervention targeting aspirations to boost economic outcomes.

Second, within the literature on the psychology of poverty, we add to the currently very limited evidence on the long-term impacts of psychological interventions, by showing how a population-wide light-touch intervention can have long-term impacts on economic outcomes.[3] A growing body of intervention-based studies examine the effects of other

3. Our study has very low attrition (less than 10 per cent) compared to recent long-term follow-ups of experiments in low- and middle-income countries (Bouguen et al., 2019).

psychological characteristics on decision-making in poor settings beyond aspirations, including self-regulation, self-efficacy, grit, and preferences (Heller et al., 2016; Blattman, Jamison, and Sheridan, 2017; Ashraf et al., 2022; Campos et al., 2017; Alan, Boneva, and Ertac, 2019; Alan and Ertac, 2018; McKelway, 2021; Bossuroy et al., 2022) and of interventions directly targeting mental health (Baranov et al., 2020; Bhat et al., 2022; Haushofer, Mudida, and Shapiro, 2020; Angelucci and Bennett, 2021; Barker et al., 2022). Among these, Baranov et al. (2020) and Bhat et al. (2022) provide evidence on long-term impacts (respectively after seven and five years), including on economic outcomes, but on a specific populations of depressed adults in Pakistan and India, respectively, using targeted psychotherapy. Our study provides the first longer term evidence of a light-touch non-targeted population-wide behavioural intervention on economic outcomes, showing how overcoming internal psychological constraints faced by households can unlock investment.

Finally, we contribute to work on the effect of role models on investments, as well as their exposure through visual media. Female role models affect girls' and young women's selection into and performance in male-dominated fields in high-income countries (Greene, Sullivan, and Beyard-Tyler, 1982; Stout et al., 2011; Porter and Serra, 2020) and girls' education investments and women's fertility in low- and middle-income countries (Jensen and Oster, 2009; Chong, Duryea, and La Ferrara, 2012; Beaman et al., 2012; Bhan, 2020; Riley, 2022). Exposure to role models also affects investments in both boys' and girls' education (Macours and Vakis, 2018). We add to this literature in three ways: by providing evidence on how exposure to role models has persistent effects on adults' labour supply and investment; by using an experimental design to provide clean identification of the causal link between exposure to the documentary featuring role models, and changes in aspirations and behaviour; and by examining a range of psychological mechanisms through which role model effects might occur. In the process, we provide a further example of how using visual media with stories featuring role models can affect behaviour — for a review see La Ferrara (2016). The placebo group allows us to separate the effects due to the content of the documentaries featuring the role models, from the exposure to visual media per se.

The implications for the design of poverty-reduction interventions are potentially important. Our study illustrates that a relatively low-cost intervention to change individuals' beliefs about what is possible in the future can in turn change their economic behaviour persistently. That a light touch intervention in the form of a one-hour documentary not only induces psychological but also behavioural changes that persist after five years suggests a promising avenue for further research and poverty-related policy interventions. We are nevertheless cautious about the external validity of the specific intervention: the study area is remote with limited exposure to other forms of media, which may have contributed to the persistent and relatively substantial impacts.

The rest of the paper has the following structure. Section 2 provides a simple theoretical framework to model aspirations, combining elements from existing theoretical models of aspirations. Section 3 describes the context of the study. Section 4 discusses our intervention, design, and gives a brief description of our main estimation strategy and tests for experimental integrity. Section 5 provides the results of the intervention on the main investment decisions and indicators of household well-being five years after the experiment. It also describes effects on aspirations and alternative mechanisms. Section 6 reports a series of tests for the presence of spillovers. Section 7 concludes.

2 Theoretical framework

The paper focuses on how an intervention targeting an exogenously induced change in aspirations might affect economic decision-making. This section sets up a theoretical framework to derive predictions about the effects of such an intervention.

2.1 Setup of the reference-dependent model

Existing economic models of aspirations formation and its consequences capture the idea that achieving goals may yield utility (Heath, Larrick, and Wu, 1999; Dalton, Ghosal, and Mani, 2016; Genicot and Ray, 2017). These models use reference-dependent utility (Kőszegi and Rabin, 2006; Kőszegi, 2010) and interpret goals as reference points. In some cases, such models match observed patterns of labour supply, job search and consumer choice better than more traditional models (O'Donoghue and Sprenger, 2018).

Aspirations enter our model as a reference point: instantaneous utility $v(c_t, l_t; a_t)$ is defined not just over consumption c_t and leisure l_t, but is anchored by the aspirations one has for one's economic position a_t. More specifically, we assume that $v(c_t, l_t; a_t) = u(c_t, l_t) + z(c_t - a_t)$, with $u_{c_t}, u_{l_t} > 0$ and $u_{c_t c_t}, u_{l_t l_t} < 0$.[4] The function z can be seen as a loss-gain function: not fulfilling one's aspirations reduces welfare, so $z(c_t - a_t) \leq 0$ if $c_t \leq a_t$. Or equivalently, starting from below and getting closer to one's goal increases one's utility. Overachieving, when $c_t > a_t$, is assumed to be adding utility or $z(c_t - a_t) > 0$. This loss-gain function is assumed to be increasing and concave in c_t, i.e. $z_{c_t} > 0$, $z_{c_t c_t} \leq 0$.

We explore the effect of a change in aspirations, or the reference point, on effort and investment, in a simple multi-period model of allocating effort and resources for future benefit versus consuming more or enjoying more leisure now. We consider a unitary household, with an infinite time horizon, maximising discounted lifetime utility at each moment t, $W_t = \sum_{s=0}^{\infty} \beta^s v(c_{t+s}, l_{t+s}; a_{t+s})$, with the discount factor being $0 < \beta \leq 1$. At the start of each period t, the household has revenue y_t and assets A_t available, based

4. We use throughout the notation $\frac{\partial g(x_t)}{\partial x_t} = g_{x_t}$ for any function g.

on decisions at $t-1$. Total resources $A_t + y_t$ in each period t can be allocated to either consumption or used to produce future revenue. Revenue at $t+1$ is obtained from allocating both effort $e_t = 1 - l_t$ and capital $k_t = A_t + y_t - c_t$ in period t. The transition equation for future revenue is $y_{t+1} = f(k_t, e_t)$, with $f_{k_t}, f_{e_t} > 0$ and $f_{k_t k_t}, f_{e_t e_t} < 0$. Allowing for some depreciation δ from using capital, the transition equation for assets is $A_{t+1} = (1-\delta).k_t$.

Maximising W_t, subject to the two transition equations for revenue and assets defined for each period $t+s$, allows us to derive the following Euler equations from the first order conditions governing decisions about consumption c_t and leisure l_t:

$$u_{c_t} + z_{c_t} = \beta.(1 + f_{k_t} - \delta).(u_{c_{t+1}} + z_{c_{t+1}}) \tag{1}$$

$$u_{l_t} = \beta.f_{e_t}.(u_{c_{t+1}} + z_{c_{t+1}}) \tag{2}$$

Equation (1) governs choices between consumption today versus saving and investing for future consumption; Equation (2) governs taking leisure today or putting in effort with a return tomorrow. These are familiar Euler equations, except for the terms defined by the loss function. Without the loss function, the model yields the standard intertemporal results, whereby the marginal utility of present consumption (or leisure) will equal the discounted marginal utility of future consumption generated from returns to savings (or effort).

2.2 Model predictions from a change in aspirations

The model predicts that a change in future aspirations can affect decisions about consumption and savings, as well as about leisure and effort. If aspirations for the future (a_{t+1}) increase at t, current effort and/or investment will increase. The intuition is captured by considering how an increase in a_{t+1} affects the Euler equations. z is a concave function in its argument $(c_{t+1} - a_{t+1})$ for a given a_{t+1}. Thus $\frac{\partial z_{c_{t+1}}}{\partial a_{t+1}} > 0$. For a given past level of aspirations, a_t, an increase in aspirations for the future, a_{t+1}, will boost the right-hand side of both Equation (1) and (2). For both equations to hold simultaneously after this change, the left-hand side of each equation needs to go up too and/or the other terms on the right hand side need to go down. To restore equality in Equation (2), a reduction in leisure today is required: investment in the future through effort will increase u_{l_t} and reduce the marginal product of labour f_{e_t}. To restore equality in Equation (1) the household will need to consume less, and save and invest more at t so that future consumption increases. In turn, this decrease in present consumption will increase the left-hand side of Equation (1), as consuming less at t will increase marginal

utility u_{c_t}, as well as z_{c_t}. More savings will also reduce the marginal product of capital f_{k_t} on the right side of (1) and reduce $u_{c_{t+1}}$ until equality across both Equation (1) and (2) is restored.

It follows that someone with lower aspirations for the future will limit investment and effort relative to someone otherwise identical in all other characteristics but with higher aspirations. Laboratory studies on goals in psychology (Schunk, 1983; Zimmerman, Bandura, and Martinez-Pons, 1992) and on reference points in economics (Abeler et al., 2011; Gneezy et al., 2017) are consistent with low aspirations or goals reducing effort.

The model yields a more ambiguous prediction on how an upward shift in aspirations would affect consumption. Equation (2) offers a rule for the path of consumption, not for the level in each period. Boosting aspirations will boost future wealth, as there is more incentive to shift resources to the future for a given discount rate. In turn, increased future wealth will boost undoubtably consumption at some point in the future. Given the stronger incentives to save and invest, whether higher aspirations will also lead to higher levels of consumption in the near future will depend on individual preferences, in particular the inter-temporal substitution elasticity and other features of the underlying utility function (Deaton, 1992). In particular, the change in aspirations for the future increases the opportunity cost of consuming today. This generates both a *income effect* — the value of lifetime assets increase because they yield higher returns in the future — and a *price effect* — the opportunity cost of consumption at any moment in time increases as well. The income effect allows for more consumption at any moment in time, but the price effect will encourage to move consumption to the future. Preferences will determine when the former will outweigh the latter across the consumption path.

Finally, we highlight three implications of the assumptions of our model. First, we remain agnostic about where the reference point comes from, beyond that it is not a decision variable. Reference points have been found to be consistent with individual past attainment, reflecting endowment effects or status quo bias (Kahneman, Knetsch, and Thaler, 1986; Madrian and Shea, 2001), goals (Heath, Larrick, and Wu, 1999; Markle et al., 2018) or peer comparisons (Neumark and Postlewaite, 1998). If the reference point could be set as part of the optimisation problem, then it follows that if there is a gain from overachieving, then, to maximise utility, the reference point would be set to be as low as possible, which would be a trivial result. We also abstract from any endogenous revision of aspirations within the model, such as in response to past attainment. Second, our assumptions imply a loss from underachieving, with marginal losses increasing for higher levels of underachievement. This setup is consistent with Dalton, Ghosal, and Mani (2016)'s assumptions for underachieving, while Genicot and Ray (2017) assume a gain from overachieving, i.e. when $c_t > a_t$, but no effect from underachieving (i.e. frustration does not come at a cost). Our assumption and these other formulations of utility around the reference point yield the same underlying intuition: if aspirations are

low relative to what could be achieved, boosting aspirations will provide incentives to put in more effort. Third, we assume a unitary household. Our focus is on how aspirations affect effort and investment for the household as a whole, rather than across its individual members, as we measure economic outcomes mostly at the household-level in a survey with the household head.

3 Context and sample

Our study took place in Doba, a remote rural administrative district of Ethiopia, 380 kilometres east of the capital city of Addis Ababa. At the time of the experiment, Doba was one of the poorer districts in the country: it was one of the first districts selected in 2005 for the national social protection programme targeted at the most chronically food-insecure districts in Ethiopia.

Doba was also extremely remote: most surveyed villages were accessible only by 4x4 vehicle and some further required camel transportation. Exposure to life outside of the district was also limited. At baseline there was limited exposure to television: only 10 per cent of respondents watched TV once a week or more, 28 per cent watched at least once a month and 62 per cent watched about once a year, if ever. Only 4 per cent of the households owned a cellphone, and no household owned a television.

Over the course of our study, Ethiopia's GDP grew by almost 10 per cent annually, making it one of the world's fastest-growing economies. Yet, despite halving the poverty headcount since 2000, the official national rate in 2015 remained at around 30 per cent, using the global benchmark of $1.90 2011 PPP per person per day (International Monetary Fund, 2018). Even for Ethiopian standards, the households in our study remain extremely poor: we estimate that 69 per cent had consumption per person per day below the $1.90 PPP level in 2016, and the rest not far above this level.

3.1 Sampling and data collection

We implemented the experiment in 64 villages, selected from the Central Statistical Agency's list of 189 rural villages in Doba with a population of 50 to 100 households from the 2007 census. Within each village, we compiled a list of all households with the assistance of the community (*kebele*) leader (who runs three or four neighbouring villages). We randomly sampled 18 households from each of the 64 villages to survey, with replacement for households that were away, ill or had just given birth.

The main sample of analysis consists of 1152 households and 2112 individuals surveyed at baseline (and any subsequent follow-up) in these 64 villages. We visited villages for the baseline survey and intervention (round 1, between September and November 2010), the midline follow-up survey six months after the baseline (round 2, between March and

May the following year), and a long-run follow-up survey five years after the baseline survey (round 3, between December 2015 and January 2016), which we refer to as the endline survey. Appendix Figure A.1 shows the timeline of the surveys. Surveys were conducted at households' homes by enumerators blind to household treatment status. The household head answered questions on issues like household composition, assets and children's schooling, so all economic variables are at the household-level. We also collect individual-level information from both the household head and spouse on beliefs and preferences, such as aspirations. Spouses were interviewed separately, usually by interviewing them simultaneously by two different enumerators, either in or around their house. Appendix Section B details the construction of the variables used in our analysis.

3.2 Characteristics of the sample

Table 1 describes the economic lives and living standards of our sample at baseline. The sample consists of small farm households, on average 5.6 members, as common in rural Ethiopia. Crop agriculture, livestock products and live livestock sales make up the majority of the households' incomes. Farm herds are small, with an average livestock value of about $411 PPP per adult, corresponding to just over one cow (worth about $370 PPP). Holdings of tools are low, at $24 PPP per adult. Households hold limited savings, with only 36 per cent holding any savings and an average amount of $7.5 PPP for those who do. Education levels are low, with adult men holding on average 3 years and women 1 year of schooling. Most of the generation before the respondents had no education at all: only 13 per cent of the respondents' fathers and 5 per cent of their mothers completed any years of education. Although enrolment levels have increased with free primary education policies, 42 per cent of children aged 7 to 15 were not enrolled in school at baseline.

Being better-off in these villages is correlated with more investment in agriculture and livestock, and effort on their farms. We split the sample by terciles of the value of durable assets at baseline, a proxy for wealth. Even if for any of the indicators used, the richest tercile is by no means well-off, their levels of assets, housing quality and value, livestock, education, and food security levels are all significantly higher than the poorest tercile.

Aspirations are also higher for the richest tercile relative to the poorest tercile. We use locally validated survey measures of aspirations (Bernard and Taffesse, 2014). These capture the level of income, assets, or children's education individuals hoped to achieve in their lifetime.[5] We find that, at baseline, aspirations levels are strongly correlated

5. To measure each dimension of aspirations, respondents were asked "*What is the level of [X] that you would like to achieve?*" where *[X]* was either: (i) annual income (from all agricultural and non-agricultural activities, or social protection programmes); (ii) value of assets (including house, furniture, consumer goods like a TV and fridge and any transport vehicles); or (iii) oldest child's education. To help respondents conceptualise the level they aspired to, they were previously asked "*What is the level of [X] you have at present?*".

with wealth. Aspirations for income, wealth and education are all significantly higher for relatively better off households (Column 5). Although the sample aspirations appear high, they were reasonable given Ethiopia's rapid economic growth during the study period. In 2020, the national GDP per capita for a family of 5.6, the average in our sample, was $13,656 PPP. On average, households aspired to achieve 70 per cent more than the average GDP per household in income, slightly less in wealth, and a few more years of education beyond completing secondary school for their oldest child. We find similar differences across wealth terciles even in the "gap" between aspirations and the current level reported in each of the dimensions (Appendix Table A.1).

In comparison, there are fewer differences in other beliefs and preferences by wealth terciles (Appendix Table A.1). There are fewer patterns in risk or time preference between these groups. Poorer households showed more patience, but we see no clear patterns for other measures of risk or time preferences. Wealth did not correlate with the belief that poverty is caused by individual-specific traits or with the belief that outcomes are contingent on an individuals' behaviour (internal locus of control). However, fewer better-off individuals believed in luck or supernatural causes of poverty.

That aspirations are strongly correlated with wealth obviously does not prove anything. It is nevertheless consistent with models of aspirations and/or of reference points that argue not only for their relevance in decision-making but also that posit that their formation is linked to past attainment (Dalton, Ghosal, and Mani, 2016; Kahneman, Knetsch, and Thaler, 1986).

4 Experimental design and estimation strategy

4.1 Content of the video intervention

Our intervention consisted of inviting randomly selected individuals to a screening session within which four short documentaries were screened to the audience.[6] The documentaries narrate motivational life stories of real people, from a similar socio-economic background as the study participants, who improved their economic circumstances through hard work and by setting, working towards, and achieving goals. Each documentary is 15 minutes long and in Oromiffa, the local language in the study site. Two stories are about male and two about female characters.

The documentary had four common themes intended to make audience members re-evaluate their own aspirations through exposure to the lives of role models who were similar to them and had succeeded in improving their economic position. First, the

6. The documentaries, with English subtitles, and one of four placebo segments are available at https://www.youtube.com/channel/UCqfoNjCzt8YPjTRWQaMQfAg. Appendix Section A summarises two documentaries and one placebo segment.

Table 1: Economic activities and aspirations by terciles of durable assets

	(1) Whole sample	(2) Lower tercile	(3) Middle tercile	(4) Upper tercile	(5) p-value	(6) Observations
Assets, per adult equivalent, USD PPP						
Value of durable assets	18.72	0.00	5.49	51.25	0.00	1119
Value of tools	24.32	16.28	17.63	38.18	0.00	1111
Total savings	8.18	7.21	5.43	8.04	0.75	1110
% holding any savings	0.36	0.34	0.38	0.36	0.67	1119
% holding any credit	0.60	0.62	0.62	0.56	0.09	1115
Livestock, per adult equivalent, USD PPP						
Value of livestock	411.49	281.09	348.13	590.91	0.00	1110
Value of cattle	366.87	235.44	296.43	524.95	0.00	1118
Value of sheep or goat	39.07	28.96	30.98	54.82	0.00	1118
% cattle owners	0.85	0.76	0.87	0.92	0.00	1118
% goat or sheep owners	0.63	0.56	0.64	0.68	0.00	1118
Labour supply and endowments						
Household size	5.61	5.36	6.28	5.30	0.72	1119
Daily minutes in paid work per adult aged above 15	11.67	14.27	9.72	10.11	0.20	1109
Daily minutes on family farm per adult aged above 15	308.90	300.24	306.35	318.95	0.06	1110
Human capital investment						
Share of children at school in the 7-15 age-group	0.58	0.57	0.57	0.61	0.20	802
Schooling expenditure per child aged 7-15 USD PPP	17.47	16.92	17.65	18.41	0.53	1110
Highest education level among male adults (years)	3.43	3.02	3.22	4.03	0.00	1045
Highest education level among female adults (years)	1.08	0.88	0.97	1.45	0.00	1025
Housing and food security						
Value of house per ad. equiv. USD PPP	371.53	228.67	323.08	531.02	0.00	1082
Non-organic roof	0.51	0.30	0.56	0.70	0.00	1077
Own toilet	0.76	0.71	0.76	0.80	0.01	1079
Food security index: z-score	-0.00	-0.19	-0.04	0.21	0.00	1119
Aspirations: what would you like to achieve?						
Income (USD PPP)	22382.36	18012.02	21106.74	27358.05	0.00	2017
Wealth (USD PPP)	12816.36	9143.45	11319.29	17365.90	0.00	2025
Education (years)	14.06	13.90	13.88	14.39	0.00	2000

Notes: Sample mean for the variables reported on the left (column 1). Conditional sample means for the variables reported on the left, conditional on the household being in the lower, middle and upper terciles (Columns 2-4) of the value of durable assets (excluding tools) at baseline, an approximation for living standards. Columns 5 reports the p-value from a t-test of equality between the mean of the lower and upper tercile. Columns 6 reports the number of observations. Variables are measured at the household level (except the aspirations variables, which are measured for both the household head and spouse) at baseline. The number of observations varies slightly across rows because some respondents do not answer all questions. Livestock and durable assets are valued using self-reported hypothetical sale prices. The OECD adult equivalence scale gives weight 0.5 to each individual younger than 16 and weight 0.7 to all other adults that are not the household head. Durable assets include radios, mobile phones, jewellery, and furniture. Tools include ploughs, hoes, axes. Household savings refers to the value of savings held inside and outside the home. The value of house is assessed by asking the household head how much their house would cost to build today (in current prices), including materials and labour costs. We use a version of the United States Department of Agriculture's food insecurity questionnaire (Bickel et al., 2000) adapted for Ethiopia (Hadley et al., 2008), to construct a z-score of the weighted sum of the answers. To measure aspirations, respondents are asked the levels of outcomes the respondent would like to achieve, on three dimensions. Annual income is the amount of cash income the household earns from all agricultural and non-agricultural activities in a year. Wealth is durable wealth (including housing, vehicles, furniture and other valuable durables). Aspired education is measured as the 'years of education that you would like your oldest child to achieve'. Variables are defined in detail in Appendix B. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using the national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The conversion is described in Appendix B.1.

documentaries emphasise the importance of working hard. Second, the documentaries highlight the importance of setting goals, planning, and persisting despite obstacles. The documentaries are filmed in a motivational and inspirational style and describe the emotional and mental processes of setting, working towards and achieving a goal. Characters highlight that progress takes time and success is incremental. Third, depicted individuals take actions which are possible for the audience. The characters succeeded largely through their own efforts. In some cases, they were able to draw assistance from community members or local agricultural extension agents, but in no case did they rely on external support that would not be available to others. Any concrete information in the videos was unlikely to be new for viewers, although the documentaries may have made existing information more salient, which we test for in Section 5.2.3. Fourth, all the subjects take slightly different courses of action to those around them: starting a small business, diversifying their source of income, or improving their farming practices.

As suggested by social learning theory, we ensure the characters featured in the documentaries were very similar to their audience. In psychology, social learning theory argues people often change goals or aspirations based on a "vicarious experience" of another person's life, either through observing them directly or through vivid stories about them (Bandura, 1977a,b). Stories often create a sense of identification between the subject and the viewer: a viewer imagines "being that character" (Cohen, 2001, 251). Stories can also be resonant and memorable, "transporting" the viewer and making them more likely to accept the information they contain (Green and Brock, 2000; Slater and Rouner, 2002). In economics, La Ferrara (2016) and Mani and Riley (2021) summarise recent examples of video-based narratives that aim to shift behaviour.

We selected the subjects of the documentaries by inviting agricultural extension agents and NGO staff to submit descriptions of life stories of potential role models who lived in their area. We worked with an Ethiopian production company to film their life stories. All the subjects were ordinary rural residents who were either initially poorer than those around them or of similar socio-economic status, so their achievements would seem attainable to our sample.[7] When those who saw the documentary were asked in the six month follow-up about the story they found the most relevant to them, 52 per cent of audience members thought the documentary subjects had initially been worse off than they currently were. However, 73 per cent of the audience said that the documentary subjects eventually became better off than they were currently.

At the time of the screening, being shown a video in itself may have been a rare event in these villagers' lives. To account for potential changes as a result of the screening alone, we also invited another group of households, in the same villages, to a "placebo"

7. Three documentary subjects were from other districts in the region and one was from a neighbouring region. It was almost impossible that respondents would know anyone in the videos and there is no evidence that this happened.

screening of an Ethiopian comedy TV show about rural life. The placebo consisted of four 15-minute segments of the comedy TV show that we selected for its entertainment value only.

4.2 Randomisation and compliance

We compare households that were invited to watch the documentaries (treatment group) to a placebo group as well as to two control groups surveyed but not shown any videos, one sampled inside the treated villages, and the other sampled from non-treated villages.

Our randomised design has three elements, illustrated in Appendix Figure A.2. The first element is an individual-level randomisation. In each of the 64 villages, we randomly allocated 18 sampled households into treatment, placebo and control groups. Our main analysis compares those treated households with the placebo and control groups. Comparing the treatment with the placebo group identifies the effects of the intervention, holding constant exposure to media and outside facilitators. Comparing the treatment with the control group identifies the policy-relevant effect of the whole intervention, assuming no spillovers.

The two other elements aim to identify potential spillovers of the intervention. The second element of our design involved setting up a pure control group of 10 additional villages from our original sampling frame and use them as an alternative counterfactual to test for within-village spillovers. We allocated villages to the intervention or the pure control group based on logistical considerations that we discuss in Section 6, along our tests for spillovers once we also include this group. Our third element was to set up a randomised saturation design (Baird et al., 2018), randomly splitting the 64 villages where we ran the intervention into two groups of 32. In one group of villages, "treatment-intense" villages, an additional 18 households per village were invited to watch the documentaries, but were not surveyed. In the second group of villages, "placebo-intense" villages, an additional 18 households per village were invited to placebo screenings. We exploit this saturation design in Section 6 to test for differential effects of our intervention by intensity of treatment at the village-level.

The analysis presented in the remainder of this section and in Section 5 focuses on the sample of 64 villages, without including the pure control group.

Compliance with our individual-level randomisation was high. At the end of their baseline interview, the household head and spouse in treatment and placebo households received non-transferable tickets for a screening session in a few days time.[8] Household heads and their spouses had the same treatment status and both were invited to the screening. On screening days, a dedicated team of facilitators checked farmers' identity

8. 95 individuals were single or widowed so the household was only given one invitation.

and the date and time of the ticket. Only 2 per cent of the surveyed individuals or households did not comply with treatment allocation, by either missing their screening or going to the wrong one (Appendix Table A.5). There are no differences in compliance rates across treatment and placebo groups.

4.3 Empirical strategy

Our main specification is:

$$y_i = \alpha + \delta T_i + \rho P_i + X'_{i1}\pi + \varepsilon_i \tag{3}$$

where y_i is a household-level outcome, $T_i = 1$ if a household was invited to watch the documentary, $P_i = 1$ if they were invited to watch the placebo movie and the omitted category is within-village control households. X_{i1} is a pre-specified vector of village-level fixed effects and controls measured at baseline: the age, gender, marital status and highest school grade completed are for the head of the household.

We will use Equation (3) to test our predictions: first, that after five years our intervention indeed increased effort and investment, such as in productive activities and education; second, and how this impacts standard of living indicators; and third, that the intervention mechanism is through rising aspirations and not through alternatives such as risk and time preferences, information transmission or beliefs about returns to innovation.

For aspirations, beliefs and preferences outcomes, which we observe separately for household heads and spouses, y_i is an individual-level outcome, and we control for age, gender, marital status and education of each individual and standard errors are clustered at the household level, the unit of randomisation. Appendix Section C.3 tests the main results for robustness to controls for the baseline value of the outcome. We pre-registered analysis for the five-year follow-up. In Appendix Section C we provide a list of deviations from the registered Pre-Analysis Plan.[9]

In the analysis, a number of related variables are grouped within table panels following our plan. A table panel corresponds to a group of variables which link to the same theoretical concept in the model. To correct for multiple testing, we use the Benjamini, Krieger, and Yekutieli (2006) resampling procedure, which we apply for each panel in the reported tables. In other words, we calculate sharpened q-values which correct p-values for multiple tests across outcomes within each panel, but do not adjust p-values across all of the outcomes.

To summarise impacts five years after the experiment, we report impact estimates on standardised inverse-covariance-weighted indices (Anderson, 2008) constructed from all

9. See `https://www.socialscienceregistry.org/trials/1483` for the trial registration.

outcomes reported in our main exhibits. Following Bessone et al. (2021) and Kling, Liebman, and Katz (2007) we also aggregate the standardised indices into a single omnibus index. We focus on these indices to test for heterogeneous effects and for the presence of spillovers in our experiment in Section 5.4 and 6, though we had not pre-specified these summary indices in our analysis plan.

4.4 Balance and attrition

Appendix Table A.2 suggests few imbalances in demographic characteristics across treatment groups. The maximum pairwise difference between treatment groups across demographic variables is 0.13 standard deviations, a relatively small difference. The exception is that there are imbalances in the number of children age 7 to 15 which are robust to correction for multiple hypothesis testing, with slightly more children in treated households. We add a control for the number of children in the household at baseline in our main specification (and all alternative models) analysing educational outcomes.

Attrition is low, with 94 per cent of baseline respondents re-interviewed after five years. Few covariates predict attrition, and a joint F-test shows that key covariates have no significant effect on attrition in any follow-up rounds. However, individuals invited to the documentary screening are slightly more likely to respond in the five-year follow-up after controlling for covariates. Individual attriters come from slightly smaller households and are 5 per cent less likely to have lived outside the district in the last six months. Given overall attrition rates are low and the covariates differences are also small, we do not believe these differences would affect our main results significantly. Household-level attrition is even lower, with a response rate of 96 per cent after five years, which is notably high compared to other long-run follow-ups of randomised controlled trials in development economics (Bouguen et al., 2019).

5 Results

This section first presents results five years after the intervention, based on the predictions from our conceptual framework. Next, we discuss results on a smaller subset of outcomes we collected after six months. Finally, we discuss effects on potential psychological mechanisms which might explain effects.

All tables follow the same structure. The columns present estimates of the parameters in Equation (3): δ (Column 1), ρ (Column 2), a test for $(\delta - \rho) = 0$ (Column 3), and the mean of the dependent variable in the control group (Column 4). A significant treatment effect compared to the control group (Column 1) indicates that the intervention had an overall impact; whereas a significant difference from the placebo effect (Column 3) indicates the impact of the intervention, holding constant having attended a screening.

5.1 Effects on economic outcomes five years after the screening

5.1.1 Effort, investment and productive assets

In Table 2, we show that the intervention had an impact on effort and investment in productive activities in these communities, most notably related to agriculture. After five years, households exposed to the documentaries work significantly more than both the control and placebo group, robust to multiple hypothesis testing (top panel). The effect is equivalent to about 7 per cent of the control mean, or nearly an hour a day across all adult household members. As most households have one female and one male adult member, this is roughly half an hour per spouse and per day.[10]

The treatment increased investments in modern inputs (second panel), especially on the extensive margin (whether or not the household has spent any resources on these inputs). Treated households are 6 percentage points more likely to have invested in modern agricultural inputs like seeds and fertiliser than the placebo group and 14 percentage points more likely to have invested in modern livestock inputs. For improved seeds and inorganic fertilisers, this translates into a 22 per cent increase in overall spending (intensive margin) compared to the households in the control group, but we cannot reject the absence of difference with the placebo group. For livestock inputs, intensive margin effects are positive but not significant. Consistent with increases in work supplied to the family farm, treated households are less likely to hire non-family labour in crop cultivation activities relative to the control group. There is no change in land area under cultivation, potentially because land is allocated by local authorities with no possibility to buy or sell land, while rental markets are also limited.[11]

The third panel explores the intervention's impact on household productive asset holdings, defined as those that may be used in agriculture or businesses. Treated households have higher values of productive tools compared to control and placebo households, significant relative to controls (third panel). The value of their livestock holdings are 9-13 per cent higher than in the control and the placebo group, respectively. The statistical significance of our treatment effects on livestock and productive tools can be sensitive to different specifications. Here we report the pre-specified specification. Results are similar but not significant when adding controls and the baseline outcomes, as reported in Appendix Table A.10.

We combine all outcomes from the first three panels of Table 2 into a single agricultural investment index. Treated households significantly increased this index of investment by

10. Results remain broadly robust to alternative specifications presented in Appendix Table A.10. The magnitude of the effects remains similar across different models, but when controlling for the baseline value of the outcome, the q-value on the treatment effect goes up to 0.14 and we do not find a significant difference between the treatment and placebo groups.

11. Only 14 households rented any land in and four households rented out any land at the five-year follow-up.

0.14 and 0.18 standard deviations relative to the placebo and control group, five years after exposure to the role models in the videos. Overall, Table 2 gives support to our theoretical predictions.

5.1.2 Educational investments

In rural Ethiopia, parents often perceive their children's education as a means to economic security in their old age (Woldehanna et al., 2008). We assess education-related investments and outcomes through enrolment and grade attainment, time in school and studying, and school-related expenses. We look at two cohorts of children for education-related outcomes, besides school-related expenses. These cohorts were pre-specified and correspond to the primary and post-primary school-going ages at the time of our follow-up (see Appendix Figure A.1 for a timeline). "Cohort 1" are aged 16 to 20 (post-primary school-going age) at the five-year follow up and 11 to 15 (upper primary school-going age) at the time of the intervention. "Cohort 2" are those aged 7 to 15 (primary school-going age) during the five-year follow-up and 2 to 10 during the intervention. We study all households in the sample, including 71 households without children in this age range, to make sure our results are comparable with other findings.

The intervention increased investment in children's education among children of post-primary school-going age at the five-year follow-up ("Cohort 1"). The first panel of Table 3 shows the treatment increases the number of children in a household aged 16 to 20 enrolled in school at endline by 35 per cent relative to both placebo and control group. In the control and placebo groups, 0.17 children in this age group per household are enrolled, compared to 0.23 children per households in the treatment group. However, the result compared to the placebo group is marginally not robust to multiple hypotheses testing.[12] Children aged 16-20 from treated households spend more time in school (relative to both the placebo and control groups) and studying (although differences are only significant relative to the control group). Most notably, there is an increase in education attainment in this group: they are 8 percentage points more likely to have completed upper primary school, relative to the control group. The increase in attainment is nearly a doubling relative to the placebo and control group, albeit from a very low base. Only 39 of our control group households, 7 per cent, report having children aged 11-15 at the time of the intervention who have completed upper primary school.

The second panel shows more modest effects on children of primary school-going age at the five-year follow-up ("Cohort 2"). There are no significant increases in enrolment, although this may reflect higher overall enrolment rates. Primary education enrolment

12. As noted in Section 4.4, all outcomes in Table 3 control for the number of children aged less than 16 at baseline since there is a baseline imbalance in the number of children. The increase in enrolment loses statistical significance in our robustness specifications, but still represents at least a 23 per cent increase in the number of children in school aged 16 to 20.

rates increased from 57 to 65 percent in the control group between baseline and the five-year follow-up. There are marginally significant increases in time at school and studying, of a similar magnitude to the older age group, but these are more noisily estimated and not robust to multiple hypothesis testing.

In the third panel, we show that treatment increases schooling expenditures five years after the intervention. Schooling expenditures in the treatment group are 46 per cent higher than in the control group and 35 per cent higher than in the placebo group.

Overall, the treatment increases an index of all outcomes in Table 3 by 0.23 and 0.21 standard deviations relative to the placebo and control group respectively, five years after exposure to the video intervention. This is again consistent with a model where higher aspirations lead to higher investment, but in an even longer-term investment than in agriculture.

5.1.3 Consumption, durable goods and well-being

Table 4 shows the impact on indicators of the standard of living five years after the screenings. We find that the intervention after five years increased wealth in the form of consumer durables and housing, and improved housing quality and some indicators of food security and subjective wellbeing, albeit not indicators of current food and basic non-food consumption as we measured them.

Treated households perceive themselves to be less at risk of hunger (top panel). Households have had fewer periods without food. We ask the number of months in the last year that the household had problems satisfying their food needs. Treated households faced 0.32 and 0.35 fewer months with difficulty satisfying food needs relative to the control and placebo groups, respectively — the control group faced 2.71 months with these difficulties. However, there is no difference between groups on a qualitative scale measuring food security capturing, for example, how frequently households skip meals or run out of money to buy food (Bickel et al., 2000).

There are few effects on food or frequent non-food consumption and marginal increases in non-food infrequent expenses, such as on clothing, services or ceremonies (second panel, Table 4). Treated households also reported higher values for a self-reported measure of general economic position relative to both the control and placebo group, though this increase is only significant at 10 per cent level and not robust to multiple hypothesis testing. As was discussed in Section 2.2, how consumption is affected by the intervention, even after five years, depends on the individuals' preferences. Our prediction is that the intervention increases lifetime wealth, which might increase current consumption, if income effects dominate. However, treated individuals may continue to move spending to the future if (intertemporal) substitution effects dominate, reducing current consumption. Hence, the effects on consumption of the intervention are theoretically ambiguous; the

Table 2: Effort, investment and productive assets

After five years	(1) Treatment	(2) Placebo	(3) Treat. vs. placebo	(4) Control mean (SD) Total obs.
Labour effort:				
Daily minutes working	56.19**	10.19	46.00*	750.26
	(23.91)	(24.69)	(24.99)	(316.21)
	[0.06]*	[0.95]	[0.17]	1075
Daily minutes in leisure	1.37	-32.88	34.25	1979.38
	(55.99)	(53.82)	(56.77)	(754.33)
	[0.98]	[0.95]	[0.55]	1076
Spending on family crop labour (USD PPP)	33.33*	1.27	32.06	387.81
	(19.73)	(19.39)	(20.08)	(258.03)
	[0.14]	[0.95]	[0.17]	1079
Agricultural investment:				
% with any spending on modern crop inputs	0.10***	0.04	0.06*	0.58
	(0.03)	(0.03)	(0.03)	(0.49)
	[0.01]***	[0.51]	[0.30]	1089
Spending on seed or fertiliser (USD PPP)	7.33**	3.80	3.53	33.49
	(3.07)	(3.32)	(3.31)	(43.54)
	[0.04]**	[0.51]	[0.39]	1078
% with any spending on feed or vet supplies	0.10***	-0.04	0.14***	0.45
	(0.03)	(0.03)	(0.03)	(0.50)
	[0.01]***	[0.51]	[0.00]***	1089
Spending on feed or vet supplies (USD PPP)	2.68	-1.84	4.52	29.30
	(4.81)	(4.81)	(4.63)	(70.92)
	[0.67]	[0.80]	[0.39]	1081
% with any spending on hired crop labour	-0.05**	-0.02	-0.03	0.36
	(0.02)	(0.02)	(0.02)	(0.48)
	[0.04]**	[0.51]	[0.39]	1089
Spending on hired crop labour (USD PPP)	-1.30	-4.97	3.67	54.16
	(5.45)	(5.51)	(5.42)	(93.01)
	[0.81]	[0.51]	[0.50]	1078
Area cultivated (hectares)	0.01	-0.01	0.02	0.55
	(0.02)	(0.02)	(0.02)	(0.30)
	[0.67]	[0.80]	[0.39]	1071
Assets:				
Value of livestock (USD PPP)	184.58	-124.53	309.11**	2018.22
	(135.92)	(130.92)	(130.43)	(1921.09)
	[0.17]	[0.34]	[0.04]**	1080
Value of tools (USD PPP)	27.51**	12.06	15.44	106.02
	(11.60)	(12.35)	(13.66)	(126.90)
	[0.04]**	[0.34]	[0.26]	1077
Summary index:				
Agricultural investment index	0.18***	0.03	0.14**	-0.00
	(0.07)	(0.07)	(0.06)	(1.00)
	[0.01]***	[0.94]	[0.03]**	1090

Notes: OLS estimates of within-village treatment and placebo effects five years after the intervention (columns 1-2). Column 3 tests for differences in parameters obtained in first two columns. The comparison group comprises households from the 64 treated villages that were not invited to any screening. Column 4 displays the control mean, standard deviation, and total number of observations. All columns control for village fixed effects and characteristics of the household head: age, years of education, an indicator for being single, and an indicator for being male. Heteroskedasticity-robust standard errors are in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The unit of observation is the household. The number of observations varies slightly across rows because some respondents do not answer all questions. Time spent on work and leisure for each adult member on a typical day in March, reported by the household head. Crop inputs include seeds, fertilizers, and pesticides. Livestock inputs include animal feed and veterinary supplies. Spending on family crop labour and hired labour is the product of the average village daily wage and the number of person-days of family or hired labourers in the most recent long rains season, respectively. Productive assets and livestock are valued using self-reported replacement costs and sale prices, respectively. Land plot areas are converted to hectares from local units. The agricultural investment index is a weighted average of all these outcomes, with leisure time re-coded as negative, following Anderson (2008). The q-values for the agricultural investment index are calculated across all other summary indices reported in Appendix Table 7.

Table 3: Educational investments

After five years	(1) Treatment	(2) Placebo	(3) Treat. vs. placebo	(4) Control mean (SD) Total obs.
Cohort 1: Children of post-primary school-going age				
Children aged 16-20 in school	0.06*	-0.00	0.06*	0.17
	(0.03)	(0.03)	(0.04)	(0.41)
	[0.08]*	[0.96]	[0.11]	1078
Daily minutes in school for children aged 16-20	30.50**	0.50	30.00**	58.64
	(12.92)	(11.36)	(13.27)	(149.88)
	[0.04]**	[0.96]	[0.05]**	1077
Daily minutes studying for children aged 16-20	7.86*	0.59	7.27	17.82
	(4.52)	(4.25)	(4.90)	(52.12)
	[0.08]*	[0.96]	[0.14]	1070
Children aged 16-20 that attained 8th grade	0.08***	0.01	0.07**	0.07
	(0.03)	(0.02)	(0.03)	(0.26)
	[0.01]***	[0.96]	[0.05]**	1078
Cohort 2: Children of primary school-going age				
Children aged 7-15 in school	0.01	-0.07	0.08	1.22
	(0.07)	(0.07)	(0.07)	(1.18)
	[0.86]	[0.48]	[0.26]	1078
Daily minutes in school for children aged 7-15	11.47	-34.37	45.84*	527.12
	(25.84)	(25.12)	(25.31)	(437.21)
	[0.86]	[0.48]	[0.21]	1068
Daily minutes studying for children aged 7-15	15.13*	5.54	9.59	91.29
	(8.36)	(8.09)	(8.57)	(115.61)
	[0.21]	[0.49]	[0.26]	1069
For all children				
Schooling expenditure (USD PPP)	8.20***	1.32	6.88**	19.17
	(2.86)	(2.54)	(3.06)	(32.73)
				1074
Summary index:				
Educational investment index	0.21***	-0.02	0.23***	0.00
	(0.07)	(0.06)	(0.07)	(1.00)
	[0.00]***	[0.94]	[0.00]***	1082

Notes: OLS estimates of within-village treatment and placebo effects five years after the intervention (columns 1-2). Column 3 tests for differences in parameters obtained in first two columns. The comparison group comprises households from the 64 treated villages that were not invited to any screening. Column 4 displays the control mean, standard deviation, and total number of observations. All columns control for village fixed effects and characteristics of the household head: age, years of education, an indicator for being single, and an indicator for being male. All regressions additionally control for the number of children aged 0-15 at baseline to account for the baseline imbalance in the number of children. Heteroskedasticity-robust standard errors are in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The unit of observation is the household. The number of observations varies slightly across rows because some respondents do not answer all questions. We label "Cohort 1" those children aged 16 to 20 at the five-year follow up, who were 11 to 15 at the time of the intervention. We label "Cohort 2" those children aged 7 to 15 at the time of the five-year follow-up, who were aged 2 to 10 at the time of the intervention, so some but not all were of school-going age. We examine all households in the sample, including 71 households who have no children in this age group in any of the rounds, to ensure the sample is comparable with other results. Daily minutes of an activity are the sum of schooling-age household members' daily minutes. School expenditures include the amount spent on uniforms, stationery and books, textbooks, and donations to the school. We do not disaggregate schooling expenditure by age group but measure it for the whole household. The educational investment index is an inverse-covariance-weighted average of all outcomes reported in the table, following Anderson (2008). The q-values for the educational investment index are calculated across all other summary indices reported in Appendix Table 7.

findings are consistent with the substitution effect balancing out the income effect, at least in the time frame considered and for our measures of food and frequent non-food consumption. The goods and services included in our measure of infrequent non-food consumption, such as clothing or ceremonies, are likely to have a higher income elasticity and appear to have a dominating income effect. Finally, none of these measures include any estimate of the service flow value from consumer durables or housing — goods that also are likely to have a higher income elasticity, and may have been accumulated since the intervention.

In fact, households report a higher stock of consumer durables such as furniture, kitchenware or phones, aggregated in our results as durable assets (third panel, Table 4). They report 31 per cent higher value of these assets, suggestive of more spending on goods with a higher income elasticity and therefore a perceived lifetime income effect. Treated households have invested more in the quality of their housing: they report an increase in the estimated value of their house (measured as the cost of rebuilding it, in materials and labour) that is 30 and 25 per cent higher than the control and the placebo groups respectively. These effects are robust to alternative specifications and multiple hypotheses testing. This result is consistent with direct observations by our enumerators: treated households are more likely to have been found to have a non-organic roof and their own toilet facility than the control group, although differences are not significant relative to the placebo group.

Treated households also score significantly higher on a Cantril ladder of self-reported wellbeing (fourth panel, Table 4). Treated participants score about a quarter of a step higher than control and placebo groups, although there is no significant effect when they are asked the same question in relation to happiness rather than life satisfaction.

Overall, these patterns suggest that treated household have (modestly) improved their standard of living, in addition to having increased their effort and investments. The treatment effect on an index of all outcomes in Table 4 (combining the outcomes reported in the top four panels) is positive but not statistically significant relative to both the placebo and control group. Most of the increase is driven by changes in the treated household's perceptions of food security, value of housing and durable assets.

5.2 Where do these results come from?

5.2.1 Early impact

Our video intervention began to change household behaviour soon after the screenings. In Table 5, we report on a shorter survey, collected six months after the experiment, to understand some of the behaviours which led to persistent changes in outcomes. Bernard et al. (2014) reported preliminary results from this survey soon after the intervention. The patterns are consistent with what was observed five years later, although the short-run

Table 4: Consumption, durable goods and well-being

After five years	(1) Treatment	(2) Placebo	(3) Treat. vs. placebo	(4) Control mean (SD) Total obs.
Food security:				
Months of food insecurity	-0.32**	0.03	-0.35**	2.71
	(0.14)	(0.15)	(0.14)	(2.13)
	[0.05]*	[0.85]	[0.03]**	1088
Food security index: z-score	-0.06	-0.10	0.04	0.48
	(0.06)	(0.06)	(0.06)	(0.92)
	[0.31]	[0.20]	[0.54]	1084
Consumption:				
Food consumption per ad. equiv. monthly (USD PPP)	-1.98	-2.29	0.32	53.91
	(2.05)	(1.92)	(2.07)	(29.98)
	[0.33]	[0.58]	[0.88]	1076
Frequent non-food per ad. equiv. (1m recall, USD PPP)	0.44	0.04	0.40	4.08
	(0.28)	(0.28)	(0.30)	(3.69)
	[0.19]	[0.94]	[0.25]	1076
Infrequent non-food consumption per ad. equiv. monthly (12m recall, USD PPP)	0.70	-0.54	1.24**	7.47
	(0.51)	(0.43)	(0.48)	(6.35)
	[0.21]	[0.58]	[0.05]*	1079
Spending on alcohol and tobacco (USD PPP)	0.20	0.04	0.17	0.80
	(0.13)	(0.12)	(0.13)	(1.66)
	[0.19]	[0.94]	[0.25]	1078
General economic position (scale 1 to 4)	0.09*	0.00	0.09*	2.10
	(0.05)	(0.05)	(0.05)	(0.73)
	[0.19]	[0.94]	[0.21]	1088
Non-productive durables and housing:				
Value of durable assets excluding tools (USD PPP)	21.87**	-3.05	24.93**	70.55
	(10.74)	(9.22)	(11.18)	(127.39)
	[0.05]*	[0.74]	[0.05]*	1077
Value of house (USD PPP)	412.38***	62.20	350.18***	1384.27
	(93.87)	(87.04)	(93.47)	(1235.57)
	[0.00]***	[0.63]	[0.00]***	1076
Non-organic roof	0.06**	0.04	0.02	0.68
	(0.03)	(0.03)	(0.03)	(0.47)
	[0.05]*	[0.39]	[0.49]	1087
Own toilet facility	0.07*	0.04	0.02	0.38
	(0.03)	(0.03)	(0.03)	(0.49)
	[0.05]*	[0.39]	[0.49]	1088
Wellbeing:				
Best life	0.23**	0.06	0.17	4.83
	(0.11)	(0.11)	(0.12)	(1.80)
	[0.09]*	[0.61]	[0.28]	1909
Happiest life	0.11	0.12	-0.01	6.05
	(0.14)	(0.14)	(0.14)	(2.19)
	[0.42]	[0.61]	[0.95]	1909
Summary index:				
Welfare index	0.07	-0.02	0.09	0.00
	(0.07)	(0.07)	(0.08)	(1.00)
	[0.38]	[0.94]	[0.25]	1092

Notes: OLS estimates of within-village treatment and placebo effects five years after the intervention (columns 1-2). Column 3 tests for differences in parameters obtained in first two columns. The comparison group comprises households from the 64 treated villages that were not invited to any screening. Column 4 displays the control mean, standard deviation, and total number of observations. All columns control for village fixed effects and characteristics of the household head: age, years of education, an indicator for being single, and an indicator for being male. Heteroskedasticity-robust standard errors are in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The unit of observation is the household, except for subjective well-being outcomes that are observed for both household head and their spouse. The number of observations varies slightly across rows because some respondents do not answer all questions. Months of food insecurity are defined as the number of months in the last 12 Ethiopian months that the household had problems satisfying their food needs. We use a version of the United States Department of Agriculture's food insecurity questionnaire (Bickel et al., 2000) adapted for Ethiopia (Hadley et al., 2008), to construct a z-score of the weighted sum of the answers. Food consumption (7 day recall) and non-food consumption (30 day and 12 month recall) are reported per adult equivalent (PAE) and converted into monthly figures (the 7 day recall is divided by 7 and multiplied by 30, the 12 month recall is divided by 12). Adult equivalents are constructed using the OECD scale, and the results are robust to being generated per capita. We collect disaggregated data on the source of food consumed in the past 7 days: purchased, produced for self consumption, and received as gifts and loans. We follow Beegle et al. (2012) to construct food prices. Non-food consumption with a 30 day recall is the sum of expenses related to: toiletries, transportation costs, mobile phone costs, energy, cigarettes and tobacco, repair, tailor, barber, other services and other small purchases (less than 100 ETB, or 11.5 USD PPP). Non-food consumption with a 12 month recall is obtained from expenses related to: clothing and footwear, utensils, bedding, school expenses, health expenses, funerals, weddings, religious expenses, contribution to community projects, land taxes and other large purchases (more than 100 ETB, or 11.5 USD PPP). Spending on alcohol and tobacco is measured for the 30 days before the survey. General economic position is measured on a scale from 1 to 4, where 4 corresponds to the household reporting that they are "doing well [and are] able to meet household needs by own efforts and 1 corresponds to the household reporting that they are "unable to meet their needs [and rely on external support]". To measure durable assets we collect the number durable assets (such as furniture, kitchenware, and phones) owned by the household, as well as the replacement value of each asset. Value of the house is calculated by asking the household how much their house would cost to build today (in current prices), including materials and labour costs. The roof variable is coded in to reflect relative quality of the building materials and the sanitation facilities are coded in to reflect the degree of privacy or excludability. The subjective well-being is measured using two items indicating best and happiest life. Best life is measured by showing respondents a picture of a ladder with 10 steps (Cantril, 1966). Respondents think of a ladder, with the best possible life for them being a 10, and the worst possible life, and rate their current position from 0 to 10. The welfare index is an inverse-covariance-weighted average of all outcomes reported above in the table, with months of food insecurity in the last year and consumption of sin goods recoded to be negative, constructed following Anderson (2008). The welfare index averages over the household head's subjective well-being outcomes. The q-values for the welfare index are calculated across all other summary indices reported in Appendix Table 7.

findings were noisier.

In the first panel, we show that households had already changed their labour supply decisions in response to the intervention: consistent with the results five years after the screening, they had increased time spent on the family farm. The effects after six months are somewhat smaller than those reported in Table 2, so that differences are not significant relative to the control group but only relative to the placebo group. Comparison across rounds are only suggestive, as the five-year measure of labour supply also included off-farm employment, which we did not collect at the six-month follow-up. However, these results indicate that labour supply increased soon after the intervention.

In the second panel, we show that soon after intervention, treatment nearly doubled stocks of savings, a difference significant relative to both the placebo and control groups. This is consistent with an increase in future-oriented behaviour and with increases in the value of assets that we observe after five years. The large increases in percentage terms are partly due to a low mean in the level of savings at the time of the survey. At baseline, only 36 per cent of the control group had any savings, which amounted on average to \$7.50 PPP. Savings behaviour is a good short-run indicator of increased propensity to invest, as it is unlikely poor households could immediately make new asset purchases given limited resources. Those effects do not persist in the long-run (Appendix Table A.17) suggestive of cash savings that were later invested in relatively lumpy assets such as livestock as reported in Table 2. We do not observe increases in actual loans, although households may struggle to access credit. We observe some positive effects on a variable capturing the hypothetical amount individuals would ask for if offered a ten-year loan with no interest, although there is also a large placebo effect on this variable.

Finally, we examine if the treatment had already induced investments in education early on. We focus our discussion on children whose outcomes after five years had been reported on in Table 3. As in Table 3 before, we add as a control the number of children aged 0 to 15 at baseline, which was imbalanced at baseline. We separately analyse the outcomes of children aged 7 to 10, of lower primary school-going age at the six-month follow-up, and those aged 11 to 15 that were of upper primary school-going age. Six months after the screening, we find some small and marginally significant changes mostly in this older cohort. This older cohort corresponds to those children for whom we found higher education attainment in the five year follow-up (Cohort 1); whereas the younger cohort corresponds to those in Cohort 2 for whom we collected educational outcomes at the six-months follow-up, since we did not measure these outcomes for children below school-going age. There is a 16 per cent increase in the number of children enrolled in the treatment group relative to the control group, though this is marginally not significant when correcting for multiple hypotheses testing. There is a 18 per cent increase in the time spent studying relative to the control mean, although no effect on time spent in school. We also find a 24 per cent increase in school-related household expenditures

relative to the control group, but no significant difference relative to the placebo. There is a slightly larger change relative to the control mean after five years (42 per cent) than after six months (24 percent). However, comparison of effects across the two follow-ups is not straightforward, as we used quarterly recall for the six-months follow-up and annual recall after five years and surveys occur at different times in the school year.

In sum, we find evidence that households had already made some changes in future-oriented behaviour after six months. Although we do not measure all variables we capture in the long-term survey in this shorter-term follow-up, where we have similar variables in both rounds, we observe clear consistency across the two rounds in patterns of behaviour change. Many of the changes households make at six months — in labour supply, education investment and asset accumulation — persist after five years.

5.2.2 Impact on aspirations and expectations

Next, we provide some evidence on the potential psychological mechanisms at play. We collected data on aspirations at baseline, straight after the screening, after six months and after five years. Our aspiration measures capture three components: the individual respondents' level of income, assets, or education for their eldest children that they *would like to* achieve in their lifetime. We complement this measure with expectations — what level of these three outcomes individuals *think* they will reach in ten years. We summarise aspirations and expectations using an Anderson (2008) index of the aspired/expected level of income, wealth, and education for their children. The expectations index and the aspirations index have a correlation coefficient of 0.42 at baseline. We further aggregate all components of these two indices into a single aspirations and expectations aggregate index.

Figure I displays effects on our indices of aspirations and expectations, across survey rounds after our intervention. Appendix Table A.12 reports results across all three indices and their individual components. Strikingly, after five years we see positive and strongly significant effects on both the aspirations and expectations indices, relative to both the treatment and the placebo group (third column). These are driven by increases in all dimensions of aspirations and expectations. The size of the effect on the indices is modest but consistently between 0.1 and 0.2 standard deviations relative to the placebo across measures. These results provide support for the intervention increasing aspirations and expectations, and that these effects are not driven by a screening effect.

In the first column of the same graph, we report the same effects on the same indicators, collected the same day after the screening of the videos. We find small and significant effects on the aspirations index of 0.1 standard deviations relative to the placebo group but no significant effect relative to the control group. We see similar patterns on the other measures. We are able to exclude a screening effect immediately after the screening,

given the significant difference between treatment and placebo groups. Two qualifications should be noted about the post-screening results: first, the interviews immediately after the screening for the control group were conducted at the respondents' homes, rather than the screening site. However, this does not affect the treatment versus placebo comparison. Second, 81 individuals left before they were surveyed and 22 individuals missed the screening altogether, resulting in a smaller sample size for the post-screening survey relative to the six-month follow-up.

In the second column, we report effects six months after the intervention. We find small and not statistically significant effects across all indices of about 0.1 standard deviations relative to the control group. We only detect differences between the treatment and placebo group six months after the screening in the expectations index, significant at the 10 per cent level. Most of the effect is driven by aspired and expected children's education (Appendix Table A.12). The effects on these variables may evolve over time, as individuals become more confident of the results of their own investments.

Although expectations and aspirations are strongly correlated at baseline, their response to the intervention over time differs, as expectations respond more to treatment after six-months, with aspirations increasing over a longer time-horizon. A plausible interpretation consistent with this dynamic pattern is that expectations contribute to the formation of aspirations. Our results suggest that aspirations may adjust gradually as individuals start investing more and seeing the returns of their investments and also as they start expecting to do better in the future.

Overall, the five year results are consistently stronger in terms of statistical significance, but it is striking that the impact was visible in the data almost immediately after the screening, both in terms of the pattern and size of the effects, and mostly significantly so. It means, at least, that we cannot reject our hypothesis that aspirations were lifted through the intervention, leading to future oriented behaviour through effort and investment, and with aspirations being still high afterwards. All results are qualitatively similar when using an alternative measure of aspirations and expectations, the "aspirations (or expectations) gap": the level on a dimension a participant would like (or think) to attain minus the level they reported to have reached at baseline (Appendix Table A.13).

5.2.3 Alternative mechanisms

While our results appear consistent with the predictions regarding the role of aspirations, we also test our findings against several alternative mechanisms that might be affected by our intervention and might have yielded similar changes in economic behaviour. We consider three plausible alternative mechanisms, conceptually and empirically, as well as social desirability bias caused by the study itself. Appendix Section B.4 details the construction of the measures used to test these alternative mechanisms.

Table 5: Economic changes after six months

After six months	(1) Treatment	(2) Placebo	(3) Treat. vs. placebo	(4) Control mean (SD) Total obs.
Labour effort:				
Daily minutes on family farm	17.78	-24.45	42.23**	683.89
	(22.66)	(21.99)	(21.34)	(344.53)
	[0.81]	[0.53]	[0.10]*	1126
Daily minutes in leisure	13.11	-22.14	35.25	1982.38
	(55.02)	(55.61)	(55.21)	(834.23)
	[0.81]	[0.69]	[0.52]	1125
Savings and credit:				
Total savings (USD PPP)	21.66**	1.19	20.47*	28.63
	(10.45)	(7.98)	(11.28)	(103.13)
	[0.08]*	[0.88]	[0.28]	1121
Credit amount (USD PPP)	4.79	2.81	1.98	19.38
	(3.27)	(3.31)	(3.50)	(42.80)
	[0.19]	[0.53]	[0.76]	1130
Hypothetical loan (1 year, USD PPP)	24.03	237.19	-213.16	2381.84
	(234.28)	(243.18)	(256.28)	(3210.14)
	[0.92]	[0.53]	[0.76]	1137
Hypothetical loan (10 years, USD PPP)	3616.50**	3584.63**	31.87	9452.19
	(1747.55)	(1601.78)	(2065.16)	(15581.81)
	[0.08]*	[0.10]	[0.99]	1142
Cohort 1: Children of post-primary school-going age at the five-year follow-up				
Children aged 11-15 in school	0.09*	0.04	0.05	0.56
	(0.05)	(0.05)	(0.05)	(0.73)
	[0.10]	[0.49]	[0.51]	1126
Daily minutes in school for children aged 11-15	21.81	10.96	10.85	188.71
	(16.51)	(15.76)	(16.36)	(248.36)
	[0.19]	[0.49]	[0.51]	1118
Daily minutes studying for children aged 11-15	11.09*	5.06	6.03	58.11
	(6.01)	(5.96)	(6.26)	(86.58)
	[0.10]	[0.49]	[0.51]	1117
Cohort $2^{(a)}$: Children of primary school-going age at the five-year follow-up				
Children aged 7-10 in school	0.08	-0.01	0.09*	0.60
	(0.05)	(0.05)	(0.05)	(0.73)
	[0.34]	[0.85]	[0.19]	1126
Daily minutes in school for children aged 7-10	14.78	-6.40	21.17	198.10
	(16.28)	(16.38)	(16.29)	(250.25)
	[0.55]	[0.85]	[0.19]	1117
Daily minutes studying for children aged 7-10	-1.61	-8.03*	6.42	45.08
	(4.86)	(4.56)	(4.69)	(70.78)
	[0.74]	[0.24]	[0.19]	1119
For all children				
Schooling expenditure (USD PPP)	9.01**	4.84	4.17	37.75
	(3.68)	(3.83)	(4.10)	(51.39)
	[0.01]**	[0.21]	[0.31]	1118

Notes: OLS estimates of within-village treatment and placebo effects six months after the intervention (columns 1-2). Column 3 tests for differences in parameters obtained in first two columns. The comparison group comprises households from the 64 treated villages that were not invited to any screening. Column 4 displays the control mean, standard deviation, and total number of observations. All columns control for village fixed effects and characteristics of the household head: age, years of education, an indicator for being single, and an indicator for being male. All regressions additionally control for the number of children aged 0-15 at baseline to account for the baseline imbalance in the number of children. Heteroskedasticity-robust standard errors are in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The conversion is described in Appendix B.1. The unit of observation is the household. The number of observations varies slightly across rows because some respondents do not answer all questions. Daily minutes of an activity are the sum of schooling-age household members' daily minutes. Daily minutes of an activity are the sum of adult household members' daily minutes. Cohort $2^{(a)}$ is not directly comparable to cohort 2 in Table 3, because some of the children in cohort 2 were not of primary school-going age at the time of six-months follow-up and as they would have been between 2 and 6 years old; we did not collect data for children in this age range.

Time and risk preferences — The intervention could have increased the discount factor (β in our framework) leading to the observed increase in future-oriented behaviour, as the future is more valued.[13]

We do not find that time preferences have shifted (top panel, Table 6). After the intervention, there is no change in the share of patient, impatient, very impatient respondents, or present-biased respondents — categorised as in Ashraf, Karlan, and Yin (2006) — indicating no impact on the discount factor. After five years, there is a small negative treatment effect on the share of time-inconsistent respondents that are patient now and impatient later, but this effect does not occur six months after intervention nor survive multiple testing.[14]

While risk does not enter explicitly the theoretical framework, it could be trivially extended. For example, if future returns are risky, then higher risk aversion would induce less effort and investment in the future. The increased salience of a plausible future through the intervention may have reduced risk aversion, leading to the observed effects.

Risk preferences have not shifted either (top panel, Table 6). In fact, if anything, our risk aversion measure — adapted from Binswanger (1980) — has increased relative to the placebo group after six months, although the effect is not robust to multiple test correction and does not persist after five years.

Perceived returns to own effort and causes of success — The intervention may have altered people's beliefs about about their ability to change their own outcomes, or, in economic terms, the underlying beliefs about the return to their own effort.

Six months after the screening, we find the treatment increased internal locus of control and the extent to which people believe poverty is an issue of individual agency (second panel, Table 6). For locus of control, this effect is also present relative to the placebo group and robust to testing for multiple hypotheses. Conversely, we find no evidence of effects on our measure of grit after six months or five years. We also find that treatment decreased the extent to which people believe poverty is caused by fate, though this effect is only significant and not robust to multiple hypothesis testing (third panel, Table 6). However, none of the measures in these two panels have persistent effects after five years.

We cannot rule out that some part of the mechanism behind the effects of the intervention is driven by changes in locus of control. Indeed, such variables are often correlated with aspirations (Levenson, 1974; Locke and Latham, 2002). At baseline we observed a

13. Alternatively, Gabaix and Laibson (2017) theorise that improving the extent to which households can visualise the future may lead to more patient behaviour.

14. Among adults, John and Orkin (2021) observed no impact on time preferences with a light-touch visualisation-based intervention, while Blattman, Jamison, and Sheridan (2017) reported temporary effects on patience through an intensive therapy program, which did not persist. In contrast, Alan and Ertac (2018) found impacts on patience among children three years after an intensive intervention that aimed to foster forward-looking behaviour.

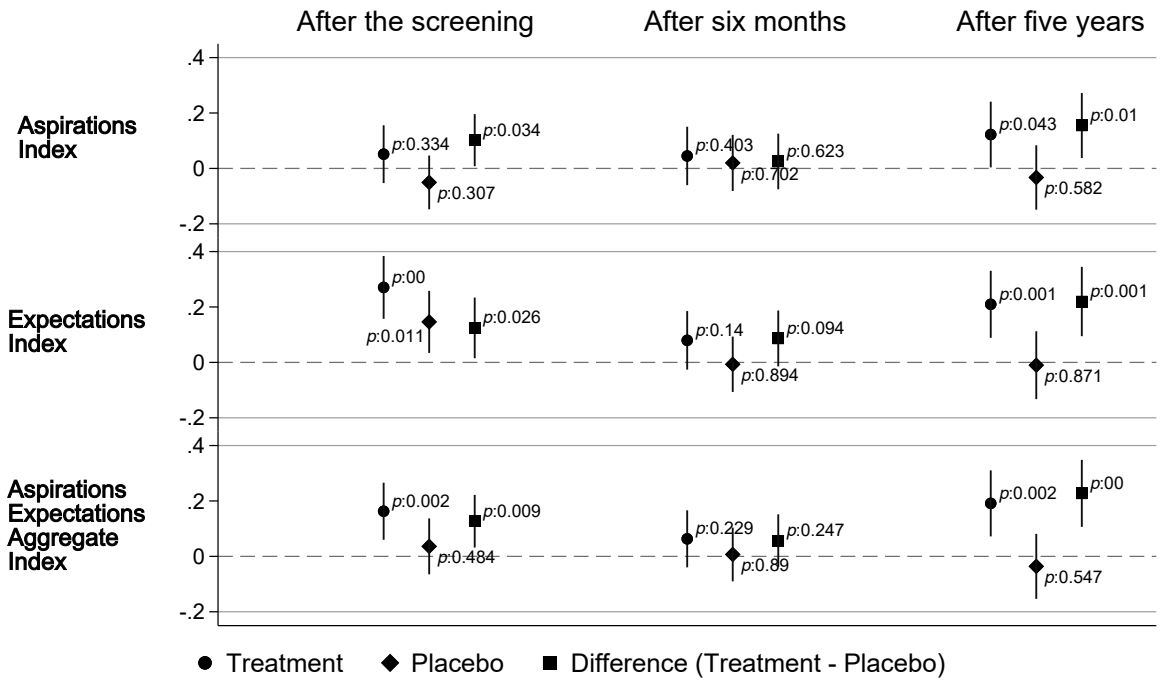


Figure I: Treatment effects on the aspirations and expectations indices.
Notes: Treatment and placebo intention-to-treat effects on aspiration and expectations indices across survey rounds. The first column shows effects on three indices collected right after the screening of the videos took place (or a few days after the baseline for the control group). The second column shows effects six months after the screening of the videos took place. The third column shows the effects five years after the screening of the videos. The comparison group comprises households from the 64 treated villages that were not invited to any screening. Column 4 displays the control mean, standard deviation, and total number of observations. All columns control for village fixed effects and characteristics of the respondent: age, years of education, an indicator for being single, and an indicator for being male. Heteroskedasticity-robust standard errors are clustered at the household level. Bars correspond to 95 per cent confidence intervals. Square-shaped markers report the estimated difference between the treatment and placebo effects. p-values are reported next to the markers. The aspirations index is an Anderson (2008) index combining what individuals would like to achieve in their life in terms of reported income, wealth, and years of education for their eldest child. The expectations index similarly combines what individuals think they will achieve in ten years time in terms of the same three dimensions (income, wealth, and children's education). The aspirations and expectations aggregate Anderson (2008) index combines six dimensions of reported income, wealth and years of education for their eldest child, for aspirations and expectations. Appendix Table A.12 reports results across all three indices and their individual components.

Table 6: Testing mechanisms

	After six months				After five years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Treatment	Placebo	Treat. vs. placebo	Control mean (SD) Total obs.	Treatment	Placebo	Treat. vs. placebo	Control mean (SD) Total obs.
Risk and time preferences:								
% that is patient	-0.01	-0.02	0.01	0.30	-0.01	-0.01	-0.01	0.17
	(0.03)	(0.02)	(0.02)	(0.46)	(0.02)	(0.02)	(0.02)	(0.37)
	[0.85]	[0.63]	[0.80]	2078	[0.62]	[0.88]	[0.72]	1955
% that is somewhat impatient	-0.02	-0.02	0.01	0.15	-0.02	-0.00	-0.01	0.10
	(0.02)	(0.02)	(0.02)	(0.36)	(0.02)	(0.02)	(0.02)	(0.31)
	[0.85]	[0.37]	[0.80]	2078	[0.55]	[0.88]	[0.58]	1955
% that is most impatient	0.03	0.04	-0.01	0.55	0.03	0.01	0.02	0.73
	(0.03)	(0.03)	(0.03)	(0.50)	(0.03)	(0.02)	(0.03)	(0.45)
	[0.85]	[0.36]	[0.80]	2078	[0.55]	[0.88]	[0.58]	1955
% that is present biased	0.03	0.04	-0.01	0.34	0.03	0.05*	-0.02	0.53
	(0.03)	(0.03)	(0.03)	(0.47)	(0.03)	(0.03)	(0.03)	(0.50)
	[0.85]	[0.36]	[0.80]	2053	[0.55]	[0.49]	[0.58]	1955
% that is patient now and impatient later	-0.00	0.00	-0.01	0.22	-0.05**	-0.02	-0.03	0.19
	(0.02)	(0.02)	(0.02)	(0.41)	(0.02)	(0.02)	(0.02)	(0.39)
	[0.92]	[0.84]	[0.80]	2053	[0.10]	[0.85]	[0.58]	1955
Risk aversion:	0.04	-0.11	0.15*	3.20	0.02	-0.04	0.06	2.52
most to least risk averse (1 to 5)	(0.08)	(0.08)	(0.08)	(1.51)	(0.09)	(0.09)	(0.09)	(1.54)
	[0.85]	[0.36]	[0.39]	2076	[0.79]	[0.88]	[0.58]	1955
Perceived returns of own effort:								
Internal locus of control	0.23*	-0.07	0.30**	12.96	-0.00	0.05	-0.05	12.26
	(0.12)	(0.12)	(0.12)	(2.10)	(0.11)	(0.11)	(0.11)	(1.91)
	[0.09]*	[0.58]	[0.04]**	2078	[1.00]	[0.94]	[0.94]	1956
Individual causes of poverty	0.26*	0.20	0.06	9.20	0.02	0.00	0.02	9.15
	(0.14)	(0.14)	(0.14)	(2.40)	(0.14)	(0.13)	(0.13)	(2.04)
	[0.09]*	[0.42]	[0.66]	2077	[1.00]	[0.99]	[0.94]	1956
Grit index	0.03	-0.06	0.10	0.00	-0.06	-0.06	0.00	0.02
	(0.06)	(0.06)	(0.06)	(1.00)	(0.06)	(0.06)	(0.06)	(0.99)
	[0.56]	[0.42]	[0.17]	2079	[1.00]	[0.94]	[0.94]	1956
Perceived external causes of success:								
Chance locus of control	0.01	-0.02	0.03	13.35	-0.01	-0.05	0.04	12.67
	(0.16)	(0.16)	(0.17)	(2.70)	(0.15)	(0.15)	(0.15)	(2.36)
	[0.94]	[0.95]	[0.87]	2075	[0.94]	[0.98]	[0.79]	1956
Fate causes of poverty	-0.23*	-0.01	-0.22*	7.19	-0.03	0.00	-0.03	6.69
	(0.14)	(0.13)	(0.14)	(2.31)	(0.11)	(0.10)	(0.10)	(1.80)
	[0.18]	[0.95]	[0.20]	2077	[0.94]	[0.98]	[0.79]	1956
Awareness and perceived returns of agricultural technologies:								
Information index					-0.03	-0.06	0.03	-0.00
					(0.07)	(0.07)	(0.07)	(1.00)
					[0.67]	[0.39]	[1.00]	1104
Expected fertiliser yields index					0.10	0.10	-0.00	0.00
					(0.07)	(0.07)	(0.07)	(1.00)
					[0.35]	[0.33]	[1.00]	1085

Notes: OLS estimates of within-village treatment and placebo effects after six months (columns 1-2) and after five years (columns 5-6) of the intervention. Columns 3 and 7 test for differences in parameters obtained in previous two columns. The comparison group comprises households from the 64 treated villages that were not invited to any screening. Column 4 and 8 display the control mean, standard deviation, and total number of observations. Heteroskedasticity-robust standard errors are clustered at the household-level in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. The unit of observation is the individual respondent (household head or their spouse), except for information and fertiliser beliefs indices (which are at the household-level and were only measured after five years of the intervention). The number of observations varies slightly across rows because some respondents do not answer all questions, though indices aggregate all non-missing outcomes. The locus of control variables are based on the Internal, Powerful Others, and Chance Scale (Levenson, 1974) and capture if people see outcomes as contingent on their behaviour (internal locus of control), as a result of chance, luck or fate (chance locus of control). We use survey-based instruments to calculate risk (Binswanger, 1980) and time preferences (Ashraf, Karlan, and Yin, 2006) and provide details of the measures in the appendix. Grit includes answers to two survey questions about how the respondent would characterise themselves. The additional measures presented in the lowest panel are standardized Anderson (2008) indices built on survey instruments and are described in detail in the Appendix B. The information index is constructed from indicator variable that take value one if the household head reported having performed some of the behaviours that were described in the documentaries. The expected fertiliser yields index is based on the household heads' expected increase in output from the application of different quantities of fertiliser for maize and sorghum in an hypothetical good or bad season.

positive correlation coefficient of 0.17 between our measure of internal locus of control and the aspirations and expectations aggregate index. For instance, the intervention's six-months impact on economic behaviour could be due to changes in both these traits and aspirations. However, we can conclude that locus of control is a less probable mechanism for sustained economic effects, as its changes do not persist.

Information and expected returns to innovation — In the bottom panel of Table 6, we investigate two additional mechanisms not included in the model but empirically examined. Firstly, we assess if treated households adopted activities mentioned in the videos, such as purchasing pumps or using specific technologies. However, there is no effects on a summary index of a set of fourteen pre-specified activities, nor on households' engagement in each activity. Secondly, we explore whether exposure to the documentaries influenced households' beliefs about the returns to modern agricultural technologies, like fertilisers, despite these not directly featuring in the videos. Yet, there is no effect observed on households' beliefs about the returns to fertilisers. Both findings suggest that the households might have been aware of these activities even without the videos, as they are common within their villages.

Social desirability bias — Finally, one might worry that experimenter demand effects or self-reporting issues might bias our results: treated households might have reported better outcomes in order to please the study team. However, while we cannot fully rule them out, such effects seem unlikely to persist over a five-year horizon. Moreover, we do not find effects in outcomes that households may have directly linked to the goals of the study. Not having found increased adoption in the practices shown in the videos partly allays this concern.

In sum, we cannot prove that the behavioural change is caused by the aspirations shift, but we definitely cannot reject it as the most plausible explanation: we find little evidence in favour of these alternative mechanisms, although we cannot rule out another psychological mechanism we do not observe.

5.3 Discussion of main results

Our intervention showed documentaries to the treated households in which people that have been successful despite initially living in poverty tell their story. In line with social learning theory in psychology, this has exposed the treated group to individuals whose initial life conditions they could relate to and whose success they could see as reachable.

The patterns of results in Section 5 are consistent with households being induced to aspire to and emulate what better off households in their communities do, even though they had lived with them well before the intervention. Treated households engage more in the kind of activities and investments the top tercile in Table 1 were doing at baseline,

such as investing in livestock and working more on the farm, rather than specifically doing what was portrayed in the videos. We also see increases in effort, and investment into their children's education as well as the locally common productive activities, crops and livestock, that persist after five years from the experiment.

Finding both more investment and likely higher perceived and actual standard of living outcomes helps to counter concerns about some of the possible alternative negative consequences of boosting aspirations and aspirations gaps: it does not appear that the intervention gave "false hope" (encouraging households to take decisions that make them worse off) or made them "frustrated" as aspirations have been raised too high, leading to less investment and effort, a possibility highlighted in Ray (2006), Genicot and Ray (2017), and McKenzie, Mohpal, and Yang (2022).

Overall, the results support our predictions: the intervention led to positive changes in aspirations and economic outcomes. We finally combine all the summary indices (agricultural investment, educational investment, welfare, and the aspirations and expectations aggregate index) into a single omnibus index, which finds the intervention yields a positive and significant effect on outcomes relative to the placebo and control groups, five years after exposure to the role models in the videos (Table 7).

5.4 Lack of heterogeneous effects

We find little or no heterogeneity of our treatment effects on our summary indices across the baseline measures we had pre-specified (Appendix Figure A.3).[15] The effects on the agricultural investment do not vary by our pre-specified dimensions after accounting for multiple hypothesis testing. We find that the educational investment index shifts more for those who had higher baseline expectations, but not for any other dimension. Despite the lack of average effects, we find some evidence that our welfare index shifts for treated individuals who had above median aspirations and assets (including livestock and productive assets) at baseline. These patterns suggest that our intervention may have had more benefit for households with higher goals for their future at baseline.

Finally, we find no heterogenous effects on the aspirations and expectations indices. In particular, while spouses may have differing baseline goals, treatment changes for both spouses' aspirations and expectations are not significantly different from each other.[16]

15. In an exploratory analysis, we also test whether effects vary by the terciles of durable assets used to categorise our sample in Table 1. We find that effects on the aspirations and expectations aggregate index are larger among households in the middle tercile of durable assets relative to those in the bottom tercile (Appendix Table A.14).

16. We observe aspirations and expectations for both the primary man and woman in the household. Whereas for the economic indices, the heterogeneity by sex of the respondent refers to whether the household head was a woman.

Table 7: Summary indices in within-village analysis

After five years	(1) Treatment	(2) Placebo	(3) Treat. vs. placebo	(4) Control mean (SD) Total obs.
Agricultural investment index	0.18***	0.03	0.14**	-0.00
	(0.07)	(0.07)	(0.06)	(1.00)
	[0.01]***	[0.94]	[0.03]**	1090
Educational investment index	0.21***	-0.02	0.23***	0.00
	(0.07)	(0.06)	(0.07)	(1.00)
	[0.00]***	[0.94]	[0.00]***	1082
Welfare index	0.07	-0.02	0.09	0.00
	(0.07)	(0.07)	(0.08)	(1.00)
	[0.38]	[0.94]	[0.25]	1092
Aspiration index	0.12**	-0.03	0.15***	0.02
	(0.06)	(0.06)	(0.06)	(1.00)
	[0.05]*	[0.94]	[0.01]**	1956
Expectations index	0.21***	-0.01	0.22***	0.01
	(0.06)	(0.06)	(0.06)	(1.00)
	[0.00]***	[0.94]	[0.00]***	1955
Aspirations and expectations aggregate index	0.19***	-0.04	0.23***	0.01
	(0.06)	(0.06)	(0.06)	(1.00)
	[0.00]***	[0.94]	[0.00]***	1956
Omnibus index	0.27***	-0.00	0.28***	0.00
	(0.07)	(0.07)	(0.07)	(1.00)
	[0.00]***	[0.94]	[0.00]***	1093

Notes: OLS estimates of within-village treatment and placebo effects five years after the intervention (columns 1-2). Column 3 tests for differences in parameters obtained in first two columns. The comparison group comprises households from the 64 treated villages that were not invited to any screening. Column 4 displays the control mean, standard deviation, and total number of observations. All columns control for village fixed effects and characteristics of the household head: age, years of education, an indicator for being single, and an indicator for being male. Regressions on the educational investment index additional control for the number of children aged 0-15 at baseline to account for the baseline imbalance in the number of children. Heteroskedasticity-robust standard errors are in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. The unit of observation is the household, except for the aspirations and expectations indices (which are are observed for both household head and their spouse). The number of observations varies slightly across rows because some respondents do not answer all questions, though the indices aggregate all non-missing outcomes. The outcomes are inverse-covariance-weighted averages standardised relative to the within-village control group, following Anderson (2008). The agricultural investment index includes all outcomes reported in Table 2, with daily minutes in leisure being recoded to be negative. The educational investment index includes all outcomes reported in Table 3. The welfare index includes all outcomes reported in Table 4, with months of food insecurity in the last year and consumption of sin goods recoded to be negative. The welfare index averages over the household head's subjective well-being outcomes. The aspirations and expectations aggregate index is made of the reported income, wealth and years of education for children, for aspirations and expectations. The omnibus index aggregates the four standardised indices into a single index, following Bessone et al. (2021) and Kling, Liebman, and Katz (2007). As the omnibus index is for the whole household, we use the household head's aspirations and expectations aggregate index.

6 Spillovers

In this section, we test for potential spillover effects between treated and other households. The findings in this section do not alter our interpretation of the main treatment effects. To test for spillovers, we included ten additional control villages to the main sample as a comparison group. These villages were not visited during the baseline or six-month follow-up surveys. During our five-year follow-up, we surveyed 18 randomly selected households per pure control village. Our allocation of villages between our main sample of treatment villages and the pure control villages proceeded in three steps. First, we randomly selected 84 villages from the census list of villages in Doba district. Second, we identified 16 screening venues near these villages, such as classrooms or agricultural facilities, capable of accommodating at least fifty individuals with controlled access. Third, out of these 84, we selected for treatment the 64 villages closest in distance to these venues, with a maximum of four villages per screening site. The remaining ten closest villages to these venues constitute the pure control group: they would have been part of the experiment if we had allocated quintuplets or sextets of villages per screening site.[17]

The validity of our approach rests on the comparability of the pure control villages to the treatment villages, in the absence of the intervention. In Appendix Table A.3 we report balance tests between treatment and pure control villages for a set of village-level characteristics. Overall, this table suggest that balance cannot be rejected.[18] For sake of parsimony, we focus our spillover analysis on the summary indices that we had also used to summarises our main results, standardised relative to the pure-control group. Given our limited sample size, focusing on summary indices can increase power to identify potential spillover effects.

We employ three empirical strategies to test for spillovers. The first strategy compares households in treated villages with households in pure control villages, estimating the following empirical specification:

$$y_{i3} = \alpha + \delta^s T_i + \rho^s P_i + \varphi C_i + X'_{i2}\pi^s + \varepsilon_i \tag{4}$$

where X_{i2} includes the same set of pre-specified controls, and village-level controls and screening fixed effects to replace the village-level fixed effects. C_i is an indicator being equal to one for households in treatment villages that were not invited to watch the documentary nor the TV show. For these specifications, standard errors are clustered at the village-level at which treatment is now allocated. The superscript "s" is added

17. See Appendix Figure A.4 for the location of the pure control and treatment villages. Given the location of the screening sites, the remaining ten out of 84 villages were too remote to be considered as part of either treatment or control villages and were not further considered in the analysis.

18. For the sources of these variables see Appendix Table A.6.

Table 8: Summary indices in spillover analysis

After five years	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Treatment	Placebo	Control	Treat. vs. placebo	Treat. vs. control	Placebo. vs. control	Pure Control mean (SD) Total obs.
Agricultural investment index	0.09	-0.07	-0.10	0.16**	0.19***	0.03	-0.00
	(0.11)	(0.10)	(0.10)	(0.07)	(0.05)	(0.06)	(1.00)
	[0.41]	[0.80]	[0.57]	[0.02]**	[0.00]***	[1.00]	1223
Educational investment index	0.18**	-0.04	-0.02	0.22***	0.20***	-0.02	0.00
	(0.07)	(0.07)	(0.06)	(0.07)	(0.06)	(0.06)	(1.00)
	[0.02]**	[0.80]	[0.76]	[0.00]***	[0.00]***	[1.00]	1219
Welfare index	0.21**	0.14	0.14	0.07	0.07	-0.00	0.00
	(0.09)	(0.09)	(0.09)	(0.08)	(0.06)	(0.08)	(1.00)
	[0.02]**	[0.80]	[0.57]	[0.40]	[0.30]	[1.00]	1224
Aspiration index	0.22***	0.06	0.09	0.17***	0.13**	-0.04	0.00
	(0.08)	(0.07)	(0.07)	(0.05)	(0.06)	(0.06)	(1.00)
	[0.01]***	[0.80]	[0.57]	[0.00]***	[0.05]*	[1.00]	2231
Expectations index	0.26***	0.03	0.06	0.24***	0.21***	-0.03	-0.00
	(0.08)	(0.08)	(0.08)	(0.06)	(0.07)	(0.06)	(1.00)
	[0.01]***	[0.87]	[0.57]	[0.00]***	[0.01]***	[1.00]	2230
Asp. and exp. aggregate index	0.24***	0.01	0.06	0.23***	0.18***	-0.05	-0.00
	(0.08)	(0.07)	(0.07)	(0.05)	(0.06)	(0.06)	(1.00)
	[0.01]***	[0.89]	[0.57]	[0.00]***	[0.01]***	[1.00]	2231
Omnibus index	0.34***	0.06	0.07	0.27***	0.26***	-0.01	0.00
	(0.09)	(0.07)	(0.08)	(0.07)	(0.06)	(0.06)	(1.00)
	[0.00]***	[0.80]	[0.57]	[0.00]***	[0.00]***	[1.00]	1225

Notes: OLS estimates of between-village effects five years after the intervention (columns 1, 2 and 3). Column 4 tests for differences in parameters obtained in first two columns. Column 5 tests for differences in parameters obtained in first and third columns. Column 6 tests for differences in parameters obtained in second and third columns. The comparison group comprises households from the ten pure-control villages that were first surveyed five years after the intervention. Column 7 displays the mean, standard deviation for the pure-control group, and total number of observations. All regressions control for screening-site fixed effects, individual characteristics of the respondent (age, years of education, an indicator for being single, and an indicator for being male) and village-level controls (the number of inhabitants, hectares covered by forest, an indicator for whether sorghum is the main crop, costs of trip to nearest market, an indicator for whether the village has a first cycle school, percentage of households with radio, distance to the next market place, distance to the school, distance to the next farmers training centre, distance to the next health centre, distance to the next river). Regressions on the educational investment index additional control for the number of children aged 0-15 currently in the household to account for the baseline imbalance in the number of children. Heteroskedasticity-robust standard errors are clustered at the village-level and are in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. The unit of observation is the household, except for the aspirations and expectations indices (which are are observed for both household head and their spouse). The number of observations varies slightly across rows because some respondents do not answer all questions, though the indices aggregate all non-missing outcomes. The outcomes are inverse-covariance-weighted averages standardised relative to the pure-control group, following Anderson (2008). The agricultural investment index includes all outcomes reported in Table 2, with daily minutes in leisure being recoded to be negative. The educational investment index includes all outcomes reported in Table 3. The welfare index includes all outcomes reported in Table 4, with months of food insecurity in the last year and consumption of sin goods recoded to be negative. The welfare index averages over the household head's subjective well-being outcomes. The aspirations and expectations aggregate index is made of the reported income, wealth and years of education for children, for aspirations and expectations. The omnibus index aggregates the agricultural investment, educational investment, welfare, and aspirations and expectations aggregate standardised indices into a single index, following Bessone et al. (2021) and Kling, Liebman, and Katz (2007). As the omnibus index is for the whole household, we use the household head's aspirations and expectations aggregate index.

to parameters δ and ρ to distinguish them from the previous within-treatment village estimates. We can assess whether spillover effects may bias the previous results by estimating φ, which measures the extent to which control households within treatment villages were (indirectly) affected by the treatment. Testing whether φ is different from zero is a statistical test of the existence of positive or negative spillovers. We find no systematic evidence of spillovers: the estimated difference between the within-village control and the pure control group is never statistically significant (Column 3, Table 8). The estimated direct treatment effects relative to the pure control group (Column 1) are all within the 90 per cent confidence interval of our within-village estimates reported in Section 5. Our within-village comparisons of treated household relative to placebo and within-village controls remain very close to those reported throughout Section 5 (Column 4 and 5).

In our second empirical specification to identify spillovers, we expand the model in Equation (4) by exploiting our randomisation saturation design. This design allows us to test how individuals in intensely-treated villages compared five years after the screening to individual in pure control villages, and to individuals in villages with fewer individuals exposed to the documentary and more individuals exposed to the placebo intervention. We interact our treatment, placebo and control indicators with an indicator for being in a village where 18 more households were exposed to the documentaries (treatment-intense) and an indicator for being in a village with 18 more households exposed to the placebo videos (placebo-intense). The comparison group remains the pure control villages.[19] We do not find systematic evidence of spillover effects within treated villages relative to the pure control group, even in villages were more individuals had been exposed to the role model documentaries (Appendix Table A.15). For example, control households in "treatment-intense" villages are not different to those in pure control villages (Column 3). Neither treated, control, or placebo group households respond differently in intensely treated villages relative to those in villages with more households exposed to placebo videos. We find no statistically significant differences between treated individuals in "treatment-intense" villages and those in "placebo-intense" villages (Column 7). Control individuals in "treatment-intense" villages and those in "placebo-intense" villages also show no significant differences (Column 8). Similarly, surveyed individuals in the placebo group are not different across the two groups of treated villages (Column 9).

In our third empirical specification to identify spillovers, we relax the assumption that spillovers may only occur within the same village. We do so by testing whether the number of individuals within a 1km radius, conditional on the number of study villages within the same radius, significantly changes our outcomes of interest beyond the household-treatment assignment. Our specification is similar to Miguel and Kremer (2004).[20] We

19. Appendix Section C.3 provides the equation of the estimated specification.

20. In Appendix Section C.3 we provide details of how we empirically selected the length of the radius to be 1km, following the methodology of Egger et al. (2022).

include two additional terms to Equation (4): the number of treated households within a 1km of household i (excluding household i) and the number of villages within 1km (excluding household i's village). The number of treated households that are within 1km is plausibly exogenous, conditioning on the number of villages within 1km. For this specification, we account for spatial dependence using Conley standard errors with a uniform kernel up to 1km (Conley, 1999). Appendix Figure A.4 graphically illustrates the spatial variation in treatment intensity in our sample.[21] We find no evidence of spillovers across villages (Column 4, Appendix Table A.16). The additional effects of treated households within a 1km radius are not statistically significant across all indices. The signs of the coefficients do not have a consistent pattern and the estimates are very small across all our indices.

Overall, we find no evidence of spillover effects. Nevertheless, given the small sample of villages, the results in this section are more indicative than conclusive, since we are not perfectly powered to detect spillover effects. We conclude that, at least in our setting, there is either no evidence of spillovers within communities or that such effects, if present, take longer than five years to occur or a larger number of villages to detect them.

7 Conclusion

We randomly exposed individuals in a poor and isolated area to a one-hour video documentary in which four people from similar backgrounds to the audience tell their life story of getting out of poverty. After five years, we find persistent effects on whether households invest for the future, and some indicators of their standard of living. These results are meaningful. The size of the effects are not very large — a few dollars more spending on education, some more durable assets. But we still find effects five years after an intervention lasting an hour, simply showing inspirational stories about the lives of people similar to those watching. Something is triggered that affects forward-oriented behaviour. We find evidence consistent with a change in aspirations being the main psychological mechanism.

Our research has shown that a light-touch easily scalable population-wide behavioural intervention can have persistent economic impacts after five years on people living in poor settings. Is this intervention giving false hope? We cannot fully judge this but the persistence after five years of impacts on assets, with a at least as good or better standard of living than the counterfactual, suggests not. And we did not suggest to individuals — rightly or wrongly — what path would lead them out of poverty, unlike most interventions that offer 'solutions' in microcredit, health or education. We only invited our treatment group to listen to stories told by individuals from similar backgrounds.

21. Within a 1km radius, the median number of treated households is 29, whereas more than two thirds of the villages have another village within a 1km radius.

References

Abeler, J., A. Falk, L. Goette, and D. Huffman. 2011. "Reference Points and Effort Provision." *American Economic Review* 101 (2): 470–92.

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 (3): 1121–1162.

Alan, S. and S. Ertac. 2018. "Fostering Patience in the Classroom: Results from a Randomised Educational Intervention." *Journal of Political Economy* 126 (5): 1865–1911.

Anderson, M. L. 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 (484): 1481–1495.

Angelucci, M. and D. Bennett. 2021. "The Economic Impact of Depression Treatment in India." *Institute for the Study of Labour (IZA) Discussion Paper* 14393: 1–64.

Appadurai, A. 2004. "The Capacity to Aspire: Culture and the Terms of Recognition." In *Culture and Public Action*, edited by Vijayendra Rao and M. Walton. Palo Alto, California: Stanford University Press, 59–84.

———. 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*: 1–80.

Ashraf, N., D. S. Karlan, and W. Yin. 2006. "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." *Quarterly Journal of Economics* 121 (2): 635–672.

Baird, S., J. A. Bohren, C. McIntosh, and B. Özler. 2018. "Optimal Design of Experiments in the Presence of Interference." *Review of Economics and Statistics* 100 (5): 844–860.

Bandura, A. 1977a. "Self-Efficacy: Toward a Unifying Theory of Behavioural Change." *Psychological Review* 84 (2): 191–215.

———. 1977b. *Social Learning Theory*. Englewood Cliffs, New Jersey: Prentice-Hall.

Baranov, V., S. Bhalotra, P. Biroli, and J. Maselko. 2020. "Maternal Depression, Women's Empowerment, and Parental Investment: Evidence from a Randomized Controlled Trial." *American Economic Review* 110 (3): 824–59.

Barker, N., G. Bryan, D. Karlan, A. Ofori-Atta, and C. Udry. 2022. "Cognitive Behavioral Therapy among Ghana's Rural Poor Is Effective Regardless of Baseline Mental Distress." *American Economic Review: Insights* 4 (4): 527–45.

Batista, C. and J. Seither. 2019. "Aspirations, Expectations, Identities: Behavioral Constraints of Micro-entrepreneurs." *NOVAFRICA Working Paper Series* 1906: 1–28.

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 (6068): 582–586.

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 (1): 3–18.

Benabou, R. 1996. "Equity and Efficiency in Human Capital Investment: The Local Connection." *The Review of Economic Studies* 63 (2): 237–264.

Benjamini, Y., A. M. Krieger, and D. Yekutieli. 2006. "Adaptive Linear Step-up Procedures that Control the False Discovery Rate." *Biometrika* 93 (3): 491–507.

Bernard, T., S. Dercon, K. Orkin, and A. S. Taffesse. 2014. "The Future in Mind: Aspirations and Forward-Looking Behaviour in Rural Ethiopia." *Bureau for Research and Economic Analysis of Development Working Paper* 429: 1–42.

Bernard, T. and A. S. Taffesse. 2014. "Aspirations: An Approach to Measurement with Validation Using Ethiopian Data." *Journal of African Economies* 23 (2): 189–224.

Besley, T. 2016. "Aspirations and the Political Economy of Inequality." *Oxford Economic Papers* 69 (1): 1–35.

Bessone, P., G. Rao, F. Schilbach, H. Schofield, and M. Toma. 2021. "The Economic Consequences of Increasing Sleep Among the Urban Poor." *The Quarterly Journal of Economics* 136 (3): 1887–1941.

Bhan, P. C. 2020. "Do Role Models Increase Student Hope and Effort? Evidence from India." *University of Glasgow Working Paper* : 1–67.

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: 1–76.

Bickel, G. W., M. Nord, C. Price, W. Hamilton, and J. Cook. 2000. *Guide to Measuring Household Food Security.* Alexandria, Virginia: Food and Nutrition Service, United States Department of Agriculture.

Binswanger, H. P. 1980. "Attitudes Toward Risk: Experimental Measurement in Rural India." *American Journal of Agricultural Economics* 62 (3): 395–407.

Bisin, A. and T. Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory* 97 (2): 298–319.

Blattman, C., J. Jamison, and M. Sheridan. 2017. "Reducing Crime and Violence: Experimental Evidence from Cognitive Behavioral Therapy in Liberia." *American Economic Review* 107 (4): 1165–1206.

Bogliacino, F. and P. Ortoleva. 2013. "The Behavior of Others as a Reference Point." *Columbia Business School Research Paper* (13–55): 1–35.

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 (7909): 291–297.

Bouguen, A., Y. Huang, M. Kremer, and E. Miguel. 2019. "Using Randomized Controlled Trials to Estimate Long-Run Impacts in Development Economics." *Annual Review of Economics* 11 (1): annurev–economics–080218–030333.

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 (6357): 1287–1290.

Cantril, H. 1966. *The Pattern of Human Concern*. New Brunswick, New Jersey: Rutgers University Press.

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.

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

Cohen, J. 2001. "Defining Identification: A Theoretical Look at the Identification of Audiences with Media Characters." *Mass Communication and Society* 4: 245–264.

Conley, T. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92 (1): 1–45.

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

Deaton, A. 1992. *Understanding Consumption, Clarendon Lectures in Economics*. Oxford: Oxford University Press.

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

Eble, A. and M. Escueta. 2022. "Demand, Supply, and Learning in a Very Low-income Context." *EdWorkingPaper No . 21-473* (20): 1–74.

Egger, D., J. Haushofer, E. Miguel, P. Niehaus, and M. W. Walker. 2022. "General Equilibrium Effects of Cash Transfers: Experimental Evidence from Kenya." *Econometrica* 90 (6): 2603–2643.

Gabaix, X. and D. Laibson. 2017. "Myopia and Discounting." *National Bureau of Economic Research Working Paper* 23254: 1–51.

Genicot, G. and D. Ray. 2017. "Aspirations and Inequality." *Econometrica* 85 (2): 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. 2022. "Sex Workers, Stigma, and Self-Image: Evidence from Kolkata Brothels." *The Review of Economics and Statistics* 104 (3): 431–448.

Gneezy, U., L. Goette, C. Sprenger, and F. Zimmermann. 2017. "The Limits of Expectations-Based Reference Dependence." *Journal of the European Economic Association* 15 (4): 861–876.

Green, M. and T. Brock. 2000. "The Role of Transportation in the Persuasiveness of Public Narratives." *Journal of Personality and Social Psychology* 79 (5): 701–721.

Greene, A. L., H. J. Sullivan, and K. Beyard-Tyler. 1982. "Attitudinal Effects of the Use of Role Models in Information About Sex-Typed Careers." *Journal of Educational Psychology* 74 (3): 393–398.

Hadley, C., D. Lindstrom, F. Tessema, and T. Belachew. 2008. "Gender Bias in the Food Insecurity Experience of Ethiopian Adolescents." *Social Science & Medicine* 66 (2): 427–438.

Haushofer, J., R. Mudida, and J. Shapiro. 2020. "The Comparative Impact of Cash Transfers and a Psychotherapy Program on Psychological and Economic Well-Being." *National Bureau of Economic Research* 28106.

Heath, C., R. P. Larrick, and G. Wu. 1999. "Goals as Reference Points." *Cognitive psychology* 38 (1): 79–109.

Heller, S. B., A. K. Shah, J. Guryan, J. Ludwig, S. Mullainathan, and H. A. Pollack. 2016. "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago." *The Quarterly Journal of Economics* 132 (1): 1–54.

International Monetary Fund. 2018. *The Federal Republic of Ethiopia, 2017 Article IV Consultation Staff Report, IMF Country Report No.18/18.* Washington D.C: International Monetary Fund.

Janzen, S. A., N. Magnan, S. Sharma, and W. M. 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." *Quarterly Journal of Economics* 124 (3): 1057–1094.

John, A. and K. Orkin. 2021. "Can Simple Psychological Interventions Increase Preventive Health Investment?" *Journal of the European Economic Association* Jvab052.

Kahneman, D., J. L. Knetsch, and R. Thaler. 1986. "Fairness as a Constraint on Profit Seeking: Entitlements in the Market." *American Economic Review* : 728–741.

Kearney, M. S. and P. B. Levine. 2020. "Role Models, Mentors, and Media Influences." *The Future of Children* 30 (1): 83–106.

Kling, J. R., J. B. Liebman, and L. F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Kőszegi, B. 2010. "Introduction to Reference-Dependent Preferences." `https://neuroeconomics.org/documents/Koszegi_Workshop2010.pdf`.

Kőszegi, B. and M. Rabin. 2006. "A Model of Reference-Dependent Preferences." *The Quarterly Journal of Economics* 121 (4): 1133–1165.

Kremer, M., G. Rao, and F. Schilbach. 2019. "Behavioral Development Economics." In *Handbook of Behavioral Economics: Applications and Foundations 1*, vol. 2. Elsevier, 345–458.

La Ferrara, E. 2016. "Mass Media and Social Change: Can We Use Television to Fight Poverty?" *Journal of the European Economic Association* 14 (4): 791–827.

———. 2019. "Presidential Address: Aspirations, Social Norms, and Development." *Journal of the European Economic Association* 17 (6): 1687–1722.

Levenson, H. 1974. "Activism and Powerful Others: Distinctions Within the Concept of Internal-External Control." *Journal of Personality Assessment* 38 (4): 377–383.

Locke, E. A. and G. P. Latham. 1990. *A Theory of Goal Setting and Task Performance.* Englewood Cliffs, New Jersey: Prentice Hall.

———. 2002. "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35 Year Odyssey." *American Psychologist* 57 (9): 705–717.

Lubega, P., F. Nakakawa, G. Narciso, C. Newman, A. N. Kaaya, C. Kityo, and G. A. Tumuhimbise. 2021. "Body and Mind: Experimental Evidence from Women Living with HIV." *Journal of Development Economics* 150: 102613.

Lybbert, T. and B. Wydick. 2018. "Poverty, Aspirations and the Economics of Hope." *Economic Development and Cultural Change* 66 (4): 709–753.

———. 2019. "Hope as Aspirations, Agency, and Pathways: Poverty Dynamics and Microfinance in Oaxaca, Mexico." In *Economics of Poverty Traps*, edited by C. B. Barrett, M. R. Carter, and J.-P. Chavas, chap. 4. Chicago, IL: University of Chicago Press, 153–177.

Macours, K. and R. Vakis. 2018. "Sustaining Impacts When Transfers End. Women Leaders, Aspirations, and Investments in Children." In *The Economics of Poverty Traps*, edited by C. B. Barrett, M. Carter, J.-P. Chavas, and M. R. Carter. Chicago: University of Chicago Press, 325–356.

Madrian, B. C. and D. F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *The Quarterly Journal of Economics* 116 (4): 1149–1187.

Mani, A. and E. Riley. 2021. "Social Networks, Role Models, Peer Effects, and Aspirations." In *Social Mobility in Developing Countries: Concepts, Methods, and Determinants*, edited by V. Iversen, A. Krishna, and K. Sen, chap. 17. Oxford, United Kingdom: Oxford University Press, 424–447.

Markle, A., G. Wu, R. White, and A. Sackett. 2018. "Goals as Reference Points in Marathon Running: A novel Test of Reference Dependence." *Journal of Risk and Uncertainty* 56 (1): 19–50.

McKelway, M. 2021. "Women's Employment in India: Intra-Household and Intra-Personal Constraints." *Dartmouth College Working Paper* : 1–102.

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

Miguel, E. and M. Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72 (1): 159–217.

Neumark, D. and A. Postlewaite. 1998. "Relative Income Concerns and the Rise in Married Women's Employment." *Journal of Public Economics* 70 (1): 157–183.

O'Donoghue, T. and C. Sprenger. 2018. "Reference-Dependent Preferences." In *Handbook of Behavioral Economics - Foundations and Applications 1*, *Handbook of Behavioral Economics: Applications and Foundations 1*, vol. 1, edited by B. D. Bernheim, S. DellaVigna, and D. Laibson. North-Holland, 1–77.

Orkin, K., R. Garlick, M. Mahmud, R. Sedlmayr, J. Haushofer, and S. Dercon. 2023. "Aspiring to a Better Future: Can a Simple Psychological Intervention Reduce Poverty?" *University of Oxford Working Paper* : 1–43.

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 (3): 226–54.

Ray, D. 2006. "Aspirations, Poverty and Economic Change." In *Understanding Poverty*, edited by Banerjee, Abhijit and Benabou, Roland and Mookherjee, Dilip. Oxford: Oxford University Press, 409–422.

Riley, E. 2022. "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. I., B. Wydick, and T. J. Lybbert. 2021. "Can Hope Elevate Microfinance? Evidence from Oaxaca, Mexico." *Oxford Economic Papers* 74 (1): 236–264.

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

Schunk, D. H. 1983. "Developing Children's Self-Efficacy and Skills: The Roles of Social Comparative Information and Goal Setting." *Contemporary Educational Psychology* 8: 76–86.

Serneels, P. and S. Dercon. 2021. "Aspirations, Poverty, and Education. Evidence from India." *The Journal of Development Studies* 57 (1): 163–183.

Slater, M. D. and D. Rouner. 2002. "Entertainment-Education and Elaboration Likelihood: Understanding the Processing of Narrative Persuasion." *Communication Theory* 12 (2): 173–191.

Stout, J. G., N. Dasgupta, M. Hunsinger, and M. A. McManus. 2011. "STEMing the Tide: Using Ingroup Experts to Inoculate Women's Self-Concept in Science, Technology, Engineering and Mathematics (STEM)." *Journal of Personality and Social Psychology* 100 (2): 255–270.

Woldehanna, T., A. Mekonnen, N. Jones, B. Tefera, J. Seager, T. Alemu, and G. Asgedom. 2008. "Education Choices in Ethiopia: What Determines Whether Poor Households Send Their Children to School?" *Ethiopian Journal of Economics* 17 (1): 1–38.

Zimmerman, B. J., A. Bandura, and M. Martinez-Pons. 1992. "Self-Motivation for Academic Attainment: The Role of Self-Efficacy Beliefs and Personal Goal-Setting." *American Educational Research Journal* 29: 663–676.

For Online Publication: Appendix

Section A gives a fuller description of two of the individual stories features in the documentaries alongside with a description of the content of the placebo screenings. Section B provides details on additional data sources as well as procedures to construct all variables used throughout the paper. Section C provides a detailed description of how the analysis presented departs from our original Pre-Analysis Plan and presents the specifications used to test for the robustness of the main results. Section D presents a series of additional tables and figures that provide further details on our experimental integrity (balance, attrition, and compliance), additional descriptive statistics and data sources, robustness of our intention-to-treat and spillover estimates to alternative specifications. We also show detailed estimates on the aspirations and expectations measures across rounds, alternative mechanisms, heterogeneity of our results by baseline durable assets, and a map providing the geographic distribution of villages in our sample.

A Summary of documentaries and placebo

The treatment consisted of four documentaries about two men and two women. Two documentaries are described below. Two documentaries are not summarised here, "Immortal Treasure", about Ayelech Fikre, and "The Exemplary Achievement", about Waki Feyyera. The four documentaries and an example of the placebo segments are available at: https://www.youtube.com/channel/UCqfoNjCzt8YPjTRWQaMQfAg.

Beshir Malim Yisak, in the video "The Fast Journey"

Beshir Malim Yisak is a farmer living 658 kilometres south of Addis Ababa. He is 27 years old, married, with two children. He has no formal education but is considered a model farmer in the area for his considerable achievement in a short period of time. Five years ago, in an area where most of the inhabitants usually breed cattle, Beshir started crop production. He consulted an agricultural expert in a local NGO about good farming practices and implemented everything he learned. He started planting vegetables, which he sold at the market, and bought a pair of oxen after a good harvest. Three years later, Beshir used money he had saved to purchase a water pump from Addis Ababa, with the help of the agricultural expert. Beshir was able to water a larger area with his pump than with buckets, so he rented additional land to expand his farm. He started planting papaya, sugar cane and maize and increased his productivity by improving his soil fertility. He gradually built up a large herd of cattle. He started keeping bees for honey and producing tree seedlings for sale. During 2007, when tree planting was encouraged by village administrations, he produced and distributed seedlings to seven

peasant associations and a local NGO in the area. Extension agents and fellow farmers speak of him as an innovator and hard worker.

Teyiba Abdella, in the video "The Life-Transforming Flour Trade"

Teyiba Abdella lives in a district in the Eastern Hararghe zone of the Oromia Region. Most people in the district are involved in cultivating crops and livestock and in trade. Teyiba is engaged in both trade and farming. She married her husband, Aliya Yousuf, by choice. Her parents refused to bless her marriage, so Teyiba and Aliya started their married life with hardly any income or assets. Their fellow villagers contributed one birr each to help them start their life together. Using the neighbours' contributions as seed money, Teyiba began trading wheat flour. She used to walk three hours to market carrying 50 kilograms of flour on her back. A woman who owns a flour mill in the market town observed her efforts and offered her credit to purchase flour. After selling the flour she obtained on credit, Teyiba paid back her debt and saved her profits. Because she paid her debts on time, the miller started giving her up to 100 kilograms of flour on credit. Teyiba also began trading eggs and chickens and bought a donkey to carry loads to the market. Then she and her husband opened their own shop. They built themselves a house and bought land in the nearby village to build another house. Teyiba's husband does most household chores while she runs the businesses. Other villagers used to criticise Teyiba for being the major breadwinner, but she rejected their criticisms. People in the village now have a high regard for her hard work and commitment. Aliya, Teyiba's husband, admires her strength and believes she is a great role model for people in their village.

Example segment from placebo treatment

The clip's title "Boru Bari", literally meaning "Tomorrow Morning", is meant to suggest the idea that "tomorrow is another day". It is a humorous take on rural life. The main character describes his current life to a journalist. He says everything is great but he looks unhappy. When pushed, he explains the reason with great hesitation, albeit humorously: his wife is having an extra-marital affair. Like the documentaries, the segment is in Oromiffa.

B Data and measures

This section provides additional details on the construction of variables used in the paper. The list is non-exhaustive: we only provide these additional details for those variables for which the main text may provide insufficient information to the reader, due to space constraint.

B.1 Conversions from Ethiopian birr to USD PPP

The survey collected data in Ethiopian birr at the time of the survey. All monetary values in the tables and figures are displayed in 2016 USD PPP.

To convert baseline and midline (six months after the screenings) values to 2016 prices, we divide the reported values in ETB by the monthly non-food national consumer price index (CPI) series (averaged over the months in which our survey took place and re-based so that it was equal to 1 in January 2016, the midpoint in our endline survey, five years after the screenings). We use the Central Statistical Authority publicly-available CPI reports (`https://web.archive.org/web/20191115152931/http://www.csa.gov.et/price-indices/consumer-price-index/category/109-cpi-2016?limitstart=0`, accessed 17/08/2021). For baseline, we divide the monetary values by 0.514. For midline, we divide the monetary values by 0.592.

To convert 2016 values to USD PPP, we use an exchange rate of 8.67 ETB per 1 USD PPP, the World Bank PPP conversion factor for private consumption in 2016. The price level ratio of PPP conversion factor (GDP) to ETB market exchange rate for 2016 was 0.41 (`https://data.worldbank.org/topic/economy-and-growth?view=chart`, accessed 27/08/2019).

B.2 Aspirations and expectations

To measure aspirations, respondents are asked the levels of outcomes they would like to achieve, on different dimensions. For income, we asked for aspirations for annual income, as the amount of cash income the household earns from all agricultural and non-agricultural activities in a year. Aspired wealth is asked in relation to durable wealth (including housing, vehicles, furniture and other valuable durables). Aspired education is measured as the years of education that the respondent would like their oldest child to achieve.[22]

The aspirations and expectations indices are standardised indices of these three dimensions, constructed following Anderson (2008). For these measures to mirror a concept of aspirations as used in psychology (goals that are attainable, as in Bandura et al. (2001) or Bandura and Locke (2003), or, as reasonable reference points as considered in economics, we anchored the elicitation of these aspirations during the interview by asking respondents first to describe the current position in each dimension before reporting aspirations.

For robustness purpose, we use the same approach to collect information on respondents' expectations, measured as the levels the respondent expects to reach in ten years on the same dimensions that compose the aspiration index. To calculate the aspirations

22. We used codes for different types of post-school education, so completing a three-year university degree was 15 years of education, while a one-year diploma is 13 years.

and expectations gaps we subtract to each elicited dimension the current level reported. For the current level of education, we use the respondents' own education level.

The measures have been tested before in Ethiopia and correlated with demographic and other characteristics variables in ways as expected Bernard and Taffesse (2014).[23]

B.3 Agricultural investment

B.3.1 Modern crop and livestock inputs

Spending on crop inputs includes expenditure on seeds (bartered or purchased), fertiliser, herbicides, tractor hire and other non-labour inputs in the last long rains season.[24]

We record the number of person-days of family and hired labour for crop agriculture in the last long rains season. This is collected by plot and crop and summed. To value this labour, we multiply by the median male wage for each village across all crop-related activities (i.e. seeding, planting, weeding, harvesting). Female wages are rarely measured, reflecting that most wage labour in agriculture is male. If a village wage is not available, we use the community-level wage. If there is no wage reported in the community (*kebele*), we use the sample median of 50 ETB per day (about $5.76 PPP per day) for that *kebele* instead.

Spending on livestock and poultry inputs includes expenditure on the purchase of inputs required for livestock in the past 12 months: feed, veterinary supplies, and hired labour. Spending on purchase of livestock is the sum of spending on all animals in the last 12 months.

B.3.2 Land

Total land area under cultivation is the area cultivated by the household across all plots in the last long rainy season. It excludes land rented out but includes land rented or sharecropped in. Areas are given in local units and converted into hectares.

23. The validity and reliability tests were performed on the aspiration indicator only and rested on a slightly different wording, namely "what is the level that (they) would like to achieve in their life". The phrase "in your life" was removed so respondents would report the highest achievement they sought rather than the level at the end of their life. Results from Bernard and Taffesse (2014) suggest the measure had high reliability and validity, provided experienced enumerators are used. The enumerators in this study were all experienced. Two days of the two weeks of survey training were dedicated to the administration of the aspiration-related questions.

24. We only have prices for seed purchases by some households. We use either the household-level purchase price or, if not available, the sample level median of seed price. We were unable to find any price points for four crops after this process, which affects a handful of observations. We use the price of white teff seeds for tikur teff, grass pea for cow peas, zengada for oats, and an average of wheat and barley seeds for wasira.

B.3.3 Assets

The value of livestock and poultry is the sum of the value of all livestock varieties owned by the household. We construct prices using the sale prices reported by households for each variety of livestock. If the household has not sold the type of livestock it owns in the last 12 months, we use the first available median of the sale prices at the village, kebele or screening site level. We compute the unit price of each livestock variety. If a price is above the 99th percentile, we replace it with the first available of the median price at the village, kebele or screening site level.

B.4 Beliefs, preferences, and information

B.4.1 Time preferences

As is common in the literature, we measure aspects of time preferences by asking individuals to choose between receiving a smaller reward immediately and receiving a larger reward with some delay (Tversky and Kahneman, 1974). The specific measurement tool is from Ashraf, Karlan, and Yin (2006), in a similar context with participants with low literacy levels. We use hypothetical rather than incentivised choices, given recent evidence suggests that hypothetical and incentivized choices over money provide fairly similar results (Ubfal, 2016; Madden et al., 2004; Falk et al., 2018).

We ask individuals to consider a situation in which they were about to receive a gift. They are first asked three questions in the "near-term" frame:

1. Would they prefer the gift of 100 ETB today or could instead choose to receive a gift of 125 ETB in one month?
2. If they answer 100 ETB to question 1, they are asked if they prefer 100 ETB today or 150 ETB in one month.
3. Individuals are then asked how much they would need to receive to wait one month for the payment instead of receiving 100 ETB today, with a ceiling of 1,000 ETB, implying a discount factor of at least 0.1.

We create three indicator variables as crude measures of an individual's discount rate, the extent to which they discount rewards when they are in the future. The indicators are for whether an individual is Patient, Impatient or Very Impatient. Individuals who select 125 ETB over 100 ETB in Question 1 are classified as Patient. Individuals who select 150 ETB over 100 ETB in Question 2 (but did not select 125 ETB over 100 ETB) are classified as Slightly Impatient. Individuals who need to receive over 150 ETB are classified as Very Impatient.[25]

25. We recode 47 observations over the three rounds who give inconsistent answers as missing. They prefer 100 ETB in the first two questions but choose less than 150 ETB in one month for this question. We view them as misunderstanding the question.

We also capture whether individuals' choices are consistent with quasi-hyperbolic discounting (rather than exponential discounting). We ask the first two questions, but over a more *distant* time frame (one vs two months). As in Ashraf, Karlan, and Yin (2006), we create two indicators. Those who are coded as "Present-biased" or "Hyperbolic" choose the immediate reward in the near term frame and the delayed reward in the distant frame. Those who are coded as "Patient now and impatient later" choose the delayed reward in the near term frame and the immediate reward in the distant frame. This could arise if individuals have funds now, but think it is likely they will be liquidity-constrained in two months time (for example, due to seasonality). Table A.8 shows that the proportion of our sample who are present biased is 34 per centsix months after the baseline, compared to 28 per centin Ashraf, Karlan, and Yin (2006), and the proportion who are "Patient now and impatient later" is 22 per cent, compared to 20 per cent.

We note increases in the portion of the sample who are impatient over time (it increases from 68 to 80 per cent over five years). We also find an increase in the proportion of people who are present-biased from 34 to 53 per cent. This could be because we neglected to alter our measures to account for inflation, so we use the same amounts in the baseline and endline five years later. The increase in impatience is consistent with the monetary reward for waiting being worth less in real terms at endline than at the six-months follow-up. However, this could also reflect recent concerns in the literature that standard measures of time preference over money may be affected by prevailing credit market conditions outside the experiment. For example, Dean and Sautmann (2021) find that the discount rate is related to whether or not an individual has just suffered an adverse shock. The endline took place just after a drought. The lower discount factor (i.e. more impatience) is consistent with the idea that people on average become more present-biased after an adverse shock. Both explanations would likely affect all treatment groups similarly, so should not jeopardise estimation of treatment effects. While measures of risk and time preferences are based on hypothetical questions, not incentivised measures, recent work suggests this does not affect answers (Ubfal, 2016; Madden et al., 2004; Falk et al., 2018). Measurement issues also cannot account for any treatment effects observed as they are constant across treatment groups.

B.4.2 Risk preferences

We use a survey-based measure of risk preferences based on Binswanger (1980). In the main measure presented in text, we ask participants about a hypothetical maize sale. We ask which of five hypothetical payouts respondents would choose for this maize, if the payout was determined by a coin toss. In the first payout, they would be certain to be paid 300 ETB for one 50kg bag of maize. In the second, they would have an equal chance of receiving 200 ETB or 400 ETB. After that, there are three more payouts, which

increase in both mean and variance, as shown in Table A.7. We treat this choice as a categorical variable, with values of 1 for those who made the most risk averse choice and 5 for those who chose the most risk neutral to risk loving option.[26] We prefer this to estimating risk preference parameters assuming a specific functional form for the utility function, as this relies on all households making decisions under uncertainty in the same way.[27]

Results are robust to different methods of constructing a measure of risk aversion from the maize sale scenario and using a different measure with a similar payout structure (available upon request). In one measure, we calculate a measure of risk aversion using a constant partial risk aversion utility function as in (Binswanger, 1980). This is of the form $U = (1-S)M^{1-S}$, where U is utility, S is partial risk aversion (fixed regardless of the level of payoff), and M is the certainty equivalent of a given lottery. The survey measure only allocates respondents to have coefficients of partial risk aversion S in an interval; we allocate individuals a value within that interval as shown in Column 9 in Table A.7. For options 2-4, individuals are allocated the geometric mean of the endpoints. For option 1, individuals are allocated 3.26, the lower bound of the interval. For option 5, we assume no respondent is risk loving, so the interval has an endpoint of 0, and use the arithmetic mean of the interval.. In a second measure, we create an indicator variable. Participants are coded as risk neutral if they choose the fourth or fifth payouts. Otherwise, they are risk averse.

In a third measure, we use a different scenario, a gamble, where individuals bet on the outcome when someone flips a coin. They are again asked to choose among five payouts, which follow the same structure as the maize sale scenario, but the stakes are divided by 100. We construct the same three measures as for the maize sale. We prefer the maize sale measure. First, the hypothetical choice was in a real-world scenario related to their livelihood, while the gamble was presented as a game. The maize sale is potentially more analogous to the choices we study. Second, the maize sale had higher stakes, again more analogous to the choices we study. In incentivised choices, higher stakes are associated with more risk-averse behaviour (Holt and Laury, 2002), and we also observe this in our hypothetical choices. Third, there is a slight imbalance in the gamble measure at baseline. However, results are similar using this measure or the maize-related one. This is unsurprising as there are strong correlations between the two measures: at the six-month follow-up in the control group, correlation coefficients are 0.66 (categorical measure), 0.66 (estimating coefficient of partial relative risk aversion) and 0.5 (indicator for being risk neutral).

26. The distribution of individuals across categories is similar to results from the same measure in the Ethiopian Rural Household Survey (Hill, Hoddinott, and Kumar, 2013).

27. For example, Harrison, Humphrey, and Verschoor (2010) have shown that in Ethiopia, roughly half of households make decisions under uncertainty that are consistent with cumulative prospect theory rather than with expected utility theory.

B.4.3 Locus of Control, Perceptions of Poverty, and Grit

Locus of control We construct two sub-scales from the IPC scale (Levenson, 1981), which captures if people see outcomes as contingent on their behaviour (internal locus of control), or as a result of chance, luck or fate (chance locus of control). Responses could be from 1 on an item if respondents "Strongly disagree" to 4 if they "Strongly agree", with no neutral option. For example, higher values on the Internal scale indicate that respondents see outcomes as contingent on individual behaviour.[28]

Causes of poverty Similarly, we construct two sub-scales of the Attributions for Poverty scale (Feagin, 1972, 1975) which capture if individuals use Individualistic explanations for poverty, or Fatalistic explanations for poverty.

We assess the reliability of both locus of control and causes of poverty. Items that met any of the following criteria were removed: low corrected item-total correlation (0.25); increased Cronbach's α if item removed; low item variation (80 per cent identical responses on the item); low loading on primary unrotated factor (< 0.30), and high cross-loading (> 0.30) (Lamping et al., 2002). If respondents did not answer all items in a sub-scale, we code the items they do not answer as missing and adjust their score to generate a homogeneous score range using an appropriate multiplier. However, if a respondent is missing over 60 per cent of the items of a sub-scale or has given the same answer to all items on the scale, we take that as an indication of low reliability of the observation, and replace the sub-scale score as missing.

Grit We construct a standardised index of grit (Anderson, 2008) from two measures in the vein of Alan, Boneva, and Ertac (2019). These measures take values from 1 ("*Strongly Disagree*") to 4 ("*Strongly Agree*"). The first question asks respondent to agree/disagree with the following statement: "*I do a thorough job*". The second asks respondents whether they agree/disagree with the statement: "I make plans and follow through with them".

B.4.4 Information

To explore whether the decisions taken by farmers in terms of investment are at all related to the decisions that the subjects in the documentaries had mentioned, we construct an index (Anderson, 2008) using the following variables, equal to 1 (and zero

28. At baseline, on average, participants have higher scores on the internal locus of control (a mean of 15.54) than on the chance (12.38) or others (11.79) sub-scales. The Internal scores are almost identical to the American samples in Levenson (1981). However, the Chance scores are much higher (the mean was 61 per cent of the total possible score, compared to 37 per cent in the studies in Levenson (1981)). Unsurprisingly, poor people in an isolated, highly religious area with limited or no education are more likely to believe that fate or chance control their outcomes. Other samples in Africa behave similarly to our sample (Cheng et al., 2013; Rossier, Dahourou, and McCrae, 2005; Reimanis and Posen, 1980; Van Haaften and van de Vijver, 1999)

otherwise) if the household: (i) earns any income from trading, (ii) attends community meeting to discuss agricultural issues, (iii) seeks visits by an agricultural expert, (iv) uses any irrigation technique; takes advice by agricultural extension on (v) land preparation, (vi) seeds, or (vii) fertilisers; (viii) grows cash crops, (ix) uses a water pump, (x) builds stone bands and terracing, (xi) applies water conservation/water harvesting practices, (xii) applies crop rotation, (xiii) uses cattle in crop activities, and the (ixv) number of visits received by an agricultural extension worker (the only non-binary component of the index). A few variables that we pre-specified to be part of the index did not have sufficient variation. For example, only 6 households reported earning income from grain milling. We exclude from the information index variables that were positively answered by less than 2 per cent of our sample.

B.4.5 Expected fertiliser yields

We developed a novel battery of questions to elicit expectations about the increase in output from the use of modern (phosphate-based) fertilisers. We asked a list of questions to the household head to elicit how many kilos of output they would expect to produce on an hectare of their land if 0, 50, 100, 150 kilos of fertiliser were applied. Specifically the producers were asked "*In a [good/bad] year, how much [Sorghum/Maize] can one expect from a one hectare plot if [0/50/100/150] kg of fertiliser is applied?*" creating sixteen responses where we varied whether the hypothetical season was good or bad (in terms of agronomic conditions) and whether the crop produced was maize or sorghum. To combine these answers, we first estimate the elasticity of expected output relative to fertiliser by regressing the answers to these questions on the four quantities of fertiliser for each respondent, by crop and hypothetical season, to generate four expected yields per respondent (i.e. expected yield from an extra kilo of fertiliser in a good/bad season for sorghum/maize). Next, we combine these four expected yield estimates into a single Andersen et al. (2008) index.

B.5 Consumption, food security, housing, and well-being

B.5.1 Consumption and food security

All consumption variables are constructed in USD PPP and transformed into adult equivalent units, where adult equivalent is constructed using the OECD scale.[29]

Food consumption is the sum of the value of food consumed from various sources over the past 7 days, (divided by 7 and multiplied by 30 to obtain a monthly estimate).

29. The number of adult equivalent household members is calculated as 1 (for the household head) plus the weighted sum of the number of children (defined as individuals aged below 16) and the number of other adults (excluding the household head), where the weights are 0.5 and 0.7 for the number of children and the number of other adults, respectively.

This includes food purchased, received via barter, gifts, loans, wages in kind and self-production. Following Beegle et al. (2012), for purchased food items, we use reported prices, and for food received via barter, gifts, loans and wages in kind and self-produced food items, we construct prices using the first available level of price of purchased food from the following: household-level price, screening site level median, median from the neighbouring kebele, sample level median. Non-food small-item consumption is the sum of frequent non-food consumption, with a recall period of one month. Items included are: toiletries, transportation costs, mobile phone costs, energy, cigarettes and tobacco, repair, tailor, barber, other services and other small purchases (less than 100 ETB, or $11.5 PPP). Non-food lumpy consumption is the sum of expenses made over the past 12 months (divided by 12 to obtain monthly estimates), from the following list of items: clothing and footwear, utensils, beddings, school expenses, health expenses, funerals, weddings, religious expenses, contribution to community projects, land taxes and other large purchases (more than 100 ETB, or $11.5 PPP).

We use two measures of food security. Food shortage in the lean season is defined as the number of months in the last 12 Ethiopian months that the household had problems satisfying their food needs. We also use a version of the United States Department of Agriculture's food insecurity questionnaire (Bickel et al., 2000; Andrews et al., 2000) adapted for Ethiopia (Hadley et al., 2008).

B.5.2 Housing

The value of house is assessed by asking the household head how much their house would cost to build today (in current prices), including materials and labor costs. The value was replaced as missing if the value was reported as zero. In addition, enumerators were asked to report on non-organic roofing and presence of own toilet, through direct observation.

B.5.3 Subjective well-being

The subjective well-being is measured using two items indicating best and happiest life. Best life is measured by showing respondents a picture of a ladder with 10 steps (Cantril, 1966). They are told the top of the ladder represents the best possible life for them and the bottom step represents the worst possible. They are then asked, "*Where on the ladder do you feel you personally stand at present?*" The above question was repeated to measure happiest life, with the top and bottom of the ladder representing the happiest and most miserable possible life.

C Deviations from the Pre-Analysis Plan (PAP)

This study was pre-registered on the AEA RCT Registry (ID: AEARCTR-0001483) under the title "The Future in Mind: Aspirations and Forward-Looking Behaviour in the Short and Long Run in Rural Ethiopia". Pre-registration took place on February 15, 2017, after the data-collection was completed but before we started analysing any of the treatment effects on the five-year follow-up data.

In the rest of this subsection we list changes from the original plan and rationales for these changes.

C.1 Trimming strategy

Our PAP described that we would trim our sample for all continuous outcome variables used in the paper. We had originally described that we would trim observations that are four standard deviations or more above or below the sample mean for a continuous outcome variable. Instead we trim our sample uniformly for values of the outcome variables that are above the 99th percentile, to avoid variable-specific differential levels of trimming. In particular, due to a few outliers at baseline in the aspirations and expectations variables, we realised that this trimming strategy was not correctly removing values that were so large to affect the sample standard deviation.

C.2 Changes to family of outcomes and hypothesis

We changed our main specification and definition of focal outcomes because of unforeseen study design issues and changes in our theoretical framework.

The PAP defined three primary hypotheses, which were sub-divided into eight sub-hypotheses, and two secondary hypothesis, which were sub-divided into eight sub-hypotheses. We redefined the set of outcomes in the primary hypothesis based on the changes to our theoretical framework, as some of the outcomes did not fit well under the umbrella of the family of outcomes set in the PAP. Below, we report how we re-defined the families of outcomes across new set of four primary hypotheses.

1. **Aspirations, expectations, and self-beliefs**
   - **Sub-families of outcomes**. Our first pre-specified primary hypothesis in the PAP was meant to test whether the intervention affected self-beliefs, through four sub-families of outcomes: (i) aspirations, (ii) expectations, (iii) belief in their ability to control their own circumstances, (iv) belief in the extent to which their lives are controlled by chance. Because our theory focuses on the role of reference-point, we see other self-beliefs as a potential mechanism through which the intervention might have changed reference-points, and so decided to move

the last two sub-families of outcome out of this hypothesis. We test whether the intervention affected these beliefs separately in Table 6.

- **Index-construction method**. Our pre-specified indices of aspirations (or expectations) were measured over four dimensions: income, wealth, education and social status (the latter measured as the percentage of community members that would ask for the respondent's advice at times of important decisions). We removed social status from our index, following (Beaman et al., 2012), who also dropped a dimension with lower internal reliability from their index of aspirations.[30] We had originally proposed to weight the four dimensions according to respondents' own assessment of each dimension's significance for them, to account for heterogeneity in valued attributes of life. We would have used these weights to aggregate the standardised responses to each of the four dimensions into an index.[31] Instead, we prefer reporting Anderson (2008) indices of our three dimensions of aspirations/expectations, which are data-driven and would reliably aggregated the different dimension given that we had removed social status from the indices. Our results remain unchanged by the type of index used (results available upon request).

2. **Labour supply and human capital investments**
   - **Sub-families of outcomes**. Our second pre-specified primary hypothesis in the PAP was meant to test whether the intervention affected two sub-families of outcomes: (i) labour supply, (ii) human capital investments. Because the returns to these two activities are likely to yield changes across different time-horizons, we decide to split these two sub-families into two separate families of outcomes. Labour supply was added to the family of economic behaviours. Whereas we treat human capital investments as a stand-alone separate family.
   - **Measures of human capital investments**. We made five changes to the pre-specified measures of human capital investments.
     (a) We had pre-specified that we would measure enrolment of children aged 6-15. However, we realised that to be consistent with the primary school starting age, we should focus instead on children aged 7-15. The results are not affected by this small deviation from the plan.[32]
     (b) We added our analysis of education variables (enrolment, time-spent studying, time-spent in school) for children aged 16-20, who would have been

30. Earlier validations of our survey instruments had also found that social status had lower internal reliability (Bernard and Taffesse, 2014). The index with three dimensions had a Cronbach's alpha ranging from 0.27-0.51, which the inclusion of social status decreased by about 12 per cent.

31. If a_i^k is individual i's aspiration for dimension k, w_i^k is the weight that individual i assigned to this dimension. μ_i^k and σ_i^k measure the sample mean and standard deviation at baseline on dimension k. The standardised index was defined as the weighted average of each standardized component.

32. Schooling is compulsory from ages 7 to 15: children are supposed to enrol in Grade 1 when they have turned 7 and stay until Grade 8, when they would be about 14 or 15.

aged 11-15 at the time of the intervention. We think this cohort is of particular interest, because we had originally found positive enrolment effects among those aged 7-15 in our analysis of the six-months follow-up data (as shown in bottom panel of Appendix Table A.11). We did pre-specifiy that this would be a secondary analysis that we had planned to carry out using the household-member-level dataset, but use the household-level measures because of a coding error on CSPro used for the survey.

(c) We could not include a measure of absenteeism because of a coding error on CSPro, which made this measure inconsistently measured across observations.

(d) We add a measure of educational attainment for the 16-20 cohort to this family of outcomes.

3. **Economic behaviour**

- **Sub-families of outcomes**. Our third pre-specified primary hypothesis in the PAP was meant to test whether the intervention affected economic behaviours, through two sub-families of outcomes: (i) savings and credit, (ii) investment flows. Because savings are the likely channel through which further investments can be financed, we decide to remove the savings and credit sub-family from this family of outcomes. We found no effects on savings after five years (Appendix Table A.17).

  Instead, we added labour supply and the value of productive assets. The value of productive assets (tools and livestock) had been pre-specified to be part of a secondary hypothesis within the family of outcomes broadly related to welfare, but we prefer to present them as an additional measure of productive investments and behaviour. We see these two sub-families of outcomes (labour supply and value of productive assets) to be consistent with the types of future-oriented behaviours that involve effort towards achieving a long-term goal.

- **Measures of agricultural investments**. We made four changes to the outcomes defined in PAP that belong to the sub-family of outcomes measuring investment flows.

  (a) We had pre-specified the total spending of crop and livestock (combined) as a focal outcome for this hypothesis, but we report a more nuanced set of variables to capture allocation of investment by different activity.

  (b) We report both the intensive and extensive margin on spending on crop inputs, livestock inputs, and hired labour.

  (c) We do not report effects on the value of family labour employed in the last agricultural season to focus on overall labour supply.

  (d) We had pre-specified analysing area from land rental and sharecropping, but only 14 households rented any land in and 4 households rented out any

land, so we do not report outcomes on this variable given how rarely we observed it.

4. **Household welfare**
   - **Sub-families of outcomes**. We had pre-specified six sub-families of welfare measure: (i) consumption, (ii) food security, (iii) subjective well-being, (iv) housing, (v) income, (vi) assets. In our pre-analysis plan, we had discussed how we had not planned to "aggregate variables from the four sub-hypotheses in H[ypothesis] 5 as our theory of change does not predict that all outcomes in [this hypothesis] will move in one direction, nor in which direction they might move. Some of them (e.g. consumption and investment in assets) can even plausibly be expected to move in opposite directions in response to an intervention that promotes more future-oriented behaviour." However, given the changes to our original theory of change, and the fact that we moved productive assets to be a sub-family of the economic behaviour family, we decided to aggregate these sub-families together. We did not find any effects on income (Appendix Table A.18).
   - **Measures of consumption**. Rather than focusing on the pre-specified focal outcome total consumption, we present results on its components to get a more nuanced understanding of the intervention. We replace the focal outcome with a single index for all outcomes related to household welfare.
   - **Measures of housing quality and durable assets**. We add a new sub-family of outcomes related to housing quality that we had not pre-specified. We include in this family the value of durable assets (e.g. furniture, jewellery), which we had originally pre-specified to be part of another sub-family of asset-related outcomes, but which we think does not fit conceptually together with productive assets (such as livestock and tools). We include measures of self-reported value of the house, an indicator for whether the house's roof is not organic, and an indicator for whether the house has its own private toilet.
5. **Alternative mechanisms**
   - **Measures of self-beliefs**. We had originally pre-specified this set of outcomes to be part of our primary hypothesis, but we do not see them fitting neatly under our reference-dependent framework. We think of them more as alternative mechanisms.
   - **Measures of risk preferences**. Our pre-analysis plan was mistakenly vague about the exact variable definition we would analyse to capture risk preferences. We use the answers to the questions we had pre-specified to analyse, but construct different measures to capture risk aversion as described above in section B.4.3.
   - **Measures of time preferences**. Our pre-analysis plan mistakenly pre-specified

that we would analyse the discount factor. Instead, we report on a set of indicators to characterise the degree of impatience that is more intuitive and following other examples in the literature, as described above in section B.4.3.

- **Grit**. We add a measure of Grit to our family of outcomes related to one's perception of the returns to own effort.
- **Expected fertiliser yields**. Alongside the information index, which we had pre-specified, we also explore whether beliefs about the yields from fertiliser changed, as described above in section B.4.5.

C.3 Tests for robustness of main results and spillovers effects

We run two sets of alternative specifications to test the robustness of our main results. The first set follows the analysis plan and includes three alternative specifications that include additional control variables. In a first robustness test, we control for individual and household characteristics found to be unbalanced at baseline. In a second robustness test, we estimate treatment effects using an ANCOVA specification to account for outcomes found unbalanced at baseline. Lastly, we also include a third robustness that uses a set of village-level controls and screening site fixed effects (instead of the village fixed effects) that is most comparable to the spillover tests discussed in Section 6.

The second set of alternative specifications was originally intended to form the basis of the main specification described in the analysis plan. However, we realised that these specifications did not fully leverage the experimental variation and so prefer to use them as additional robustness checks. From the original 84 villages randomly sampled for the intervention, 64 were effectively selected to be exposed to documentaries. In the analysis plan, we had mistakenly indicated that the selected 64 were done so randomly, leaving the remaining 20 as pure control villages. The 64 villages were selected for logistical reasons, enabling the screening of documentaries to occur in groups of four neighbouring villages, in a large enough closed room (typically a classroom or an agricultural extension raining center). Results presented in Section 5 of the main text focus on treated villages only. In Section 6, we add to this sample the 10 villages that would have been selected for logistical reasons had we formed quintuplets instead of quadruplets of villages to organise the screening events. We develop three alternative strategies to assess for the presence of spillovers. The first one is based on that proposed in the PAP and is presented in the main text.

The second specification uses variation in treatment intensity across treated villages, based on our randomised saturation design. Specifically, we estimate:

(5)
$$y_{i3} = \alpha + \varphi_1 TI_v + \varphi_2 PI_v + \delta_1 T_i * TI_v + \delta_2 T_i * PI_v + \rho_1 P_i * TI_v + \rho_1 P_i * PI_v + + X'_{i2}\pi^{si} + \varepsilon_i$$

where, as before, i indexes individuals (or households) and v indexes villages, y_i denotes the outcome of interest measured in the five-year follow-up survey, TI_v is an indicator for being in a treatment-intense village, PI_v is an indicator for being in a placebo-intense village. We decompose the parameter φ into φ_1 and φ_2 to test whether the spillovers of the treatment to the control group may be different in villages with more individuals exposed to the treatment. Similarly, we compare δ_1 with δ_2 to test whether our treatment effects are different in villages that were more intensely treated. We include the same control variables as in Equation (4) and cluster the standard errors at the village-level.

Our third specification exploits the spatial distribution of households to assess how exposure to treated individuals varies across 1km radii around each observation. We follow Egger et al. (2022) in order to pick the most relevant radius distance. We estimated the effect of treatment intensity within a series of non-overlapping doughnuts, $d = 1, ..., D$, each with inner radius r and outer radius $r + 1$ kilometres. We estimated a series of nested models: with a single doughnut ($d = 1$) with $r = 0$; with two doughnuts ($d = 1, 2$) with $r \in \{0, 1\}$; ...; with ten doughnuts ($d = 1, 2, ..., 10$) with $r \in \{0, 1, \ldots, 9\}$. For each specification and outcome, we then select the model which minimises the Bayesian Information Criterion (BIC). Across all specifications and summary indices, we achieved the minimum BIC with a single doughnut with with $d = 1$ and $r = 0$.

We estimated the effects of treatment intensity on our summary indices using the following specification:

(6)
$$\begin{aligned} y_{i3} =& \alpha + \varphi_1^r C_i + \delta_1^r T_i + \rho_1^r P_i + \\ & \sum_{d=1}^{D} \left(\delta_1^d \mathrm{V}_{i\neg v}^d + \delta_2^d \mathrm{TH}_{i\neg i}^d \right) + X'_{i2}\pi_{lr} + \varepsilon_i \end{aligned}$$

where variables are defined as above, $\mathrm{TH}_{i\neg i}^d$ is the total amount of tickets *assigned* to households within doughnut d of household i (excluding individual i), and $\mathrm{V}_{i\neg v}^d$ is the number of villages within doughnut d (excluding village v). Conditional on the number of villages within a given radius, the total amount of treated individuals that are within $d$km is exogenous. Our conditional exogeneity comes from the fact that the number of households invited to the documentary is a fixed discrete number depending on the random saturation to which the village was assigned. That is,

intensely-treated villages had 24 households invited to the screening, other treated villages had 6 households invited to the screening, whereas the pure control villages had none. This equation allows us to estimate the following effects:

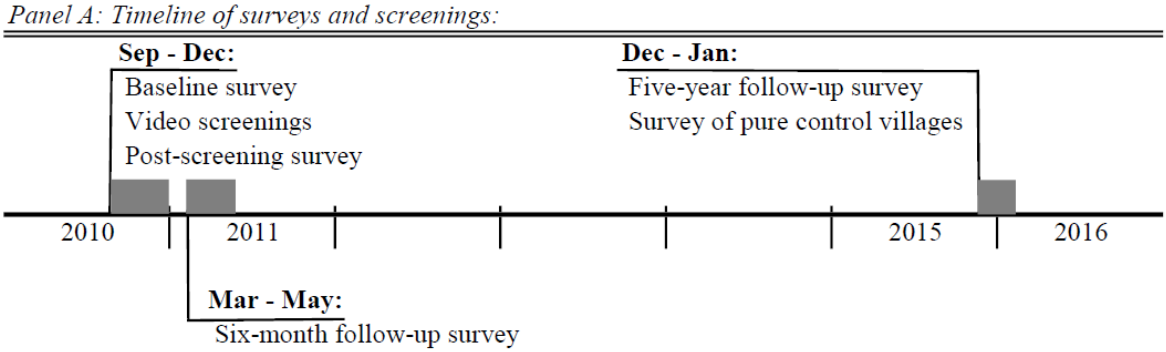
(a) δ_1^r gives us the (direct) intent-to-treat effect of the aspirational videos on individuals/households in treated villages.

(b) φ_1^r gives us the (spillover) effect of the video transfers on uninvited households in treatment villages.

(c) δ_2^d gives us the (spillover) effect of the total amount of treated individuals within the radii defined for doughnut d.

D Additional tables and figures

Panel A: Timeline of surveys and screenings:



Panel B: Schooling-age cohorts at the time of the surveys:

Cohort	Age in 2010/11	Age in 2015/16
1	11-15	16-20
2	2-10	7-15
2(a)	7-10	12-15

Figure A.1: Timeline of the study

Notes: Panel A shows the overall study timeline. Grey horizontal bars denote the periods where a survey or the screening intervention took place. Panel B shows the cohort ages of children between baseline and the five-year follow-up. These cohorts are used to define educational outcomes in the analysis.

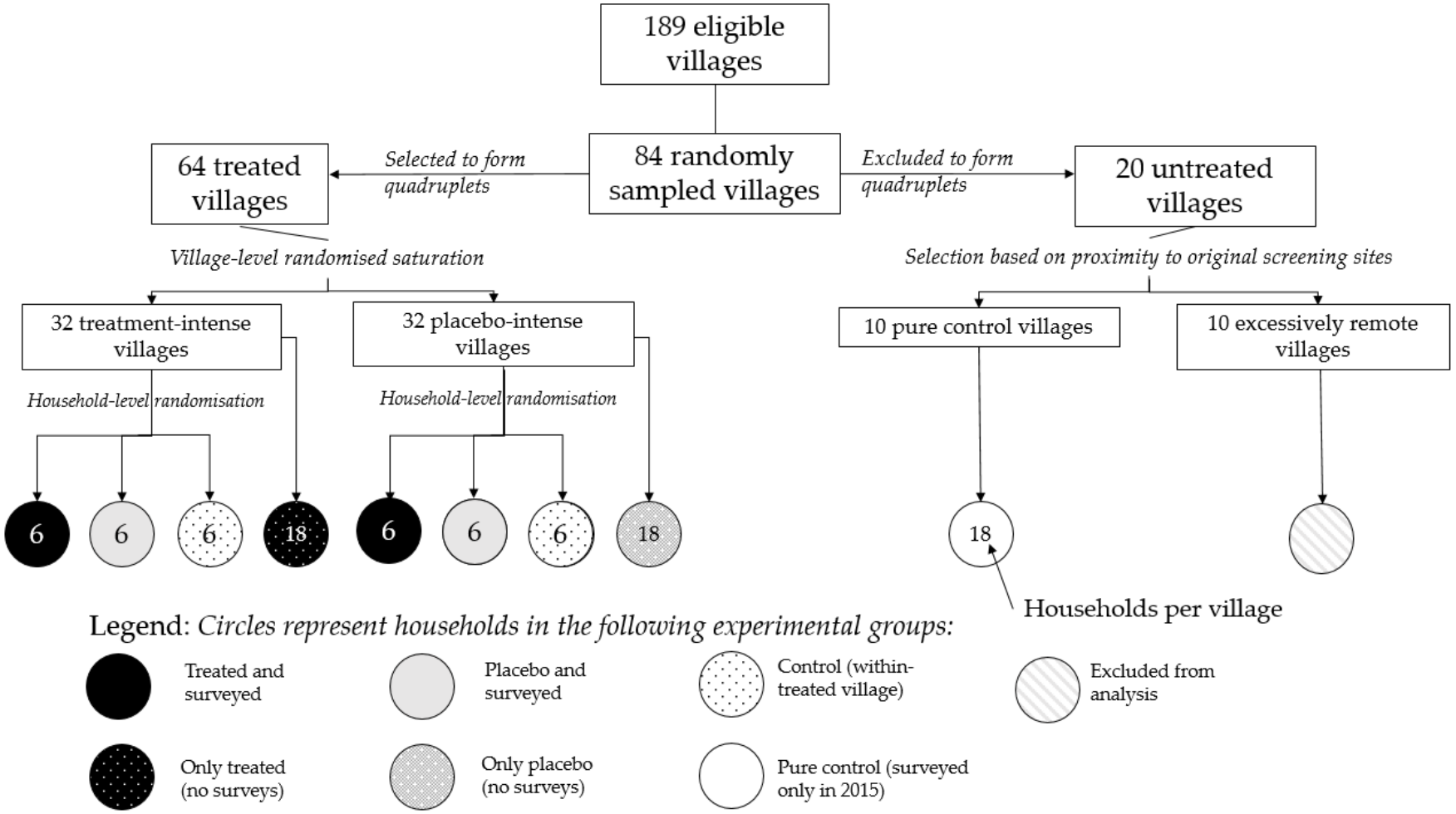


Figure A.2: Study design

Notes: Diagram of the sampling and randomisation into different experimental groups. Rectangles indicated villages, whereas the circles indicate households. Numbers inside the circle represent the number of households in each experimental group per village.

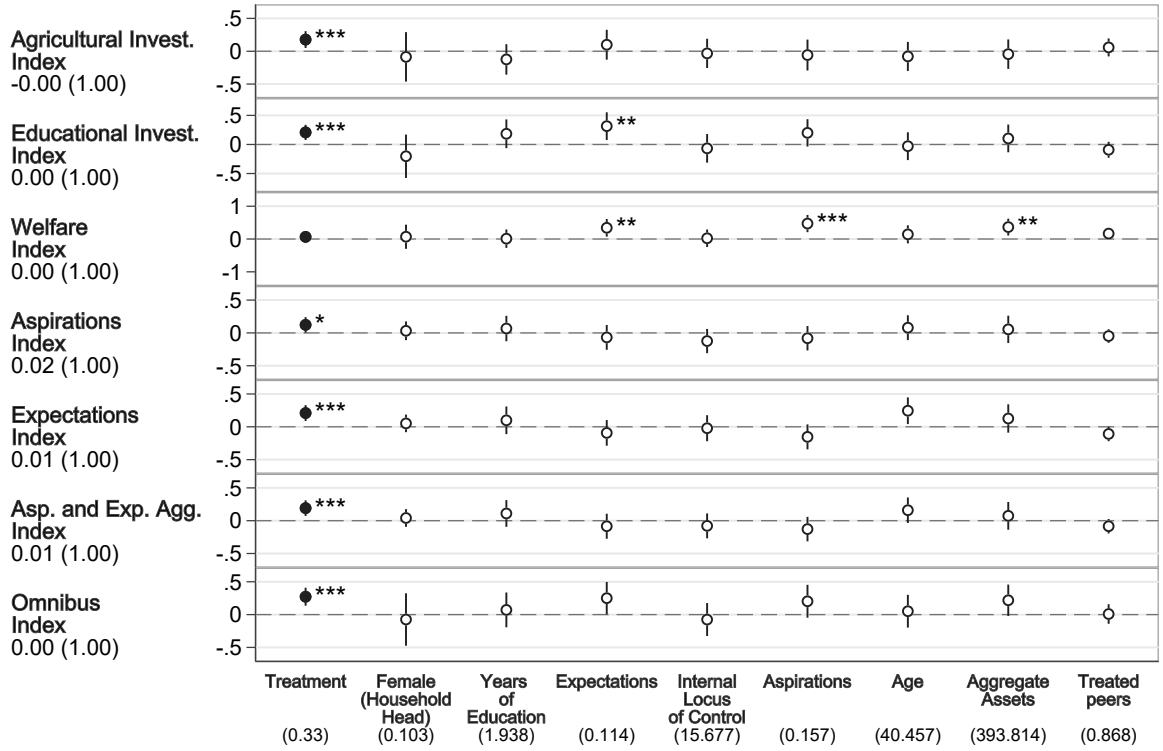


Figure A.3: Heterogeneous treatment effects on summary indices

Notes: Each coefficient represents a separate OLS regression using data five years after the intervention. The first column represents within-village intention-to-treat effects controlling for a placebo-group indicator. The comparison group comprises households from the 64 treated villages that were not invited to any screening. The second column additionally controls for an indicator equal to one if the respondent is female and report its interaction with the treatment indicator. The subsequent columns additionally control for an indicator equal to one if the baseline value of the variable reported in the horizontal axis is above the median, and report its interaction with the treatment indicator. For the household-level outcomes (agricultural investment, educational investment, welfare, and omnibus indices), the baseline heterogeneity dimensions are those reported by the household head. For the individual-level outcomes (aspirations, expectations, and the aspirations and expectations aggregate indices), the baseline heterogeneity dimensions are those reported by the individual respondent (either the spouse or the household-head), except for aggregate assets that are only observed at the household-level. The construction of the internal locus of control is described in Appendix Section B.4.3. Aggregate assets include non-productive assets, productive assets, savings, and livestock holdings. Treated peers corresponds to the number of close social connections at baseline that were invited to watch the role model videos. All regressions control for village fixed effects and characteristics of the respondent: age, years of education, an indicator for being single, and an indicator for being male. Regressions on the educational investment index additional control for the number of children aged 0-15 at baseline to account for the baseline imbalance in the number of children. Standard errors are robust to heteroskedasticity (and clustered at the household-level for the aspirations/expectations indices). Bars represent 95% confidence interval based on naive p-values. For heterogeneous effects, we correct p-values on our interaction terms for multiple testing using False-Discovery-Rate-adjusted q-values (Benjamini, Krieger, and Yekutieli, 2006). Stars correspond to significance according to the minimum q-value at which each hypothesis is rejected. These are calculated across the number of outcomes per interaction term, most relevant for determining whether heterogeneous effects are statistically significantly different than zero. We do not include FDR adjusted q-values where we correct for the number of interactions (dimensions of heterogeneity) for a given outcome, which is most relevant for determining if the magnitude of heterogeneous effects varies across dimensions of heterogeneity, as we did not anticipate being powered for such tests. The outcomes, described in Table 7, are inverse-covariance-weighted averages standardised relative to the within-village control group, following Anderson (2008).

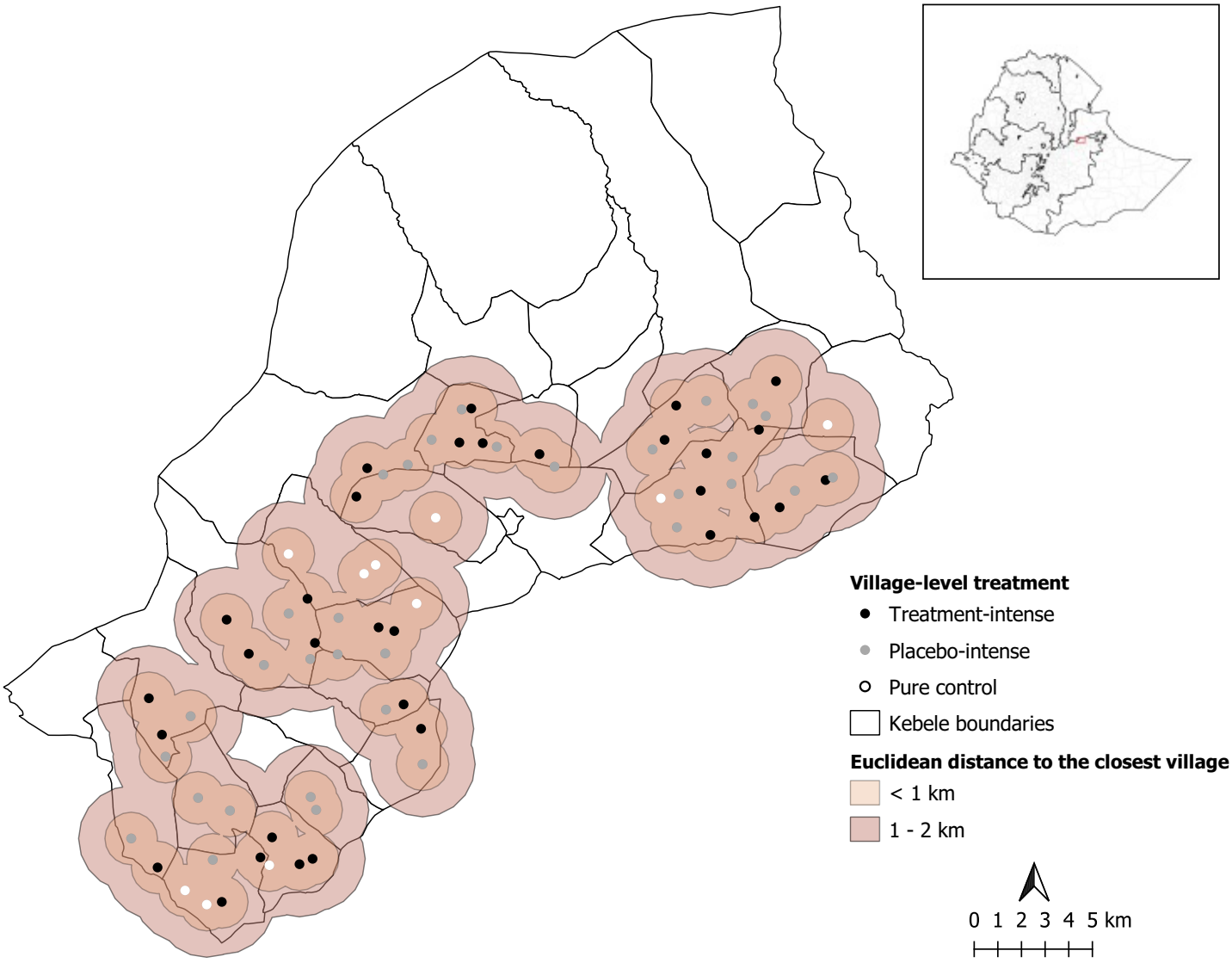


Figure A.4: Villages in the study

Notes: White dots correspond to the pure control villages used in the spillover analysis. Black dots corresponds to villages that were intensely treated. Grey dots corresponds to villages with a higher number of individuals watching the placebo videos.

Table A.1: Psychological characteristics, by terciles of durable assets

	(1) Whole sample	(2) Lower tercile	(3) Middle tercile	(4) Upper tercile	(5) p-value	(6) Observations
Aspirations gap (aspirations - current level)						
Aspired income gap (USD PPP)	20630.88	16802.07	20037.23	24972.45	0.01	2002
Aspired wealth gap (USD PPP)	10524.31	7987.79	9635.42	14079.33	0.00	2005
Aspired education gap	10.69	10.69	10.30	11.06	0.10	1939
Locus of control and causes of poverty						
Internal locus of control	12.60	12.65	12.46	12.65	0.98	2038
Individual causes of poverty	8.89	8.86	8.87	8.95	0.45	2021
Chance locus of control	12.43	12.62	12.56	12.13	0.00	2037
Fate causes of poverty	6.79	6.89	7.05	6.47	0.00	2025
Time and risk preferences						
% that is patient	0.33	0.36	0.30	0.32	0.06	2037
% that is somewhat impatient	0.13	0.13	0.13	0.13	0.93	2037
% that is most impatient	0.54	0.51	0.56	0.56	0.07	2037
% that is present biased	0.35	0.33	0.36	0.34	0.89	2004
% that is patient now and impatient later	0.22	0.20	0.22	0.25	0.02	2004
Risk aversion: most to least risk averse (1 to 5)	3.20	3.30	3.10	3.22	0.35	2010

Notes: We show descriptive statistics for the sample (column 1), divided into lower, middle and upper terciles (Columns 2-4) by the value of durable assets (excluding tools) at baseline, an approximation for living standards. Columns 5 reports the p-value from a t-test of equality between the mean of the lower and upper tercile. Columns 6 reports the number of observations. Variables are measured for the whole sample of individuals (household head and spouse) at baseline. The aspirations gaps take the measure of aspirations and subtract the current level at baseline elicited for that same dimension. To measure aspirations, respondents are asked the levels of outcomes the respondent would like to achieve, on different dimensions. Annual income is the amount of cash income the household earns from all agricultural and non-agricultural activities in a year. Wealth is durable wealth (including housing, vehicles, furniture and other valuable durables). Aspired education is measured as the "years of education that you would like your oldest child to achieve". For the current level of education, we use the respondents' own education level. The locus of control variables are based on the Internal, Powerful Others, and Chance Scale (Levenson, 1974) and capture if people see outcomes as contingent on their behaviour (internal locus of control), as a result of chance, luck or fate (chance locus of control). We use survey-based instruments to calculate risk (Binswanger, 1980) and time preferences (Ashraf, Karlan, and Yin, 2006). All other variables are defined in detail in Appendix B.

Table A.2: Balance tests — baseline household and individual characteristics

	(1) Treatment	(2) Placebo	(3) Treat. vs. placebo	(4) Control mean (SD)	(5) Max pairwise difference Total Obs.
Male	-0.01	-0.02*	0.01	0.50	0.04
	(0.01)	(0.01)	(0.01)	(0.50)	1913
	[0.64]	[0.88]	[0.57]		
Age (at baseline)	0.22	0.19	0.03	36.66	0.02
	(0.80)	(0.86)	(0.83)	(12.52)	1913
	[0.82]	[0.88]	[0.97]		
Years of education	0.25*	0.04	0.21	1.25	0.11
	(0.13)	(0.13)	(0.14)	(2.22)	1913
	[0.40]	[0.88]	[0.27]		
Marital status is single or divorced or widowed	0.03*	-0.00	0.03**	0.07	0.12
	(0.02)	(0.01)	(0.02)	(0.25)	1913
	[0.40]	[0.88]	[0.27]		
Watches television at least once a week	0.01	-0.01	0.02	0.11	0.07
	(0.02)	(0.02)	(0.02)	(0.31)	1909
	[0.78]	[0.88]	[0.32]		
Listens to radio at least once a week	0.00	-0.04	0.04	0.62	0.08
	(0.03)	(0.03)	(0.03)	(0.49)	1909
	[0.92]	[0.88]	[0.27]		
Travels outside the village within the district at least once a week	0.01	-0.00	0.01	0.28	0.04
	(0.03)	(0.03)	(0.03)	(0.45)	1913
	[0.82]	[0.88]	[0.67]		
Travels outside the district at least once a week	0.02	-0.01	0.03	0.14	0.10
	(0.02)	(0.02)	(0.02)	(0.35)	1913
	[0.64]	[0.88]	[0.27]		
Ever lived outside of current village 6 months	0.02	-0.04*	0.05**	0.17	0.13
	(0.02)	(0.02)	(0.02)	(0.38)	1913
	[0.64]	[0.88]	[0.27]		
Ever lived outside of current district 6 months	0.02	-0.02	0.04**	0.10	0.11
	(0.02)	(0.02)	(0.02)	(0.30)	1912
	[0.64]	[0.88]	[0.27]		
Joint p-value	0.35	0.17	0.02**		
Durable assets (USD PPP)	8.60	-17.07	25.67	159.88	0.07
	(25.98)	(23.60)	(20.19)	(386.12)	1077
	[0.76]	[0.80]	[0.44]		
Household size	0.20	0.15	0.06	5.50	0.10
	(0.16)	(0.16)	(0.16)	(2.15)	1090
	[0.35]	[0.80]	[0.72]		
Number of individuals aged 0-6	-0.10	0.03	-0.14*	1.41	0.12
	(0.08)	(0.08)	(0.08)	(1.12)	1090
	[0.35]	[0.80]	[0.31]		
Number of male individuals aged 7-15	0.20***	0.12*	0.07	0.75	0.21
	(0.07)	(0.07)	(0.07)	(0.89)	1090
	[0.03]**	[0.54]	[0.44]		
Number of female individuals aged 7-15	0.14**	0.02	0.12*	0.76	0.16
	(0.07)	(0.07)	(0.07)	(0.90)	1090
	[0.17]	[0.80]	[0.31]		
Adult males aged above 15	0.04	0.01	0.02	1.29	0.05
	(0.05)	(0.05)	(0.05)	(0.69)	1090
	[0.63]	[0.80]	[0.72]		
Adult female aged above 15	-0.01	-0.05	0.04	1.20	0.09
	(0.04)	(0.04)	(0.04)	(0.58)	1090
	[0.76]	[0.61]	[0.44]		
Joint p-value	0.76	0.16	0.29		

Notes: OLS estimates of baseline within-village differences across treatment arms. Outcome variables are listed on the left. The unit of observation is the individual in the upper panel and household in the lower panel. All columns include village fixed-effects. Standard errors are in parentheses and are clustered at household level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values that control for the false discovery rate (FDR), with an adjustment based on the number of outcomes tested, are reported in brackets, following Benjamini, Krieger, and Yekutieli (2006). * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. In column 5, we calculate the maximum pairwise difference between any two treatment group means and divide this by the standard deviation of the variable, following Imbens and Rubin (2015). The last row shows joint significance of the coefficients in the corresponding column from SUR estimation. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using the national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The conversion is described in Appendix B.1.

Table A.3: Village-level balance

	(1) Treatment village	(2) Control village mean (SD)	(3) Standardised difference	(4) Obs.
Altitude (in meters)	13.07	1810.70	0.32	74
	(34.75)	(175.43)		
	[0.97]			
Number of households	53.11	87.90	0.34	74
	(35.57)	(66.21)		
	[0.40]			
Area of agricultural land (hectares)	3.07	41.25	0.16	74
	(5.13)	(20.52)		
	[0.94]			
Forest (hectares)	1.69	4.40	0.21	74
	(1.49)	(8.11)		
	[0.55]			
Time to walk to nearest market (minutes)	22.69*	109.50	0.14	74
	(12.89)	(93.62)		
	[0.40]			
Distance to nearest market (in km by road)	2.01	10.82	0.12	74
	(1.24)	(9.41)		
	[0.40]			
Village has first cycle school	0.00	0.15	0.24	74
	(0.16)	(0.37)		
	[1.00]			
Village has second cycle school	-0.03	0.10	0.20	74
	(0.15)	(0.31)		
	[0.97]			
Village connected to mobile network	0.19	0.65	0.85	74
	(0.12)	(0.49)		
	[0.40]			
Percentage of hh with mobile	0.08*	0.19	0.68	73
	(0.05)	(0.12)		
	[0.40]			
Distance to next city	33.87	11632.35	0.12	74
	(523.67)	(3725.41)		
	[1.00]			
Distance to next health centre	786.55	9108.00	0.24	74
	(520.39)	(5578.49)		
	[0.40]			
Distance to next market place	460.15	10173.19	0.03	74
	(587.81)	(4695.58)		
	[0.82]			
Distance to next river	-112.67	3110.02	0.41	74
	(352.92)	(1173.63)		
	[0.97]			
Distance to next road	113.03	6953.03	0.40	74
	(611.36)	(3453.54)		
	[0.97]			
Distance to farmers training centre	-230.29	4973.67	0.63	74
	(486.56)	(3015.27)		
	[0.97]			
Distance to next school	259.38	1469.59	0.23	74
	(190.75)	(1323.95)		
	[0.44]			

Notes: OLS estimates of village differences in treatment level of village. Outcome variables are listed on the left, and described in Appendix B. Columns 1 report estimates of a village-level treatment indicator. Column 2 reports the pure-control village mean in the ten villages that were first surveyed five years after the experiment. The unit of observation is the village. Standard errors are in parentheses and are robust. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. In column 3, we calculate the standardised difference between the two group means and divide this by the standard deviation of the variable, following Imbens and Rubin (2015). All variables are from village-level questionnaires collected five years after the experiment, except for the bottom five distance variables, which come from administrative data collected prior to the intervention (before baseline).

Table A.4: Determinants of attrition

Attrited in...	(1) Any round	(2) Any round	(3) After six months	(4) After six months	(5) After five years	(6) After five years
Treatment	0.016	0.020	-0.003	-0.006	0.024	0.030*
	(0.017)	(0.017)	(0.008)	(0.008)	(0.016)	(0.016)
Placebo	0.015	0.017	-0.005	-0.006	0.016	0.017
	(0.017)	(0.017)	(0.008)	(0.008)	(0.015)	(0.015)
% male		-0.008		-0.004		-0.002
		(0.014)		(0.007)		(0.013)
Age		0.001		0.000		0.001
		(0.001)		(0.000)		(0.001)
Years of education		0.001		0.000		0.001
		(0.003)		(0.001)		(0.003)
Single		0.007		-0.005		0.006
		(0.031)		(0.014)		(0.027)
% that watches television at least once a week		-0.011		-0.010		-0.001
		(0.023)		(0.012)		(0.020)
% that listens to radio at least once a week		-0.002		-0.011		0.001
		(0.014)		(0.008)		(0.013)
% that travels outside the village within the district at least once a week		0.003		0.015		-0.006
		(0.016)		(0.009)		(0.015)
% that travels outside the district at least once a week		-0.020		-0.002		-0.024
		(0.021)		(0.013)		(0.017)
% that ever lived outside of current village 6 months		0.011		-0.013		0.010
		(0.025)		(0.013)		(0.022)
% that ever lived outside of current district 6 months		-0.042		0.016		-0.049**
		(0.028)		(0.018)		(0.023)
Durable assets (USD PPP)		0.000		-0.000		0.000
		(0.000)		(0.000)		(0.000)
Household size		-0.018**		-0.003		-0.015**
		(0.007)		(0.004)		(0.006)
Number of individuals aged 0-6		0.024**		0.002		0.018*
		(0.011)		(0.005)		(0.010)
Number of male individuals aged 7-15		0.003		0.008		-0.007
		(0.010)		(0.006)		(0.009)
Number of female individuals aged 7-15		0.009		0.004		0.006
		(0.011)		(0.005)		(0.009)
Adult males aged above 15		0.019		-0.002		0.019
		(0.015)		(0.007)		(0.014)
Adult female aged above 15		0.002		0.003		-0.005
		(0.014)		(0.008)		(0.013)
Control mean	.08	.08	.03	.03	.06	.06
F-test p-value	.55	.36	.82	.75	.28	.03

Notes: Columns (1), (3), and (5) test whether attrition differs by treatment arm by showing coefficients from a linear regression of an indicator variable for the individual not being surveyed at any follow-up, six months after the screen, five years after the screening, respectively, on treatment and placebo indicator. Columns (2), (4), and (6) test whether attrition differs by household- and individual-level characteristics by showing coefficients from a linear regression of an indicator variable for the individual not being surveyed at any follow-up, six months after the screen, five years after the screening, respectively, on treatment and placebo indicator on baseline covariates, a treatment indicator, and a placebo indicator. If a baseline covariate is missing, we replace the missing values with the sample mean and include a missing data indicator. All regressions include village fixed effects. Standard errors are clustered at the household-level, are reported in parentheses. *; **; and *** denote significance at the 10; 5; and 1 percent levels respectively. At the bottom we report the mean attrition rate in the control group and a p-value from an F-test testing that all coefficients on the covariates reported in the column are equal to zero. The number of observations is 2,112 individuals interviewed at baseline.

Table A.5: Sample and compliance

	All groups	Treatment	Placebo	Within-village control	Pure control
Number of villages	74		64		10
Individuals:					
In sample	2434	690	717	705	322
Given tickets	2112	690	717	705	0
Compliers	2070	673	698	699	0
Non-compliers	42	17	19	6	0
of which					
At wrong screening	20	3	11	6	0
Missed screening	22	14	8	0	0
% of non-compliers	.06	.025	.026	.009	0
Households:					
In sample	1322	383	378	381	180
Given tickets	1142	383	378	381	0
Compliers	1116	371	368	377	0
Non-compliers	26	12	10	4	0
of which					
At wrong screening	11	2	5	4	0
Missed screening	15	10	5	0	0
% of non-compliers	.067	.031	.026	.01	0

Notes: Observations for individuals and households by treatment and compliance to treatment. Authors' calculations.

Table A.6: Administrative data sources

GIS object	Source	Year
Cities	1994 population census	1994
Health Centers	FAO Environment and Natural Resources Service (SDRN)	2007
Market Centers	IFPRI/FAO Environment and Natural Resources Service (SDRN)	2004
Rivers	FAO Environment and Natural Resources Service (SDRN)	2007
Roads	Woody Biomass Inventory and Strategic Planning Project (WBISPP), Ministry of Agriculture and Rural Development	2004

Table A.7: Mapping of hypothetical lotteries to risk aversion coefficients

(1) Choice	(2) Payouts	(3)	(4) Exp. value	(5) Std. dev.	(6) ΔE/ ΔSD	(7) Risk aversion	(8) S	(9) Value given
	Heads	Tails						
1	2.5	2.5	2.5	0.00	0.35	Severe	3.26 - ∞	3.260
2	2	4	3	1.41	0.35	Intermediate	1.2 - 3.26	1.978
3	1.5	5.5	3.5	2.83	0.35	Moderate	0.68 - 1.2	0.903
4	1	7	4	4.24	0.35	Slight-to-neutral	0.33 - 0.68	0.474
5	0	10	5	7.07		Neutral-to-preferred	0 - 0.33	0.165

Notes: Column 1 gives the choice number. Columns 2 and 3 give the payout options of the hypothetical lotteries. Columns 4 and 5 give the mean and variance of each lottery. The successive lotteries offered increase in both mean and variance, with payouts ordered from most to least risk averse. Column 8 shows the range of coefficient of partial risk aversion based on the chosen lottery. Following Hill, Hoddinott, and Kumar (2013), we assign the geometric mean of the endpoints of the interval for options 2-4. For option 1, since only 12 percent of individuals choose this option, we assign 3.26, whereas for option 5 we use the arithmetic mean of the points, assuming that no respondents are risk-loving.

Table A.8: Tabulation of responses to hypothetical time preference questions

			Indifferent between ETB 100 ETB in one and month and X in two months			
			Patient	Somewhat impatient	Most impatient	
			X <125	125 <X <150	150 <X	Total
Indifferent between 100 ETB now and X in one month	Patient	X <125	**602**	**45**	**38**	685
			28.9%	**2.2%**	**1.8%**	32.8%
	Somewhat impatient	125 <X <150	**61**	**91**	**116**	268
			2.9%	**4.4%**	**5.6%**	12.8%
	Most impatient	150 <X	**53**	**37**	**1,044**	1134
			2.5%	**1.8%**	**50%**	54.3%
	Total		716	173	1,198	2087
			34.3%	8.3%	57.4%	100%

Notes: Tabulation of the sample individual-level sample at baseline.
■ "Present-biased": More patient over future trade-offs than current trade-offs.
■ "Patient now and impatient later": Less patient over future trade-offs than current trade-offs.
■ "Time inconsistent" (Direction of inconsistency depends on answer to the question : "How much you would need to receive to wait one month for the payment instead of receiving 100 ETB today"). Details are provided in the Appendix section B.4.3.

Table A.9: Robustness tests for individual-level outcomes

After five years	Pre-specified		HH controls		ANCOVA+HH controls		Village controls		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Treatment	Treat. vs. placebo	Treatment	Treat. vs. placebo	Treatment	Treat. vs. placebo	Treatment	Treat. vs. placebo	Control mean (SD) Total obs.
Aspirations index	0.12**	0.15***	0.10*	0.15**	0.10	0.14**	0.14**	0.16***	0.02
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(1.00)
	[0.04]**	[0.01]***	[0.08]*	[0.01]**	[0.11]	[0.01]**	[0.02]**	[0.01]***	1956
Expectations index	0.21***	0.22***	0.18***	0.20***	0.17***	0.19***	0.22***	0.23***	0.01
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)	(1.00)
	[0.00]***	[0.00]***	[0.01]***	[0.00]***	[0.01]**	[0.00]***	[0.00]***	[0.00]***	1955
Asp. and exp. aggregate index	0.19***	0.23***	0.16***	0.21***	0.16***	0.20***	0.21***	0.24***	0.01
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(1.00)
	[0.00]***	[0.00]***	[0.01]***	[0.00]***	[0.01]**	[0.00]***	[0.00]***	[0.00]***	1956
Aspired income (USD PPP)	2174.31	3133.46*	1778.74	3046.79*	1792.07	3071.95*	2663.53	3265.30*	15475.50
	(1734.24)	(1678.24)	(1753.21)	(1699.91)	(1781.87)	(1700.32)	(1740.82)	(1741.88)	(27785.72)
	[0.28]	[0.09]*	[0.46]	[0.11]	[0.44]	[0.11]	[0.19]	[0.09]*	1940
Aspired wealth (USD PPP)	1367.43	724.36	946.70	353.55	993.32	314.42	1636.66	737.85	11926.45
	(1277.15)	(1362.92)	(1269.85)	(1339.07)	(1272.74)	(1343.91)	(1302.91)	(1394.41)	(21327.22)
	[0.28]	[0.60]	[0.46]	[0.79]	[0.44]	[0.81]	[0.21]	[0.60]	1935
Aspired education (years)	0.29*	0.42**	0.29*	0.45***	0.28*	0.40**	0.31*	0.43**	14.25
	(0.16)	(0.17)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.17)	(2.60)
	[0.22]	[0.04]**	[0.20]	[0.02]**	[0.24]	[0.04]**	[0.17]	[0.03]**	1847
Expected income (USD PPP)	259.09	182.82	158.05	120.48	173.72	164.32	296.10	205.91	3413.11
	(187.94)	(192.05)	(186.19)	(189.74)	(188.09)	(189.06)	(190.82)	(197.64)	(2827.21)
	[0.17]	[0.34]	[0.40]	[0.52]	[0.35]	[0.38]	[0.12]	[0.30]	1940
Expected wealth (USD PPP)	525.91**	367.15	381.92	259.43	380.14	275.26	610.31**	365.04	4025.57
	(246.99)	(247.56)	(247.60)	(244.29)	(249.70)	(243.54)	(253.92)	(254.94)	(3991.92)
	[0.05]*	[0.21]	[0.18]	[0.43]	[0.19]	[0.38]	[0.02]**	[0.23]	1935
Expected education (years)	0.69***	0.98***	0.69***	0.99***	0.58**	0.87***	0.70***	1.05***	12.28
	(0.26)	(0.26)	(0.26)	(0.26)	(0.27)	(0.27)	(0.26)	(0.27)	(3.92)
	[0.02]**	[0.00]***	[0.03]**	[0.00]***	[0.09]*	[0.00]***	[0.02]**	[0.00]***	1847
Best life	0.20*	0.16	0.16	0.15	0.17	0.10	0.22*	0.17	4.83
	(0.11)	(0.12)	(0.11)	(0.11)	(0.11)	(0.11)	(0.12)	(0.12)	(1.80)
	[0.16]	[0.32]	[0.30]	[0.41]	[0.26]	[0.77]	[0.12]	[0.28]	1955
Happiest life	0.08	-0.02	0.07	-0.01	0.08	0.01	0.12	0.00	6.05
	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(2.20)
	[0.54]	[0.88]	[0.63]	[0.95]	[0.56]	[0.95]	[0.39]	[0.97]	1955

Notes: OLS estimates of within-village treatment effects five years after the intervention (columns 1-8). The comparison group comprises households from the 64 treated villages that were not invited to any screening. All columns control for characteristics of the household head: age, years of education, an indicator for being single, and an indicator for being male. Columns 1-2 replicate the results in the main tables of the paper, including village fixed effects. HH controls specification (columns 3-4) adds as controls: a baseline indicator for ever having lived outside of the village in the last 6 months; baseline indicator for ever having lived outside of the district in the last 6 months; the baseline value of durable assets (excluding tools); and household size. ANCOVA+HH controls specification (columns 5-6) uses the same controls as the previous two columns and additionally controls for the baseline value of the outcome. Village controls specification (columns 7-8) controls for the set of pre-specified village-level controls as in the between-village analysis and replaces the village fixed-effects with screening fixed effects. The set of pre-specified village-level controls includes the number of inhabitants, hectares covered by forest, an indicator for whether sorghum is the main crop, costs of trip to nearest market, an indicator for whether the village has a first cycle school, percentage of households with radio, distance to the next market place, distance to the school, distance to the next farmers training centre, distance to the next health centre, distance to the next river. Column 9 display the control mean; standard deviation; and total number of observations. Heteroskedasticity-robust standard errors are clustered at the household-level in parentheses, except for the specifications in columns 7-8, which are clustered at the screening-site-level. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The conversion is described in Appendix B.1. The unit of observation is the individual respondent (household head or their spouse). The number of observations varies slightly across rows because some respondents do not answer all questions, though the indices aggregate all non-missing outcomes.

Table A.10: Robustness test for household-level outcomes

After five years	Pre-specified (1)	(2)	HH controls (3)	(4)	ANCOVA+HH controls (5)	(6)	Village controls (7)	(8)	(9)
	Treatment	Treat. vs. placebo	Treatment	Treat. vs. placebo	Treatment	Treat. vs. placebo	Treatment	Treat. vs. placebo	Control mean (SD) Total obs.
Children aged 16-20 in school	0.06*	0.06*	0.05	0.06*	0.05	0.05	0.06*	0.05	0.17
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.41)
	[0.08]*	[0.11]	[0.13]	[0.13]	[0.12]	[0.15]	[0.09]*	[0.14]	1078
Daily minutes in school for children aged 16-20	30.50**	30.00**	28.12**	28.45**	30.57**	29.31**	28.29**	27.78**	58.64
	(12.92)	(13.27)	(13.01)	(13.25)	(13.49)	(13.76)	(13.18)	(13.53)	(149.88)
	[0.04]**	[0.05]**	[0.06]*	[0.06]*	[0.05]**	[0.07]*	[0.06]*	[0.08]*	1077
Daily minutes studying for children aged 16-20	7.86*	7.27	6.93	6.94	8.33*	7.31	8.06*	7.53	17.82
	(4.52)	(4.90)	(4.54)	(4.92)	(4.69)	(5.07)	(4.68)	(5.09)	(52.12)
	[0.08]*	[0.14]	[0.13]	[0.16]	[0.10]	[0.15]	[0.09]*	[0.14]	1070
Children aged 16-20 that attained 8th grade	0.08***	0.07**	0.08***	0.07**	0.09***	0.07**	0.08***	0.07**	0.07
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.26)
	[0.01]***	[0.05]**	[0.01]**	[0.05]*	[0.01]***	[0.04]**	[0.00]***	[0.06]*	1078
Children aged 7-15 in school	0.01	0.08	0.01	0.08	0.03	0.05	0.03	0.10	1.22
	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)	(1.18)
	[0.86]	[0.26]	[0.90]	[0.28]	[0.75]	[0.55]	[0.70]	[0.17]	1078
Daily minutes in school for children aged 7-15	11.47	45.84*	10.42	46.05*	20.69	36.18	13.12	56.23**	527.12
	(25.84)	(25.31)	(26.24)	(25.66)	(30.47)	(30.70)	(26.90)	(26.02)	(437.21)
	[0.86]	[0.21]	[0.90]	[0.22]	[0.75]	[0.55]	[0.70]	[0.09]*	1068
Daily minutes studying for children aged 7-15	15.13*	9.59	14.70*	9.27	16.69*	7.14	16.57*	12.33	91.29
	(8.36)	(8.57)	(8.46)	(8.65)	(9.03)	(9.24)	(8.57)	(8.91)	(115.61)
	[0.21]	[0.26]	[0.25]	[0.28]	[0.19]	[0.55]	[0.16]	[0.17]	1069
Schooling expenditure (USD PPP)	8.20***	6.88**	7.88***	6.43**	6.83**	5.51*	8.06***	6.76**	19.17
	(2.86)	(3.06)	(2.90)	(3.11)	(2.89)	(3.00)	(2.99)	(3.26)	(32.73)
	[0.00]***	[0.02]**	[0.01]***	[0.04]**	[0.02]**	[0.07]*	[0.01]***	[0.04]**	1074
Daily minutes working	55.91**	46.45*	51.00**	42.68*	43.83*	40.81	62.90***	35.53	750.26
	(23.88)	(24.98)	(24.18)	(25.39)	(23.99)	(25.21)	(24.07)	(25.12)	(316.21)
	[0.04]**	[0.13]	[0.07]*	[0.19]	[0.14]	[0.21]	[0.02]**	[0.31]	1075
Daily minutes in leisure	0.66	35.33	-1.39	33.46	-7.88	12.97	9.11	6.32	1979.38
	(55.91)	(56.79)	(56.51)	(57.46)	(56.12)	(56.84)	(56.07)	(56.16)	(754.33)
	[0.99]	[0.53]	[0.98]	[0.56]	[0.89]	[0.82]	[0.87]	[0.91]	1076
Value of livestock (USD PPP)	184.58	309.11**	120.10	261.10**	92.06	84.21	204.54	271.09*	2018.22
	(135.92)	(130.43)	(135.46)	(125.17)	(123.18)	(113.07)	(141.58)	(138.22)	(1921.09)
	[0.17]	[0.04]**	[0.38]	[0.07]*	[0.46]	[0.79]	[0.15]	[0.10]	1080
Value of tools (USD PPP)	27.51**	15.44	17.94	3.71	16.44	3.39	31.60***	14.68	106.02
	(11.60)	(13.66)	(11.01)	(13.25)	(10.74)	(13.05)	(11.76)	(14.20)	(126.90)
	[0.04]**	[0.26]	[0.21]	[0.78]	[0.25]	[0.79]	[0.01]**	[0.30]	1077
Value of durable assets excluding tools (USD PPP)	21.87**	24.93**	21.86**	21.42**	18.51*	19.60**	22.23**	21.63*	70.55
	(10.74)	(11.18)	(10.31)	(10.43)	(9.82)	(9.95)	(11.12)	(11.59)	(127.39)
	[0.05]*	[0.05]*	[0.07]*	[0.08]*	[0.12]	[0.10]*	[0.05]**	[0.12]	1077
Value of house (USD PPP)	412.38***	350.18***	361.11***	311.35***	361.56***	287.49***	438.11***	366.55***	1384.27
	(93.87)	(93.47)	(92.89)	(91.94)	(89.77)	(88.11)	(97.87)	(97.73)	(1235.57)
	[0.00]***	[0.00]***	[0.00]***	[0.00]***	[0.00]***	[0.00]***	[0.00]***	[0.00]***	1076
Non-organic roof	0.06**	0.02	0.05	0.01	0.01	-0.01	0.08**	0.02	0.68
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.47)
	[0.05]*	[0.49]	[0.13]	[0.64]	[0.75]	[0.63]	[0.03]**	[0.49]	1087
Own toilet	0.07*	0.02	0.06	0.02	0.06	0.02	0.08**	0.03	0.38
	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.49)
	[0.05]*	[0.49]	[0.13]	[0.64]	[0.16]	[0.63]	[0.04]**	[0.49]	1088

Notes: OLS estimates of within-village treatment effects five years after the intervention (columns 1-8). The comparison group comprises households from the 64 treated villages that were not invited to any screening. All columns control for characteristics of the household head: age, years of education, an indicator for being single, and an indicator for being male. Columns 1-2 replicate the results in the main tables of the paper, including village fixed effects. HH controls specification (columns 3-4) adds as controls: a baseline indicator for ever having lived outside of the village in the last 6 months; baseline indicator for ever having lived outside of the district in the last 6 months; the baseline value of durable assets (excluding tools); and household size. ANCOVA+HH controls specification (columns 5-6) uses the same controls as the previous two columns and additionally controls for the baseline value of the outcome. Village controls specification (columns 7-8) controls for the set of pre-specified village-level controls as in the between-village analysis and replaces the village fixed-effects with screening fixed effects. The set of pre-specified village-level controls includes the number of inhabitants, hectares covered by forest, an indicator for whether sorghum is the main crop, costs of trip to nearest market, an indicator for whether the village has a first cycle school, percentage of households with radio, distance to the next market place, distance to the school, distance to the next farmers training centre, distance to the next health centre, distance to the next river. Column 9 display the control mean; standard deviation; and total number of observations. All regressions on the educational outcomes additional control for the number of children aged 0-15 at baseline to account for the baseline imbalance in the number of children, expect in columns 5-6 that already control for the baseline value of the outcome (which is highly correlated with the number of children at baseline). Heteroskedasticity-robust standard errors are in parentheses, except for the specifications in columns 7-8, which are clustered at the screening-site-level. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The conversion is described in Appendix B.1. The unit of observation is the household. The number of observations varies slightly across rows because some respondents do not answer all questions.

Table A.11: Robustness test for education outcomes after six months

After six months	Pre-specified		HH controls		ANCOVA+HH controls		Village controls		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Treatment	Treat. vs. placebo	Treatment	Treat. vs. placebo	Treatment	Treat. vs. placebo	Treatment	Treat. vs. placebo	Control mean (SD) Total obs.
Children aged 7-10 in school	0.08	0.09*	0.08	0.08	0.11**	0.05	0.07	0.10*	0.60
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.73)
	[0.34]	[0.19]	[0.37]	[0.22]	[0.10]	[0.57]	[0.50]	[0.12]	1126
Daily minutes in school for children aged 7-10	14.75	21.10	13.84	20.14	21.59	9.60	15.85	30.50*	198.10
	(16.28)	(16.29)	(16.67)	(16.53)	(16.91)	(16.96)	(17.31)	(17.44)	(250.25)
	[0.55]	[0.19]	[0.61]	[0.22]	[0.30]	[0.57]	[0.54]	[0.12]	1117
Daily minutes studying for children aged 7-10	-1.62	6.40	-2.12	6.05	0.92	3.83	-1.29	7.32	45.08
	(4.86)	(4.69)	(4.99)	(4.75)	(4.94)	(4.78)	(5.12)	(5.01)	(70.78)
	[0.74]	[0.19]	[0.67]	[0.22]	[0.85]	[0.57]	[0.80]	[0.14]	1119
Children aged 11-15 in school	0.09*	0.05	0.09*	0.04	0.07	0.03	0.09*	0.05	0.56
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.73)
	[0.11]	[0.51]	[0.12]	[0.59]	[0.24]	[0.70]	[0.11]	[0.49]	1126
Daily minutes in school for children aged 11-15	21.63	10.73	19.64	8.24	13.02	6.17	21.89	10.97	188.71
	(16.52)	(16.36)	(16.86)	(16.49)	(16.10)	(16.05)	(17.52)	(17.21)	(248.36)
	[0.19]	[0.51]	[0.24]	[0.62]	[0.42]	[0.70]	[0.21]	[0.52]	1118
Daily minutes studying for children aged 11-15	11.04*	5.99	11.00*	6.03	9.64*	7.63	12.05*	6.52	58.11
	(6.02)	(6.26)	(6.16)	(6.31)	(5.83)	(6.22)	(6.36)	(6.68)	(86.58)
	[0.11]	[0.51]	[0.12]	[0.59]	[0.24]	[0.66]	[0.11]	[0.49]	1117
Schooling expenditure (USD PPP)	9.00**	4.14	9.04**	4.12	6.35*	1.98	10.06***	5.68	37.75
	(3.68)	(4.10)	(3.74)	(4.13)	(3.60)	(4.00)	(3.89)	(4.42)	(51.39)
	[0.01]**	[0.31]	[0.01]**	[0.32]	[0.08]*	[0.62]	[0.01]***	[0.20]	1118
Children aged 16-20 in school	0.01	0.04	-0.00	0.04	-0.03	0.01	0.01	0.03	0.27
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.59)
	[0.84]	[0.78]	[0.94]	[0.82]	[0.74]	[0.94]	[0.86]	[0.89]	1126
Daily minutes in school for children aged 16-20	10.12	6.20	5.21	3.06	-2.37	-2.56	5.00	1.95	76.91
	(12.85)	(13.61)	(13.07)	(13.67)	(12.19)	(12.63)	(13.37)	(14.59)	(176.69)
	[0.84]	[0.97]	[0.94]	[0.82]	[0.85]	[0.94]	[0.86]	[0.89]	1119
Daily minutes studying for children aged 16-20	-1.37	0.12	-3.62	-1.40	-4.15	-0.46	-4.01	-1.62	36.71
	(6.11)	(6.14)	(6.23)	(6.16)	(6.04)	(5.73)	(6.40)	(6.56)	(90.28)
	[0.84]	[0.98]	[0.94]	[0.82]	[0.74]	[0.94]	[0.86]	[0.89]	1120
Children aged 7-15 in school	0.17**	0.13*	0.16**	0.12*	0.15**	0.05	0.16**	0.15**	1.16
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(1.10)
	[0.04]**	[0.08]*	[0.04]**	[0.11]	[0.08]*	[0.42]	[0.04]**	[0.05]*	1126
Daily minutes in school for children aged 7-15	52.04**	46.86**	51.74**	43.95*	44.94*	21.25	55.27**	55.99**	386.98
	(22.98)	(22.85)	(23.50)	(23.00)	(22.91)	(23.39)	(24.49)	(24.60)	(370.32)
	[0.04]**	[0.08]*	[0.04]**	[0.11]	[0.08]*	[0.42]	[0.04]**	[0.05]*	1118
Daily minutes studying for children aged 7-15	7.68	10.73	7.70	10.36	8.36	8.89	9.57	12.21	105.16
	(8.19)	(8.27)	(8.37)	(8.32)	(8.11)	(8.45)	(8.68)	(8.87)	(122.83)
	[0.35]	[0.19]	[0.36]	[0.21]	[0.30]	[0.42]	[0.27]	[0.17]	1115

Notes: OLS estimates of within-village treatment effects six months after the intervention (columns 1-8). The comparison group comprises households from the 64 treated villages that were not invited to any screening. All columns control for characteristics of the household head: age, years of education, an indicator for being single, and an indicator for being male. Columns 1-2 replicate the results in Table 5 of the paper, including village fixed effects. HH controls specification (columns 3-4) adds as controls: a baseline indicator for ever having lived outside of the village in the last 6 months; baseline indicator for ever having lived outside of the district in the last 6 months; the baseline value of durable assets (excluding tools); and household size. ANCOVA+HH controls specification (columns 5-6) uses the same controls as the previous two columns and additionally controls for the baseline value of the outcome. Village controls specification (columns 7-8) controls for the set of pre-specified village-level controls as in the between-village analysis and replaces the village fixed-effects with screening fixed effects. The set of pre-specified village-level controls includes the number of inhabitants, hectares covered by forest, an indicator for whether sorghum is the main crop, costs of trip to nearest market, an indicator for whether the village has a first cycle school, percentage of households with radio, distance to the next market place, distance to the school, distance to the next farmers training centre, distance to the next health centre, distance to the next river. Column 9 display the control mean; standard deviation; and total number of observations. All regressions on the educational outcomes additional control for the number of children aged 0-15 at baseline to account for the baseline imbalance in the number of children, expect in columns 5-6 that already control for the baseline value of the outcome (which is highly correlated with the number of children at baseline). Heteroskedasticity-robust standard errors are in parentheses, except for the specifications in columns 7-8, which are clustered at the screening-site-level. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The conversion is described in Appendix B.1. The unit of observation is the household. The number of observations varies slightly across rows because some respondents do not answer all questions.

Table A.12: Aspirations and expectations after the screening, after six months, and after five years

	After the screening				After six months				After five years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Treatment	Placebo	Treat. vs. placebo	Control mean (SD) Total obs.	Treatment	Placebo	Treat. vs. placebo	Control mean (SD) Total obs.	Treatment	Placebo	Treat. vs. placebo	Control mean (SD) Total obs.
Summary indices:												
Aspirations index	0.05	-0.05	0.10**	0.00	0.04	0.02	0.03	0.00	0.12**	-0.03	0.15***	0.02
	(0.05)	(0.05)	(0.05)	(1.00)	(0.05)	(0.05)	(0.05)	(1.01)	(0.06)	(0.06)	(0.06)	(1.00)
	[0.33]	[0.46]	[0.03]**	2005	[0.40]	[0.89]	[0.62]	2079	[0.04]**	[0.87]	[0.01]***	1956
Expectations index	0.27***	0.15**	0.12**	0.00	0.08	-0.01	0.09*	-0.00	0.21***	-0.01	0.22***	0.01
	(0.06)	(0.06)	(0.06)	(1.00)	(0.05)	(0.05)	(0.05)	(1.00)	(0.06)	(0.06)	(0.06)	(1.00)
	[0.00]***	[0.03]**	[0.03]**	2005	[0.34]	[0.89]	[0.28]	2078	[0.00]***	[0.87]	[0.00]***	1955
Aspirations and expectations aggregate index	0.16***	0.04	0.13***	0.00	0.06	0.01	0.06	0.00	0.19***	-0.04	0.23***	0.01
	(0.05)	(0.05)	(0.05)	(1.00)	(0.05)	(0.05)	(0.05)	(1.01)	(0.06)	(0.06)	(0.06)	(1.00)
	[0.00]***	[0.48]	[0.03]**	2005	[0.34]	[0.89]	[0.37]	2079	[0.00]***	[0.87]	[0.00]***	1956
Aspirations: what would you like to achieve?												
Aspired income (USD PPP)	1745.07	-2295.18	4040.24	23993.62	619.92	2269.88	-1649.95	21539.46	2194.87	-934.18	3129.06*	15460.56
	(2999.28)	(2738.13)	(2715.79)	(57202.10)	(2375.94)	(2393.56)	(2503.44)	(44863.09)	(1733.56)	(1590.25)	(1677.97)	(27766.92)
	[0.84]	[0.60]	[0.20]	1994	[0.79]	[0.54]	[0.76]	2069	[0.28]	[0.62]	[0.09]*	1941
Aspired wealth (USD PPP)	-71.02	-3425.21*	3354.18**	13717.86	-1480.46	-1303.17	-177.29	14449.17	1368.33	644.16	724.17	11922.30
	(2018.08)	(1897.83)	(1678.95)	(38805.20)	(1692.32)	(1625.51)	(1618.98)	(31089.47)	(1276.48)	(1311.75)	(1362.83)	(21311.05)
	[0.97]	[0.21]	[0.14]	1993	[0.57]	[0.54]	[0.91]	2071	[0.28]	[0.62]	[0.60]	1936
Aspired education (years)	0.18	0.07	0.11	14.12	0.30**	0.09	0.21	14.05	0.30*	-0.12	0.42**	14.24
	(0.13)	(0.13)	(0.13)	(2.39)	(0.15)	(0.15)	(0.15)	(2.61)	(0.16)	(0.17)	(0.17)	(2.60)
	[0.56]	[0.60]	[0.40]	1976	[0.12]	[0.54]	[0.45]	1989	[0.20]	[0.62]	[0.04]**	1848
Expectations: what do you expect in ten years?												
Expected income (USD PPP)	1031.59***	388.49	643.10**	4792.92	177.10	22.98	154.12	5129.90	260.37	77.79	182.57	3413.18
	(284.16)	(274.38)	(281.36)	(4895.01)	(452.91)	(427.84)	(427.09)	(8613.47)	(187.91)	(176.38)	(192.03)	(2825.03)
	[0.00]***	[0.23]	[0.03]**	1958	[0.69]	[0.96]	[0.72]	2058	[0.17]	[0.66]	[0.34]	1941
Expected wealth (USD PPP)	855.68***	145.04	710.64**	4366.51	156.33	-272.69	429.02	4752.88	527.93**	161.18	366.75	4024.70
	(277.11)	(277.03)	(285.70)	(4534.00)	(328.62)	(299.30)	(317.29)	(5781.44)	(246.82)	(244.75)	(247.55)	(3988.89)
	[0.00]***	[0.60]	[0.03]**	1965	[0.69]	[0.63]	[0.26]	2043	[0.05]**	[0.66]	[0.21]	1936
Expected education (years)	0.58***	0.49***	0.09	13.33	0.47***	0.14	0.33*	13.48	0.69***	-0.29	0.98***	12.28
	(0.19)	(0.19)	(0.16)	(3.61)	(0.17)	(0.17)	(0.17)	(3.04)	(0.26)	(0.27)	(0.26)	(3.91)
	[0.00]***	[0.03]**	[0.56]	1893	[0.02]**	[0.63]	[0.17]	1905	[0.02]**	[0.66]	[0.00]***	1848

Notes: OLS estimates of within-village treatment and placebo effects right after the video screenings (columns 1-2), after six months (columns 5-6), and after five years (columns 9-10), including pre-specified individual-level controls. Columns 3, 7, and 11 test for differences in parameters obtained in previous two columns. Column 4, 8, and 12 displays the control mean, standard deviation, and number of observations across rounds. The comparison group comprises households from the 64 treated villages that were not invited to any screening. We note that after six months we have more observations than after the screening because of logistical challenges after the screening: we could not complete the surveys with 22 individuals that missed the screening and 81 individuals that attended them but left before the end of the videos. Heteroskedasticity-robust standard errors are clustered at the household-level in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The conversion is described in Appendix B.1. The unit of observation is the individual respondent (household head or their spouse). The number of observations varies slightly across rows because some respondents do not answer all questions, though the indices aggregate all non-missing outcomes. To measure aspirations, respondents are asked the levels of outcomes the respondent would like to achieve, on different dimensions. Annual income is the amount of cash income the household earns from all agricultural and non-agricultural activities in a year. Wealth is durable wealth (including housing, vehicles, furniture and other valuable durables). Aspired education is measured as the 'years of education that you would like your oldest child to achieve'. Expectations are measured as the levels the respondent expects to reach in ten years, on the same dimensions. The aspirations and expectations indices are standardised indices (inverse-covariance-weighted averages) of these three dimensions, constructed following Anderson (2008). The outcomes are indices, standardised relative to the within-village control group. The aspirations and expectations aggregate index is constructed analogously and made of the reported income, wealth and years of education for children, for both aspirations and expectations.

Table A.13: Aspirations and expectations gaps after the screening, after six months, and after five years

	After the screening				After six months				After five years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Treatment	Placebo	Treat. vs. placebo	Control mean (SD) Total obs.	Treatment	Placebo	Treat. vs. placebo	Control mean (SD) Total obs.	Treatment	Placebo	Treat. vs. placebo	Control mean (SD) Total obs.
Summary indices of gaps:												
Aspirations (minus current at baseline) gap index	0.05	-0.05	0.11**	0.00	0.03	0.00	0.02	-0.01	0.09	-0.02	0.11**	0.01
	(0.05)	(0.05)	(0.05)	(1.00)	(0.05)	(0.05)	(0.05)	(1.00)	(0.06)	(0.06)	(0.06)	(0.98)
	[0.31]	[0.32]	[0.03]**	2005	[0.60]	[0.93]	[0.64]	2079	[0.11]	[0.94]	[0.05]**	1955
Expectations (minus current at baseline) gap index	0.28***	0.17***	0.12**	0.00	0.06	-0.02	0.08	-0.01	0.18***	-0.00	0.18***	-0.00
	(0.06)	(0.06)	(0.06)	(1.00)	(0.05)	(0.05)	(0.05)	(1.00)	(0.06)	(0.06)	(0.06)	(0.99)
	[0.00]***	[0.01]**	[0.04]**	2005	[0.60]	[0.93]	[0.34]	2076	[0.01]**	[0.94]	[0.01]***	1954
Aspirations and expectations gap aggregate index	0.19***	0.05	0.14***	0.00	0.04	-0.01	0.05	-0.01	0.15**	-0.03	0.17***	0.00
	(0.05)	(0.05)	(0.05)	(1.00)	(0.05)	(0.05)	(0.05)	(0.99)	(0.06)	(0.06)	(0.06)	(0.98)
	[0.00]***	[0.32]	[0.02]**	2005	[0.60]	[0.93]	[0.43]	2079	[0.02]**	[0.94]	[0.01]***	1955
Aspirations minus current level (at baseline)												
Aspired income gap (USD PPP)	1552.77	-2846.79	4399.56	22564.40	608.01	2150.79	-1542.78	20235.67	1819.60	-858.58	2678.18	14452.79
	(3024.02)	(2721.50)	(2705.76)	(56840.75)	(2391.76)	(2376.46)	(2493.83)	(44553.51)	(1738.09)	(1611.71)	(1679.19)	(27927.35)
	[0.91]	[0.44]	[0.16]	1968	[0.80]	[0.55]	[0.80]	2038	[0.44]	[0.70]	[0.17]	1909
Aspired wealth gap (USD PPP)	149.04	-3546.93**	3695.97**	11728.91	-1454.83	-1407.28	-47.55	12519.43	814.03	508.08	305.94	10599.78
	(1885.06)	(1753.03)	(1554.94)	(37173.77)	(1613.70)	(1539.62)	(1532.00)	(29899.89)	(1264.50)	(1310.14)	(1341.01)	(21286.73)
	[0.94]	[0.13]	[0.05]*	1972	[0.55]	[0.55]	[0.98]	2039	[0.52]	[0.70]	[0.82]	1906
Aspired education gap	0.18	0.07	0.11	12.85	0.29**	0.07	0.22	12.86	0.33**	-0.08	0.42**	12.99
	(0.13)	(0.13)	(0.13)	(2.90)	(0.15)	(0.14)	(0.15)	(3.06)	(0.16)	(0.17)	(0.17)	(3.04)
	[0.56]	[0.60]	[0.40]	1976	[0.14]	[0.62]	[0.41]	1989	[0.13]	[0.70]	[0.04]**	1848
Expectations minus current level (at baseline)												
Expected income gap (USD PPP)	986.26***	415.86	570.40**	3674.11	138.51	77.13	61.38	4018.26	152.84	58.18	94.66	2349.07
	(276.66)	(263.58)	(275.14)	(4698.67)	(454.41)	(427.75)	(428.01)	(8612.38)	(185.76)	(169.83)	(185.14)	(2763.83)
	[0.00]***	[0.17]	[0.06]*	1936	[0.76]	[0.86]	[0.88]	2027	[0.41]	[0.73]	[0.61]	1911
Expected wealth gap (USD PPP)	763.59***	198.13	565.46**	2948.10	115.79	-303.06	418.85	3261.21	394.31	137.48	256.83	2637.00
	(252.98)	(252.78)	(269.74)	(4119.79)	(304.50)	(277.74)	(292.38)	(5352.84)	(251.05)	(250.84)	(254.10)	(3945.80)
	[0.00]***	[0.43]	[0.06]*	1951	[0.76]	[0.73]	[0.23]	2017	[0.17]	[0.73]	[0.47]	1906
Expected education gap	0.58***	0.49***	0.09	12.06	0.45***	0.12	0.33*	12.30	0.73***	-0.25	0.98***	11.03
	(0.19)	(0.19)	(0.16)	(3.92)	(0.17)	(0.17)	(0.17)	(3.50)	(0.26)	(0.27)	(0.27)	(4.25)
	[0.00]***	[0.03]**	[0.56]	1893	[0.02]**	[0.73]	[0.16]	1905	[0.02]**	[0.73]	[0.00]***	1848

Notes: OLS estimates of within-village treatment and placebo effects right after the video screenings (columns 1-2), after six months (columns 5-6), and after five years (columns 9-10), including pre-specified individual-level controls. Columns 3, 7, and 11 test for differences in parameters obtained in previous two columns. Column 4, 8, and 12 displays the control mean, standard deviation, and number of observations across rounds. The comparison group comprises households from the 64 treated villages that were not invited to any screening. We note that after six months we have more observations than after the screening because of logistical challenges after the screening: we could not complete the surveys with 22 individuals that missed the screening and 81 individuals that attended them but left before the end of the videos. Heteroskedasticity-robust standard errors are clustered at the household-level in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The conversion is described in Appendix B.1. The unit of observation is the individual respondent (household head or their spouse). The number of observations varies slightly across rows because some respondents do not answer all questions, though the indices aggregate all non-missing outcomes. The aspirations (or expectations) gaps take the measure of aspirations (or expectations) and subtract the current level at baseline elicited for that same dimension. To measure aspirations, respondents are asked the levels of outcomes the respondent would like to achieve, on different dimensions. Annual income is the amount of cash income the household earns from all agricultural and non-agricultural activities in a year. Wealth is durable wealth (including housing, vehicles, furniture and other valuable durables). Aspired education is measured as the 'years of education that you would like your oldest child to achieve'. Expectations are measured as the levels the respondent expects to reach in ten years, on the same dimensions. For the current level of education, we use the respondents' own education level. The aspirations and expectations gap indices are standardised indices (inverse-covariance-weighted averages) of these three dimensions, constructed following Anderson (2008). The outcomes are indices, standardised relative to the within-village control group. The aspirations and expectations gap aggregate index is constructed analogously and made of the reported gaps in income, wealth and years of education for the respondents' oldest child, for both aspirations and expectations.

Table A.14: Heterogeneous treatment effects on summary indices after five years by terciles of durable assets

After five years	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Medium	High	Treat.# Low	Treat.# Med.	Treat.# High	(5)-(3)	(5)-(4)	(4)-(3)	Mean (SD) Total obs.
Agricultural investment index	0.13	0.30***	0.16	0.21**	0.12	-0.04	-0.09	0.05	0.00
	(0.09)	(0.09)	(0.11)	(0.11)	(0.11)	(0.15)	(0.15)	(0.14)	(1.00)
	[0.20]	[0.00]***	[0.47]	[0.05]*	[0.34]	[0.85]	[0.68]	[0.85]	1061
Educational investment index	0.09	0.22**	0.09	0.31**	0.22**	0.13	-0.09	0.22	0.00
	(0.08)	(0.09)	(0.10)	(0.12)	(0.11)	(0.15)	(0.16)	(0.15)	(1.00)
	[0.34]	[0.02]**	[0.57]	[0.02]**	[0.16]	[0.82]	[0.68]	[0.32]	1061
Welfare index	0.24**	0.51***	-0.01	-0.01	0.17	0.18	0.18	-0.00	0.00
	(0.10)	(0.10)	(0.12)	(0.12)	(0.13)	(0.17)	(0.17)	(0.16)	(1.00)
	[0.04]**	[0.00]***	[0.92]	[0.91]	[0.34]	[0.82]	[0.50]	[0.99]	1063
Aspiration index	0.07	0.12	-0.01	0.26***	0.08	0.09	-0.18	0.27*	0.02
	(0.08)	(0.08)	(0.10)	(0.10)	(0.09)	(0.13)	(0.13)	(0.14)	(1.00)
	[0.36]	[0.13]	[0.92]	[0.02]**	[0.36]	[0.82]	[0.50]	[0.26]	1901
Expectations index	0.13*	0.37***	0.15*	0.27***	0.12	-0.03	-0.15	0.12	0.01
	(0.08)	(0.08)	(0.09)	(0.10)	(0.10)	(0.13)	(0.14)	(0.13)	(1.00)
	[0.20]	[0.00]***	[0.47]	[0.02]**	[0.34]	[0.85]	[0.50]	[0.49]	1900
Asp. and exp. aggregate index	0.11	0.27***	0.08	0.32***	0.11	0.03	-0.21	0.23*	0.01
	(0.08)	(0.08)	(0.10)	(0.09)	(0.10)	(0.13)	(0.13)	(0.13)	(1.00)
	[0.20]	[0.00]***	[0.57]	[0.01]***	[0.34]	[0.85]	[0.50]	[0.26]	1901
Omnibus index	0.22***	0.48***	0.11	0.25**	0.27**	0.17	0.02	0.14	0.00
	(0.08)	(0.09)	(0.10)	(0.11)	(0.12)	(0.15)	(0.15)	(0.14)	(1.00)
	[0.04]**	[0.00]***	[0.57]	[0.03]**	[0.13]	[0.82]	[0.88]	[0.49]	1064

Notes: OLS estimates of within-village heterogeneous treatment effect after five years. Columns 3 to 5 report the coefficients from interacting the treatment indicator with an indicator for each of the three baseline value of durable assets excluding tools (USD PPP) per adult equivalent (a proxy for wealth) terciles (where low medium and high value of wealth refers to individuals or households who were in the bottom or middle or highest terciles at baseline). The omitted category represents individuals or households in the within-village control group from the lowest tercile of value of durable assets. The comparison group comprises households from the 64 treated villages that were not invited to any screening. Column 6 reports the difference between the estimates reported in Columns 5 and 3; Column 7 reports the difference between the estimates reported in Columns 5 and 4; Column 8 reports the difference between the estimates reported in Columns 4 and 3. Column 9 displays the control mean, standard deviation, and total number of observations. All columns control for village fixed effects and characteristics of the household head: age, years of education, an indicator for being single, and an indicator for being male. Regressions on the educational investment index additional control for the number of children aged 0-15 currently in the household to account for the baseline imbalance in the number of children. Heteroskedasticity-robust standard errors are in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. The unit of observation is the household, except for the aspirations and expectations indices (which are are observed for both household head and their spouse). The number of observations varies slightly across rows because some respondents do not answer all questions, though the indices aggregate all non-missing outcomes. The outcomes, described in Table 7, are inverse-covariance-weighted averages standardised relative to the within-village control group, following Anderson (2008).

Table A.15: Summary indices in spillover analysis with saturation design

After five years	(1) Treatment #Treatment-intense	(2) Treatment #Placebo-intense	(3) Control #Treatment-intense	(4) Control #Placebo-intense	(5) Placebo #Treatment-intense	(6) Placebo #Placebo-intense	(7) Treat #Treat-intense vs. Treate #Placebo-intense	(8) Control #Treat-intense vs. Control #Placebo-intense	(9) Placebo #Treat-intense vs. Placebo #Placebo-intense	(10) Pure Control mean (SD) Total obs.
Agricultural investment index	0.17**	0.21**	-0.07	-0.13	0.12	0.04	-0.05	0.07	0.08	-0.00
	(0.07)	(0.09)	(0.11)	(0.10)	(0.14)	(0.07)	(0.11)	(0.07)	(0.17)	(1.00)
	[0.02]**	[0.06]*	[0.62]	[0.65]	[0.90]	[0.85]	[0.94]	[0.97]	[0.82]	1223
Educational investment index	0.22***	0.17*	-0.02	-0.01	-0.06	0.02	0.04	-0.01	-0.08	0.00
	(0.08)	(0.10)	(0.07)	(0.07)	(0.11)	(0.09)	(0.13)	(0.07)	(0.17)	(1.00)
	[0.02]**	[0.11]	[0.77]	[0.83]	[0.92]	[0.85]	[0.94]	[0.97]	[0.82]	1219
Welfare index	0.04	0.09	0.12	0.17	0.02	-0.09	-0.05	-0.05	0.11	0.00
	(0.09)	(0.09)	(0.11)	(0.11)	(0.16)	(0.13)	(0.13)	(0.11)	(0.24)	(1.00)
	[0.67]	[0.28]	[0.60]	[0.65]	[0.92]	[0.85]	[0.94]	[0.97]	[0.82]	1224
Aspiration index	0.12*	0.14	0.10	0.08	-0.09	-0.03	-0.02	0.02	-0.06	0.00
	(0.07)	(0.11)	(0.08)	(0.09)	(0.11)	(0.10)	(0.13)	(0.09)	(0.17)	(1.00)
	[0.10]*	[0.23]	[0.60]	[0.70]	[0.90]	[0.85]	[0.94]	[0.97]	[0.82]	2231
Expectations index	0.23**	0.19**	0.07	0.03	0.11	-0.10	0.04	0.04	0.21	-0.00
	(0.11)	(0.09)	(0.09)	(0.09)	(0.11)	(0.09)	(0.15)	(0.09)	(0.17)	(1.00)
	[0.07]*	[0.10]	[0.60]	[0.83]	[0.90]	[0.85]	[0.94]	[0.97]	[0.82]	2230
Asp. and exp. aggregate index	0.18**	0.18*	0.06	0.05	0.04	-0.09	0.01	0.01	0.14	-0.00
	(0.07)	(0.10)	(0.07)	(0.09)	(0.10)	(0.09)	(0.12)	(0.08)	(0.16)	(1.00)
	[0.02]**	[0.11]	[0.60]	[0.78]	[0.92]	[0.85]	[0.94]	[0.97]	[0.82]	2231
Omnibus index	0.25***	0.27***	0.07	0.07	-0.03	-0.03	-0.02	-0.00	0.01	0.00
	(0.08)	(0.09)	(0.09)	(0.09)	(0.12)	(0.09)	(0.12)	(0.08)	(0.17)	(1.00)
	[0.01]***	[0.02]**	[0.60]	[0.70]	[0.92]	[0.85]	[0.94]	[0.97]	[0.96]	1225

Notes: OLS estimates of between-village effects five years after the intervention (columns 1-6). Column 7 tests for differences in parameters obtained in first two columns. Column 8 tests for differences in parameters obtained in columns third and fourth. Column 9 tests for differences in parameters obtained in fifth and sixth columns. The comparison group comprises households from the ten pure-control villages that were first surveyed five years after the intervention. Column 10 displays the mean, standard deviation for the pure-control group, and total number of observations. All regressions control for screening-site fixed effects, individual characteristics of the respondent (age, years of education, an indicator for being single, and an indicator for being male) and village-level controls (the number of inhabitants, hectares covered by forest, an indicator for whether sorghum is the main crop, costs of trip to nearest market, an indicator for whether the village has a first cycle school, percentage of households with radio, distance to the next market place, distance to the school, distance to the next farmers training centre, distance to the next health centre, distance to the next river). Regressions on the educational investment index additional control for the number of children aged 0-15 currently in the household to account for the baseline imbalance in the number of children. Heteroskedasticity-robust standard errors are clustered at the village-level and are in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. The unit of observation is the household, except for the aspirations and expectations indices (which are are observed for both household head and their spouse). The number of observations varies slightly across rows because some respondents do not answer all questions, though the indices aggregate all non-missing outcomes. The outcomes are inverse-covariance-weighted averages standardised relative to the pure-control group, following Anderson (2008). The agricultural investment index includes all outcomes reported in Table 2, with daily minutes in leisure being recoded to be negative. The educational investment index includes all outcomes reported in Table 3. The welfare index includes all outcomes reported in Table 4, with months of food insecurity in the last year and consumption of sin goods recoded to be negative. The welfare index averages over the household head's subjective well-being outcomes. The aspirations and expectations aggregate index is made of the reported income, wealth and years of education for children, for aspirations and expectations. The omnibus index aggregates the agricultural investment, educational investment, welfare, and aspirations and expectations aggregate standardised indices into a single index, following Bessone et al. (2021) and Kling, Liebman, and Katz (2007). As the omnibus index is for the whole household, we use the household head's aspirations and expectations aggregate index.

Table A.16: Spillover analysis allowing for between-village interactions

After five years	(1) Treatment	(2) Placebo	(3) Control	(4) Treated households within 1km	(5) Villages within 1km	(6) Treat. vs. placebo	(7) Treat. vs. control	(8) Pure Control mean (SD) Total obs.
Agricultural investment index	0.09	-0.08	-0.10	0.00	-0.10**	0.16**	0.19***	-0.00
	(0.11)	(0.09)	(0.10)	(0.00)	(0.05)	(0.07)	(0.05)	(1.00)
	[0.45]	[0.71]	[0.44]	[0.68]	[0.09]*	[0.02]**	[0.00]***	1223
Educational investment index	0.06	-0.10	-0.08	-0.00	-0.04	0.15**	0.13**	0.00
	(0.07)	(0.07)	(0.06)	(0.00)	(0.04)	(0.06)	(0.05)	(1.00)
	[0.45]	[0.38]	[0.42]	[0.68]	[0.32]	[0.02]**	[0.02]**	1219
Welfare index	0.30***	0.24**	0.24**	-0.01	-0.09	0.06	0.06	0.00
	(0.11)	(0.10)	(0.12)	(0.00)	(0.06)	(0.08)	(0.07)	(1.00)
	[0.01]***	[0.12]	[0.23]	[0.68]	[0.23]	[0.40]	[0.36]	1224
Aspiration index	0.21***	0.04	0.07	0.00	-0.02	0.17***	0.13**	0.00
	(0.08)	(0.07)	(0.07)	(0.00)	(0.04)	(0.05)	(0.06)	(1.00)
	[0.01]***	[0.77]	[0.44]	[0.68]	[0.72]	[0.00]***	[0.04]**	2231
Expectations index	0.28***	0.04	0.07	0.00	-0.10**	0.24***	0.21***	-0.00
	(0.09)	(0.08)	(0.08)	(0.00)	(0.05)	(0.06)	(0.07)	(1.00)
	[0.00]***	[0.77]	[0.44]	[0.96]	[0.09]*	[0.00]***	[0.00]***	2230
Asp. and exp. aggregate index	0.23***	0.00	0.05	0.00	-0.06	0.23***	0.18***	-0.00
	(0.07)	(0.07)	(0.07)	(0.00)	(0.04)	(0.05)	(0.06)	(1.00)
	[0.00]***	[0.97]	[0.44]	[0.68]	[0.23]	[0.00]***	[0.00]***	2231
Omnibus index	0.39***	0.12	0.13	-0.00	-0.09**	0.27***	0.26***	0.00
	(0.10)	(0.08)	(0.08)	(0.00)	(0.05)	(0.07)	(0.06)	(1.00)
	[0.00]***	[0.38]	[0.38]	[0.68]	[0.09]*	[0.00]***	[0.00]***	1225

Notes: OLS estimates of between-village effects five years after the intervention, controlling for exogenous spatial treatment intensity. Each row represents a separate regression. Column 1 report estimates on household-level indicators for treatment assignment. Column 4 reports estimates of the coefficient δ_2^d from equation (6) that calculate the effect of every additional household invited to the intervention within a radius of 0-1km of the observation. The radius of 0-1km was selected after running a series of nested models as in Egger et al. (2022), selecting the model that minimised the Bayesian Information Criterion across all models for each outcome. Column 6 tests for differences in parameters obtained in first two columns. Column 7 tests for differences in parameters obtained in first and third columns. The comparison group comprises households from the ten pure-control villages that were first surveyed five years after the intervention. Column 8 displays the mean, standard deviation for the pure-control group, and total number of observations. All regressions control for screening-site fixed effects, individual characteristics of the respondent (age, years of education, an indicator for being single, and an indicator for being male) and village-level controls (the number of inhabitants, hectares covered by forest, an indicator for whether sorghum is the main crop, costs of trip to nearest market, an indicator for whether the village has a first cycle school, percentage of households with radio, distance to the next market place, distance to the school, distance to the next farmers training centre, distance to the next health centre, distance to the next river). Regressions on the educational investment index additional control for the number of children aged 0-15 currently in the household to account for the baseline imbalance in the number of children. Conley (1999) standard errors are in parentheses, accounting for spatial correlation within a 1km radius. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. The unit of observation is the household, except for the aspirations and expectations indices (which are are observed for both household head and their spouse). The number of observations varies slightly across rows because some respondents do not answer all questions, though the indices aggregate all non-missing outcomes. The outcomes are inverse-covariance-weighted averages standardised relative to the pure-control group, following Anderson (2008). The agricultural investment index includes all outcomes reported in Table 2, with daily minutes in leisure being recoded to be negative. The educational investment index includes all outcomes reported in Table 3. The welfare index includes all outcomes reported in Table 4, with months of food insecurity in the last year and consumption of sin goods recoded to be negative. The welfare index averages over the household head's subjective well-being outcomes. The aspirations and expectations aggregate index is made of the reported income, wealth and years of education for children, for aspirations and expectations. The omnibus index aggregates the agricultural investment, educational investment, welfare, and aspirations and expectations aggregate standardised indices into a single index, following Bessone et al. (2021) and Kling, Liebman, and Katz (2007). As the omnibus index is for the whole household, we use the household head's aspirations and expectations aggregate index.

Table A.17: Savings and credit

	After six months				After five years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Treatment	Placebo	Treat. vs. placebo	Control mean (SD) Total obs.	Treatment	Placebo	Treat. vs. placebo	Control mean (SD) Total obs.
Has any savings	0.05*	0.01	0.04	0.39	-0.01	-0.01	-0.00	0.20
	(0.03)	(0.02)	(0.03)	(0.49)	(0.02)	(0.02)	(0.02)	(0.40)
	[0.11]	[0.74]	[0.34]	2064	[0.99]	[0.76]	[1.00]	1949
Has outside savings	0.04*	0.02	0.02	0.26	-0.01	-0.01	0.00	0.19
	(0.02)	(0.02)	(0.02)	(0.44)	(0.02)	(0.02)	(0.02)	(0.40)
	[0.11]	[0.59]	[0.73]	2064	[0.99]	[0.76]	[1.00]	1949
Has any credit	0.04	-0.01	0.05*	0.34	-0.03	0.03	-0.06**	0.33
	(0.03)	(0.03)	(0.03)	(0.47)	(0.03)	(0.03)	(0.03)	(0.47)
	[0.16]	[0.80]	[0.34]	2064	[0.99]	[0.76]	[0.44]	1909
Has any agricultural credit					0.02***	0.01*	0.01	0.00
					(0.01)	(0.01)	(0.01)	(0.07)
					[0.06]*	[0.34]	[0.85]	1908
Total savings (USD PPP)	18.75***	4.35	14.40*	24.37	0.03	1.67	-1.64	17.37
	(6.92)	(5.40)	(7.57)	(80.68)	(3.74)	(3.69)	(3.85)	(63.24)
	[0.06]*	[0.63]	[0.34]	2026	[0.99]	[0.76]	[0.96]	1930
Total outside savings (USD PPP)	2.54*	1.66	0.88	8.88	0.06	2.54	-2.48	16.05
	(1.49)	(1.47)	(1.61)	(25.07)	(3.55)	(3.57)	(3.78)	(58.72)
	[0.13]	[0.59]	[0.73]	2030	[0.99]	[0.76]	[0.85]	1930
Credit amount (USD PPP)	4.71**	1.13	3.58	19.65	-8.39*	-0.33	-8.06	39.65
	(2.33)	(2.31)	(2.51)	(40.09)	(5.03)	(5.16)	(5.05)	(87.98)
	[0.11]	[0.74]	[0.34]	2044	[0.48]	[0.95]	[0.44]	1897
Hypothetical loan (1 year, USD PPP)	127.59	248.59	-121.00	2461.23	-48.37	-263.98*	215.62	1606.71
	(247.10)	(249.74)	(267.40)	(3255.10)	(146.01)	(135.38)	(142.52)	(1962.02)
	[0.68]	[0.59]	[0.73]	2051	[0.99]	[0.34]	[0.44]	1915
Hypothetical loan (5 years, USD PPP)	244.04	813.84	-569.80	6022.56	31.86	-287.59	319.45	3084.38
	(776.61)	(834.85)	(835.04)	(10933.82)	(251.83)	(234.17)	(240.09)	(3450.85)
	[0.75]	[0.59]	[0.73]	2060	[0.99]	[0.73]	[0.46]	1902
Hypothetical loan (10 years, USD PPP)	4338.36**	4002.98**	335.38	9865.82	-420.01	-397.06	-22.94	5797.16
	(1836.44)	(1711.77)	(2191.28)	(16200.39)	(575.35)	(599.44)	(572.63)	(7844.57)
	[0.08]*	[0.17]	[0.88]	2060	[0.99]	[0.76]	[1.00]	1807

Notes: OLS estimates of within-village treatment and placebo effects after six months (columns 1-2) and after five years (columns 5-6) of the intervention. Columns 3 and 7 test for differences in parameters obtained in previous two columns. The comparison group comprises households from the 64 treated villages that were not invited to any screening. Column 4 and 8 display the control mean, standard deviation, and total number of observations. Heteroskedasticity-robust standard errors are clustered at the household-level in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. The unit of observation is the individual respondent (household head or their spouse), except for information and fertiliser beliefs indices (which are at the household-level and were only measured after five years of the intervention). The number of observations varies slightly across rows because some respondents do not answer all questions.

Table A.18: Revenue

After five years	(1) Treatment	(2) Placebo	(3) Treat. vs. placebo	(4) Control mean (SD) Total obs.
Gross revenue (USD PPP)	104.37	-56.11	160.48*	1468.82
	(95.87)	(94.65)	(94.90)	(1273.97)
	[0.66]	[0.62]	[0.55]	1061
Revenue from crop production (USD PPP)	21.93	21.16	0.77	383.70
	(22.43)	(22.56)	(23.97)	(300.60)
	[0.66]	[0.62]	[0.97]	1077
Revenue from livestock rearing and produce (USD PPP)	-3.96	-99.66	95.70	740.53
	(73.81)	(71.95)	(71.74)	(1002.83)
	[0.96]	[0.62]	[0.55]	1087
Revenue from on- and off-farm (USD PPP)	-3.06	6.08	-9.14	25.86
	(8.04)	(9.18)	(9.07)	(111.32)
	[0.84]	[0.62]	[0.63]	1080
Revenue from non-farm enterprises (USD PPP)	21.34	14.93	6.41	159.94
	(28.79)	(29.86)	(31.79)	(353.37)
	[0.69]	[0.62]	[0.97]	1076
Transfers and remittances (USD PPP)	18.62	7.26	11.36	114.03
	(16.86)	(14.76)	(18.06)	(180.99)
	[0.66]	[0.62]	[0.79]	1079

Notes: OLS estimates of within-village treatment and placebo effects five years after the intervention (columns 1-2). Column 3 tests for differences in parameters obtained in first two columns. The comparison group comprises households from the 64 treated villages that were not invited to any screening. Column 4 displays the control mean, standard deviation, and total number of observations. All columns control for village fixed effects and characteristics of the household head: age, years of education, an indicator for being single, and an indicator for being male. Heteroskedasticity-robust standard errors are in parentheses. Stars on the coefficient estimates reflect unadjusted p-values. Minimum q-values are in square brackets and are calculated over each panel of variables. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. Outcome variables are listed on the left, and described in detail in Appendix B. All monetary values are in PPP-adjusted USD, set at 2016 (endline) prices and deflated using national non-food CPI. In 2016, USD 1 = 8.67 ETB (Ethiopian birr) PPP. The unit of observation is the household. The number of observations varies slightly across rows because some respondents do not answer all questions. Livestock revenue includes own-consumption of animals, valued at sales prices.

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 (3): 1121–1162.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutstrom. 2008. "Eliciting Risk and Time Preferences." *Econometrica* 76 (3): 583–618.

Anderson, M. L. 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 (484): 1481–1495.

Andrews, M., M. Nord, G. Bickel, and S. Carlson. 2000. "Household Food Security in the United States, 1999." *Food Assistance and Nutrition Research Report, United States Department of Agriculture* .

Ashraf, N., D. S. Karlan, and W. Yin. 2006. "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." *Quarterly Journal of Economics* 121 (2): 635–672.

Bandura, A., C. Barbaranelli, G. V. Caprara, and C. Pastorelli. 2001. "Self-Efficacy Beliefs as Shapers of Children's Aspirations and Career Trajectories." *Child Development* 72 (1): 187–206.

Bandura, A. and E. A. Locke. 2003. "Negative Self-Efficacy and Goal Effects Revisited." *Journal of Applied Psychology* 88 (1): 87–99.

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 (6068): 582–586.

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 (1): 3–18.

Benjamini, Y., A. M. Krieger, and D. Yekutieli. 2006. "Adaptive Linear Step-up Procedures that Control the False Discovery Rate." *Biometrika* 93 (3): 491–507.

Bernard, T. and A. S. Taffesse. 2014. "Aspirations: An Approach to Measurement with Validation Using Ethiopian Data." *Journal of African Economies* 23 (2): 189–224.

Bessone, P., G. Rao, F. Schilbach, H. Schofield, and M. Toma. 2021. "The Economic Consequences of Increasing Sleep Among the Urban Poor." *The Quarterly Journal of Economics* 136 (3): 1887–1941.

Bickel, G. W., M. Nord, C. Price, W. Hamilton, and J. Cook. 2000. *Guide to Measuring Household Food Security.* Alexandria, Virginia: Food and Nutrition Service, United States Department of Agriculture.

Binswanger, H. P. 1980. "Attitudes Toward Risk: Experimental Measurement in Rural India." *American Journal of Agricultural Economics* 62 (3): 395–407.

Cantril, H. 1966. *The Pattern of Human Concern.* New Brunswick, New Jersey: Rutgers University Press.

Cheng, C., S.-F. Cheung, J. H.-M. Chio, and M.-P. S. Chan. 2013. "Cultural Meaning of Perceived Control: A Meta-Analysis of Locus of Control and Psychological Symptoms Across 18 Cultural Regions." *Psychological Bulletin* 139 (1): 152–188.

Conley, T. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92 (1): 1–45.

Dean, M. and A. Sautmann. 2021. "Credit Constraints and the Measurement of Time Preferences." *The Review of Economics and Statistics* 103 (1): 119–135.

Egger, D., J. Haushofer, E. Miguel, P. Niehaus, and M. W. Walker. 2022. "General Equilibrium Effects of Cash Transfers: Experimental Evidence from Kenya." *Econometrica* 90 (6): 2603–2643.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde. 2018. "Global Evidence on Economic Preferences." *Quarterly Journal of Economics* 133 (4): 1645–1692.

Feagin, J. R. 1972. "Poverty: We Still Believe that God Helps Those Who Help Themselves." *Psychology Today* 6: 101–129.

———. 1975. *Subordinating the Poor: Welfare and American Beliefs*. Englewood Cliffs, New Jersey: Prentice Hall.

Hadley, C., D. Lindstrom, F. Tessema, and T. Belachew. 2008. "Gender Bias in the Food Insecurity Experience of Ethiopian Adolescents." *Social Science & Medicine* 66 (2): 427–438.

Harrison, G. W., S. J. Humphrey, and A. Verschoor. 2010. "Choice under Uncertainty: Evidence from Ethiopia, India and Uganda." *The Economic Journal* 120 (543): 80–104.

Hill, R. V., J. Hoddinott, and N. Kumar. 2013. "Adoption of Weather-Index Insurance: Learning from Willingness to Pay Among a Panel of Households in Rural Ethiopia." *Agricultural Economics* 44 (4-5): 385–398.

Holt, C. A. and S. K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review* 92 (5): 1644–1655.

Imbens, G. W. and D. B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.

Kling, J. R., J. B. Liebman, and L. F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Lamping, D. L., S. Schroter, P. Marquis, A. Marrelb, I. Duprat-Lomon, and P.-P. Sagnier. 2002. "The Community-Acquired Pneumonia Symptom Questionnaire: a New, Patient-Based Outcome Measure to Evaluate Symptoms in Patients with Community-Acquired Pneumonia." *Chest* 122 (3): 920–929.

Levenson, H. 1974. "Activism and Powerful Others: Distinctions Within the Concept of Internal-External Control." *Journal of Personality Assessment* 38 (4): 377–383.

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

Madden, G. J., B. R. Raiff, C. H. Lagorio, A. M. Begotka, A. M. Mueller, D. J. Hehli, and A. A. Wegener. 2004. "Delay Discounting of Potentially Real and Hypothetical

Rewards: II. Between-And Within-Subject Comparisons." *Experimental and Clinical Psychopharmacology* 12 (4): 251–261.

Reimanis, G. and C. F. Posen. 1980. "Locus of Control and Anomie in Western and African Cultures." *Journal of Social Psychology* 112: 181–189.

Rossier, J., D. Dahourou, and R. R. McCrae. 2005. "Structural and Mean-Level Analyses of the Five-Factor Model and Locus of Control: Further Evidence from Africa." *Journal of Cross-Cultural Psychology* 36: 227–246.

Tversky, A. and D. Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Biases." *Science* 185: 1124–1131.

Ubfal, D. 2016. "How General Are Time Preferences? Eliciting Good-Specific Discount Rates." *Journal of Development Economics* 118: 150–170.

Van Haaften, E. H. and F. J. R. van de Vijver. 1999. "Dealing with Extreme Environmental Degradation: Stress and Marginalization of Sahel Dwellers." *Social Psychiatry and Psychiatric Epidemiology* 34: 376–382.