Essays on rural economic development

Giulio Schinaia

Balliol College,

University of Oxford

A THESIS PRESENTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN ECONOMICS.

TRINITY 2023

Abstract

This thesis seeks to broaden our knowledge on different policies to foster rural economic development. It comprises three distinct chapters that study how we can leverage digital technologies to improve smallholder farmers' economic outcomes and the effects of a social protection programme on rural labour markets.

The first chapter studies whether, and how, providing market information to smallholder farmers can increase their revenue. I study this question leveraging a two-level cluster randomised controlled trial among 1988 cashew producers in 290 villages in Guinea-Bissau. Treated producers received free weekly messages to their mobiles during the trading seasons in 2020 and 2021. The messages provided up-to-date market news, farmgate prices, and sales advice. Treated producers reported higher prices, mostly during the 2021 season. Treated producers sold their cashews more frequently relative to the other producers, who tend to sell their cashews in a single transaction. The chapter explores several mechanisms to unpack the results, finding evidence consistent with the information increasing the bargaining power of treated producers, who negotiate better deals.

The second chapter studies the effects of a large social protection programme in rural Ethiopia on local labour markets. The programme targeted food-insecure households to provide them with food or cash transfers, as compensation for public works participation or unconditionally. Using repeated cross-sections of the Ethiopian Labour Force Survey, I show that workers shifted from agricultural to non-agricultural self-employment. I also find that the programme did not change employment rates or wages in this rural economy. I find similar results complementing my analysis with data from the Ethiopian Socio-Economic Survey.

The third chapter tests whether trying to change how people living in poverty perceive their future opportunities alters their aspirations and, through that, modifies their investment decisions. In a randomised controlled trial in remote rural Ethiopia, a treatment group of households was exposed to video documentaries of people from similar communities who improved their economic conditions through their hard work, and, as such, can serve as role models. Five years after the screening took place, the treated household that watched the videos increased their future-oriented investments in agriculture and towards their children's education. The results can be explained by an increase in aspirations in terms of lifetime goals.

Acknowledgements

First and foremost, I would like to express my immense gratitude to my supervisor, Pramila Krishnan, for all her patience and for guiding me through the research process with her broad and insightful vision of the researcher I can only aspire to be. I am also extremely grateful to Stefan Dercon and Simon Quinn for having given their advice at crucial points throughout the process of writing this thesis. I owe Stefano Caria and Doug Gollin for having encouraged me to pursue this DPhil.

The Luca D'Agliano Scholarship, the Grand Challenges Research Fund, and the bursaries of the Department of Economics have partly funded my studies. Balliol College, St. Antony's College, and the George Webb Medley Fund have financially supported travel costs. Each chapter acknowledges additional intellectual and financial support.

Many thanks to the Applied Microeconomics Research Group faculty and the members of the Centre for The Study of African Economies that have offered their time to support me in the last four years. *Nmisti tambi fala obrigadu pa tudu equipa di* Bissau Economics Lab and thanks to all the team members at the Mind and Behaviour Research Group for having always given me an additional intellectual and social home.

I am thankful to my co-authors for their continued support, encouragement, and invaluable contributions. I look forward to learning and discovering more with them.

I am deeply indebted to all participants in the studies in this thesis for all their time and patience.

For all the tips, advice, peer support, shared meals, fun, runs, and games, I am lucky to have had a fantastic PhD cohort (broadly defined). For all the hospitality and special moments shared in the last four years *merci à tous mes amis*.

Non ci sono parole per ringraziare i miei genitori per tutto il loro sostegno e affetto, cosí come per il clan Pavone-Capitani...per ora, forse, ho finito de studiá ma non de impará.

Last but opposite of least, *Sama Soppe, Pizzetta*, I am fortunate to have been sharing this journey with you. I could not dream of a more supportive, generous, reassuring, and loving partner. Thank you for teaching me so much everyday, I am thrilled to continue writing future life chapters together.

Dedico questa tesi alle mie nonne: Clara, Saia, e Lisetta.

Contents

List of Tables

List of Figures

Introduction

Most individuals living in poverty are farmers in low-income countries. Having a better understanding of the importance of various constraints hindering their ability to enhance their economic conditions remains a key area of policy and research interest. This thesis seeks to broaden our knowledge on different policies to foster rural economic development. Specifically, this thesis studies how we can leverage digital technologies to improve smallholder farmers' economic outcomes and the effects of a social protection programme on rural labour markets. Below, I provide a summary of each chapter and sketch their main contribution.

In the first chapter ("*Impact of Market Information on Cashew Producers in Guinea-Bissau*"), I study how lack of market information affects the commercialisation decisions of smallholder producers of export crops.[1] With my co-authors, I conduct a nationwide randomised controlled experiment providing reliable weekly messages with market news, prices, and sales advice to smallholder cashew-nut producers in Guinea-Bissau. Cashew nuts account for 90% of the country's exports and are a key source of income for producers. However, producers often sell at low prices because they lack information on price fluctuations, which are mostly driven by international markets. The chapter's main contribution is showing that access to market information can improve market outcomes for producers. Treated producers report higher farmgate prices than the

[1] Co-authored with Brais Álvarez Pereira, Adewusi Mendonça, Dayvikson Raiss Laval Tavares, Sebastian Schäber.

control group. The results can be explained by an increase in bargaining power.

In the second chapter ("*Labour Market Effects of Ethiopia's Social Safety Net*", I assess how a social protection programme in rural Ethiopia affects farmers' occupational choices. I provide the first assessmant of how this affects labour markets and find that the programme shifted employment from agriculture towards non-agricultural self-employment. These shifts are the result of indirect changes to the local labour market, rather than direct changes to beneficiaries. I find no changes to private sector wages or overall employment. I suggest that the rise in non-agricultural self-employment is consistent with the public works component of the programme having stimulated market access through infrastructure development.

Finally, the third chapter (*Medium-term Effects of a Role Model Intervention in Rural Ethiopia*") studies whether people in poverty may not make long-term investments if they doubt that a better life is possible.[2] This study randomly selected households in remote rural Ethiopia to watch video documentaries. The videos described role models that had improved their welfare through their own efforts. The role models came from similar communities as the study participants and worked in agriculture or business. Five-years after the screening took place, those that watched the videos changed their behaviours. They had increased investment in their farms and towards their children's education. The results can be explained by an increase in aspirations in terms of lifetime goals. This chapter's main contribution is finding persistent effects of a light-touch intervention targeting aspirations on economic investments.

The thesis is bound together by the increased use of randomised controlled trials and causal inference to accelerate poverty reduction over the last two decades. It aims to address the limitations of short-term and small-scale evaluations when policies are implemented at scale. Specifically, the thesis seeks to contribute to our understanding of different interventions' spillover effects and their impacts over multiple years.

[2] Co-authored with Tanguy Bernard, Stefan Dercon, Kate Orkin, Alemayehu Seyoum Taffesse.

Chapter 1

Impact of Market Information on Cashew Producers in Guinea-Bissau

Acknowledgements: This chapter is co-authored with Brais Álvarez Pereira (Universidade Nova de Lisboa, NOVAFRICA, BELAB, & Partnership for Economic Policy); Adewusi Mendonça (Ministry of Economy and Finance of Guinea-Bissau & BELAB); Dayvikson Raiss Laval Tavares (Universidade Lusofona de Bissau, Ministry of Economy and Finance of Guinea-Bissau, & BELAB); Sebastian Schäber (European Commission). This research work was carried out with financial and scientific support by the Partnership for Economic Policy (PEP www.pep-net.org) with funding from the Hewlett-Foundation Government of Canada through the International Development Research Center (IDRC), Private Enterprise Development in Low-Income Countries (PEDL), and the Government of Guinea-Bissau. In particular, we thank PEP for the initial support and especially our project mentor, Marcos Agurto Adrianzen (University of Piura), who provided insights that greatly assisted the research. We thank the NGO Nitidae and the mobile operator MTN Guinea-Bissau for their support and development of the cashew market information service *n'kalô* in Guinea-Bissau. We extend our gratitude to the excellent enumerators and supervisors that worked long hours to collect data across almost all of Guinea-Bissau, both in-person and over the phone. In particular, we would like to thank Jayrson Deoclecio Marino, Jair Martins, and Tales Undiga for their invaluable research assistance along the project. We thank the National Cashew Agency (ANCA) for sharing their previous experiences in developing a national market information system and André Nanque for his excellent insights on the cashew supply chain and market. We are also immensely grateful to Pramila Krishnan for her support and patience throughout the project. We are very grateful to Simon Quinn, Craig McIntosh, Binta Zahra Diop, Chris Woodruff and participants of the Applied Microeconomics Workshop, CSAE Conference 2021, Quantitative Development Workshop, CSAE Workshop, and DPhil Peer Presentations at the University of Oxford, Econometric Society Africa Region meeting 2021, NEUDC conference 2021, PacDev conference 2022, Advances with Field Experiments conference 2022, NOVAFRICA conference 2022, AFDEV conference 2022, University of Johannesburg's EDWRG seminar, CEPR/IFS/UCL/BREAD/TCD Workshop in Development Economics 2022 for comments and thoughtful suggestions. All usual disclaimers apply. This study has ethics approval from the University of Oxford (protocol # SSD/CUREC1A/ECONCIA19-20-20) and the Ministry of Economy and Finance of Guinea-Bissau.

1.1 A market information experiment

Prices in agricultural commodity markets fluctuate considerably (Deaton and Laroque, 1992). Yet, for export crops, pass-through rates from international to farmgate prices are low. Low pass-through rates could be due to high transport costs and competitive forces along the value chain (Casaburi et al., 2013; Casaburi and Reed, 2022) or informational asymmetries. Producers of export crops might lack sufficient information about price fluctuations compared to intermediary buyers who are well-integrated into global value chains. As a result, intermediary buyers are likely to be better informed about prices and market news (Fafchamps and Hill, 2008). Lack of accurate price information could weaken producers' bargaining power and lower the prices they can negotiate.

This paper studies whether and how digital technologies affect producers' outcomes by increasing the availability of accurate price and market information. Our study takes place in Guinea-Bissau among producers of cashew nuts, an export crop with high price volatility. Cashew nuts account for 90% of the Guinea-Bissau's exports and are a key source of income for producers. To support producers in their sales decisions, we introduced a new market information system in the country, which combined market and price information from both international and local markets.[1] Our market information system provided selected producers with weekly voice and text messages to their mobile phones during the trading season, which runs between March and July. These messages were tailored to be easy to understand and to provide producers with up-to-date market information.

We evaluate the effect of providing market information by implementing a two-level cluster randomised control trial. We randomised treatment assignment both across and

[1] We partnered with the Ministry of Finance and Economy of Guinea-Bissau, the National Cashew Agency, and the mobile operator MTN to introduce to Guinea-Bissau the *n'kalô* service. *n'kalô* is a market information system designed by the French NGO Nitidae, which already operates in several cashew-producing countries and is a globally trusted source of information for this commodity market. The intervention is described further in Section 1.4.1.

within villages. In treated villages, a randomly selected group of cashew producers received free messages from our market information system during the trading season across two years, in 2020 and 2021. Our research design enabled us to estimate within-village spillover effects due to the presence of both treated and untreated producers in treated villages. We assessed the intervention's impact on various outcomes through an in-person survey and a second phone-based follow-up conducted after the 2020 and 2021 trading seasons, respectively. Analysis plans were registered for each follow-up survey before their completion.[2] During the first year of the intervention, only the treated group had access to these messages. In the second year, after the in-person survey, we informed untreated producers about the service and told them how they could subscribe to it for a monthly fee charged by the mobile network operator.

The cashew-nut market in Guinea-Bissau is a useful setting to understand how market information can help producers negotiate higher prices. First, Guinea-Bissau is a price taker in the global cashew-nut market. Cashew prices fluctuate across and within the trading seasons largely due to exogenous fluctuations in international prices. Prior to our intervention, these price fluctuations were not well communicated to producers. Second, producers make a high-stakes decision when deciding at what price to sell their cashews. Most producers earn their annual revenue from selling all their production in one or two transactions. Third, every year, producers have about four months during the trading seasons to make their transactions. While producers do not store their stock across different trading seasons, most producers can store raw cashew nuts post-harvest for several months within the same trading season since raw cashew nuts are not a highly perishable commodity. Fourth, there are few opportunities for spatial arbitrage for producers. Most producers sell their cashews at the farmgate to intermediary buyers, who live or visit the village during the trading season, with very

[2] See https://www.socialscienceregistry.org/trials/4740 for the trial registration and analysis plan.

few producers selling in centralised wholesale markets.

We have three main results showing that access to market information can support producers in their sales decisions. First, we find that the information system increased farmgate prices. We find a 2% intent-to-treat effect on average prices comparing treated producers with control producers in 2021. Whereas in 2020, we find a small and not statistically significant treatment effect. *Between*-village spillovers can explain why we did not find a larger increase in average prices. Our intent-to-treat estimates could be an underestimate of the effect of the intervention, which could have reached some control producers in more intensely-treated areas. We find that an additional treated producer within a 5 km radius increases the average farmgate price by 1%, affecting also producers in control villages closer to treated villages.[3] When accounting for between-cluster spillovers we find a positive and direct treatment effect of the intervention in both trading seasons, comparing treated producers with control producers without any treated producers within a 5 km radius. Second, we show that treated producers sold their cashews more frequently relative to the other producers across both trading seasons. This result is consistent with the advice that treated producers received through our messages, which advised them to sell their stock in multiple transactions. Treated producers report that they preferred to sell more than once as a result of the advice received and because they wanted to smooth their income. We interpret the changes in the frequency of sales as evidence that better informed producers attempt less risky sales strategies. Third, treated producers earned 20-23% higher revenue from all sales and barters, relative to the control group mean. The increase in revenue is consistent with the positive price effect and is also due to a larger amount of cashews sold in market transactions, instead of barters, own consumption, or towards repayment of loans. We are cautious in not over-interpreting the magnitude of this

[3] Although we acknowledge that we had not pre-specified our between-cluster analysis, our estimates of between-cluster spillovers are robust to endogenous exposure to our exogenous cluster-level randomisation, once we follow the methodology recently proposed by Borusyak and Hull (2021).

effect, as the estimates on revenue are noisy, and we cannot fully account for the increase in the quantity sold.

The main mechanisms that explain our results is that our intervention increased bargaining power of producers. First, we find that treated producers in 2021 are 7% more likely than the control group to report having negotiated a higher price than the one that they were initially offered. Second, compared to control producers, treated producers are 6% more likely to sell to itinerant traders. Itinerant buyers can negotiate prices more freely than local buyers who typically receive a fixed fee from an exporter. Third, treated producers are less likely to rely on their buyers as a source of information for prices. Instead, treated producers are more likely to report having used the messages as a source of information to guide their sales decisions. Exploiting interviews conducted with buyers across a sub-sample of study villages, we find that treated villages had higher pass-through than control villages, as buyers earned lower margins but bought higher quantities. While these differences across villages are not statistically significant, they are consistent in their direction and magnitude with those obtained using the producers' data. Fourth, we find that our effects are more pronounced for producers with larger plots, consistent with the idea that producers with more cashews to sell have more room for negotiation.

Our findings are not consistent with other behavioural explanations. First, we do not find evidence that producers changed the timing of their sales to increase prices. Instead, we find that treated producers start selling some of their stock earlier, despite prices being lower at the beginning of the trading season. These earlier sales are consistent with the treatment reducing price uncertainty leading producers to revise downwards their reservation prices.[4] With lower reservation prices, treated producers are more likely to accept offers that they would have otherwise rejected. These earlier

[4] Consistent with this interpretation, we find some weak evidence that treated producers had a more realistic outlook on prices for the upcoming trading season, suggesting that the information revised downward producers' reservation prices.

sales also help producers smooth their income during the trading season. By shifting some transactions earlier, treated producers are less likely to barter their cashew for rice later in the trading season, when the terms of barter are worse than market transactions. Second, we do not find evidence consistent with our effects being the result of better record-keeping induced by our intervention or by changes in preference parameters, such as risk aversion. Third, we also do not find evidence consistent with the timing of the messages changing the salience of transactions. Receiving regular messages could have induced producers to sell more frequently just by making transactions more salient. We do not observe producers being more likely to report a transactions within two days of having received a message, leveraging transaction diaries that producers filled with the exact date of their transaction.

We add to the literature on the effects of information communication technologies in agricultural markets.[5] Prior work found mixed results on the benefits of market information systems for farmers.[6] Our study has three novel differences compared with previous studies. First, we focus on an export commodity, whose local price is primarily driven by international price fluctuations. Previous studies have not focused on export commodities but rather on internally consumed commodities. We argue that intra-seasonal price variation for internally consumed commodities is more likely to be known to producers than the price variation of internationally traded commodities, though both types of commodities expose farmers to price risk (Cardell and Michelson, 2023). Second, we are the first to evaluate a market information system that disseminated price information via audio messages (through robocalls and an interactive voice-response system) as opposed to text-based messages. This novel

[5] See Nakasone et al. (2014) or Aker et al. (2016) for reviews of this literature.

[6] Randomised evaluations of MIS in Colombia (Camacho and Conover, 2019) and India (Fafchamps and Minten, 2012; Mitra et al., 2018) have failed to find a significant average treatment effect on producer prices. In contrast, Svensson and Yanagizawa (2009) and Soldani et al. (2023) find that a MIS in Uganda and Ghana, respectively, increased producer prices by 7-10% for specific crops. In a related intervention, Goyal (2010) studies the expansion of information kiosks in district markets in Andhra Pradesh, and finds that the kiosks increased producer prices by about 1-3%.

means of communication addresses the barriers faced by users with low levels of literacy in interpreting the information they received. Previous studies identified low levels of literacy as a potential explanation for the lack of positive effects of similar text-based interventions (Fafchamps and Minten, 2012).[7] Third, our messages provided not only a point estimate for the wholesale market price (Mitra et al., 2018) but also a range of farmgate prices for all regions in Guinea-Bissau. The messages also included market information on expected price developments and sales advice on when and at which price to sell production.

We contribute to the literature on agricultural commodity markets by showing that imperfect information can affect the timing of sales. Information frictions could be both spatial and temporal. Producers may lack information on both *where* and *when* to sell their output to maximise profits. With the introduction of new information technologies producers have been better able to decide *where* to sell (Jensen 2007; Aker 2010; Casaburi et al. 2013; Allen 2014; Parker et al. 2016). Instead, our study shows that producers may also lack information on *when* it is better to sell and how much to sell at different points in time. Credit or storage constraints may prevent producers from selling at the optimal time (Aggarwal et al. 2018; Kadjo et al. 2018; Burke et al. 2019; Mukherjee et al. 2021). We are the first to experimentally document how information frictions affect market timing in commodity markets. Earlier studies in this literature relied on non-experimental data to look at the effects of information on market performance over time (Osborne 2004; Fafchamps and Hill 2008).

Finally, our study contributes to the literature on information exchanges in agricultural markets. A large literature has documented the role that communication networks within rural communities play in the adoption of new technologies and sharing of agricultural practices, with a strong focus on within-village spillovers (Foster and

[7] Cole and Fernando (2021) and Gupta et al. (2021) study an agricultural extension service that informed Indian farmers of different production practices via audio-based messages. But, they did not provide sales advice or price information.

Rosenzweig, 1995; Conley and Udry, 2010; Magruder, 2018; Beaman et al., 2021). Similarly to Fabregas et al. (2019); Nakasone (2013), who respectively evaluated an SMS-based agricultural advice service and a price information system, our study was explicitly designed to estimate potential *within*-cluster spillovers of our intervention.[8] More recent work has attempted to also measure the extent to which spillovers of market information systems may occur *between*-clusters, especially within groups of villages that had pre-existing strong networks (Soldani et al., 2023; Falcao Bergquist et al., 2021). Our study covers a vast geographic area, encompassing nearly the entire country, enabling us to identify between-village spillovers by leveraging the random intensity of treatment induced by our cluster-level randomisation, as in Miguel and Kremer (2004) or Egger et al. (2022).

This article is structured as follows. We begin by illustrating the main features of the market for cashew nuts in Guinea-Bissau, briefly describing its supply chain and the characteristics of the sample of producers we work with. To motivate the intervention, in Section 1.3 we lay out a simple conceptual framework to illustrate how the intervention may affect the frequency, timing, and prices of sales. In Section 1.4, we describe our intervention, our sampling and randomization protocol, and our estimation and inference strategy. Section 1.5 presents results on our main outcomes of interest. Before concluding, in Section 1.6, we analyze a set of potential competing explanations for the underlying mechanisms of our main results.

[8] A few papers study the presence of spillovers without relying on randomised saturation design. For example, (Cole and Fernando, 2021) and (Camacho and Conover, 2019) use variation induced by the experiment to also estimate spillovers *within*-clusters, such as the share of producers treated within a cluster or the share of producers' network that was randomly assigned to treatment. The design of Fafchamps and Minten (2012)'s study also allows for the estimation of *within*-village spillovers, but the authors do not report them, presumably since they did not find any direct effect of the SMS-based intervention evaluated.

1.2 Context of the study

1.2.1 International market outlook and the cashew value chain

Cashew nuts account for 90% of Guinea-Bissau's exports and are a key source of income for producers. The majority of production is exported unprocessed to countries such as India, Vietnam, and China, with only 10% serving the domestic market. The country's cashew market is significantly influenced by international supply and demand dynamics, similar to other internationally traded raw agricultural commodities like coffee and cocoa. However, the difference between export and farmgate prices varies greatly across years (see Figure 1.1), suggesting that market imperfections may reduce pass-through rates to producers.[9] To curb speculation, the government sets a reference farmgate price at the beginning of the trading season. However, this price may not always reflect market dynamics. Although it can influence market expectations, buyers are not obliged to pay this price when purchasing cashews.[10]

The cashew nut value chain in this market involves producers, intermediaries, wholesalers, and exporters. Intermediaries purchase raw cashew nuts from producers and either sell them to wholesalers or exporters. Intermediaries are either local or itinerant. Local intermediaries that reside in the village usually rely on exporters or wholesalers for pre-financing and earn a fixed commission. Itinerant intermediaries are more likely to self-finance their cashew purchases through other businesses or informal credit. Wholesalers collect cashews from smaller intermediaries and handle transportation and storage until selling to exporters. The export segment is relatively competitive.

[9] All monetary values reported are in nominal West African CFA francs (XOF), which is pegged to the euro at an exchange rate of 1 EU for 656 XOF.

[10] For example, the difference between export and farmgate prices spiked in 2018, the year before our study. Export prices were twice the farmgate price. This spike was partly due to the government setting a reference farmgate price of 1000 XOF/kg, which exceeded what buyers were willing to pay and deterred entry among buyers. Consequently, many producers eventually sold their cashews at considerably lower prices as they could not wait any longer for the expected higher price.

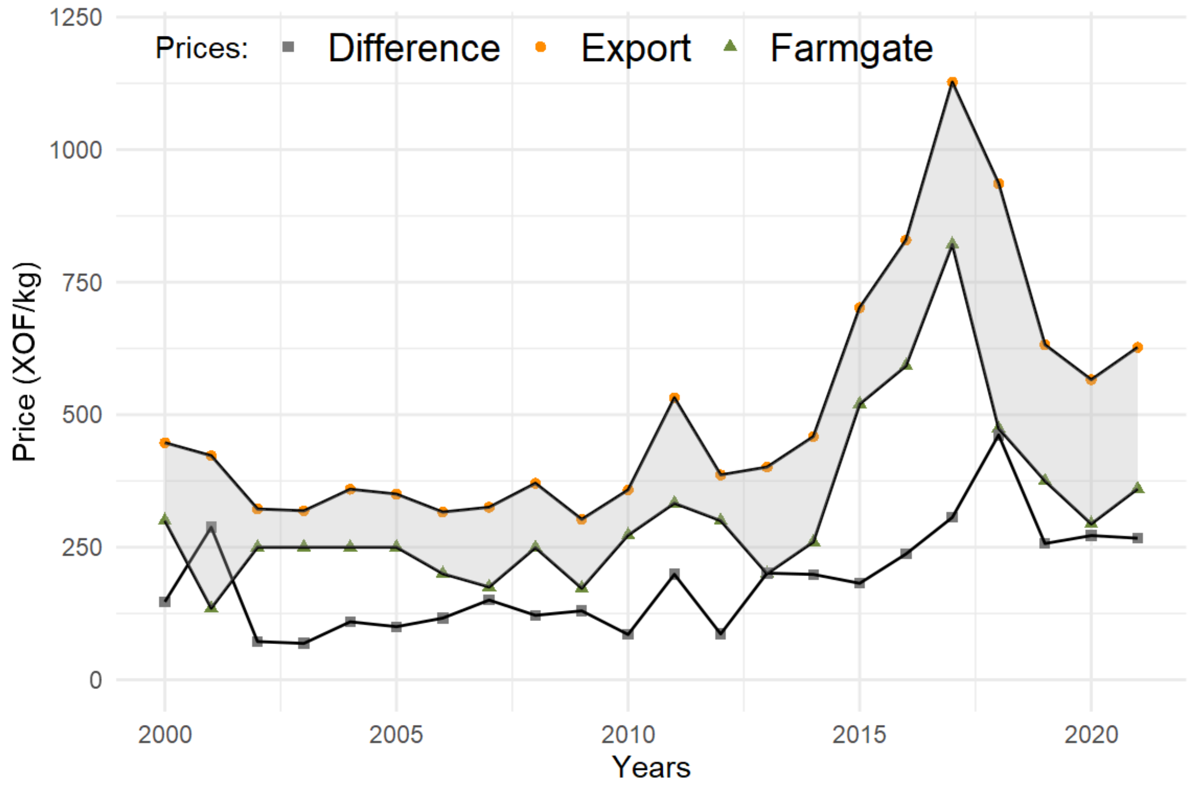


Figure 1.1: Margin between export and farmgate prices across years.[a]

[a] Source: Directorate for Macroeconomic Forecasting, Ministry of Economy and Finance of Guinea-Bissau. Raw cashew-nut prices per kg reported in nominal West African CFA francs (XOF). Export prices are free-on-board annual average prices from the Bissau port collected by the Ministry of Commerce, and farmgate prices are national annual averages collected by the National Statistics Institute.

Exporters finance most of the operations in the value chain, fix contracts with processing plants, and bear the costs from farmgate to shipping, including insurance, banking, transport, and storage (Cont and Porto, 2014).

Margins between export and farmgate prices can vary also within the same year. Figure 1.2 reports the month-by-month average prices during the 2019, 2020, and 2021 trading seasons. During these years, the difference between export and producer prices moved by almost 50 XOF (17% of the farmgate prices) across different months. In general, farmgate prices tend to increase towards the end of the trading season. However, the

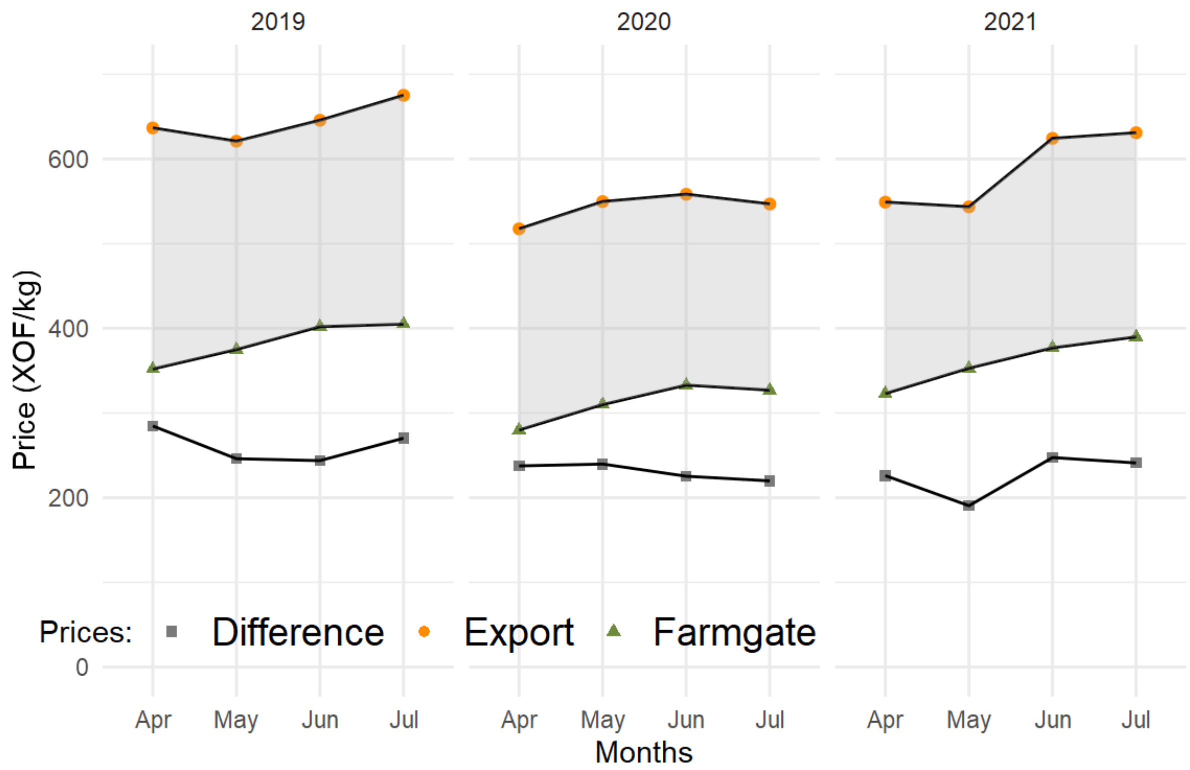


Figure 1.2: Margin between export and farmgate prices *within* years.[a]

[a] Source: Directorate for Macroeconomic Forecasting, Ministry of Economy and Finance of Guinea-Bissau. Raw cashew-nut prices per kg reported in nominal West African CFA francs (XOF). Export prices are free-on-board monthly average prices from the Bissau port collected by the Ministry of Commerce, and farmgate prices are monthly averages collected from producers in our study.

starting point and peak of prices within the trading season still exhibit volatility from year to year.[11]

1.2.2 Descriptive statistics on cashew producers

At least 40% of households in the country are involved in cashew production. The majority (85%) of raw cashew nuts producers are smallholders. The remainder of this section characterizes producers in this market using data from our baseline survey.

[11] For example, overall prices dropped from 2019 to 2020 in part as a result of the COVID-19 pandemic shock, and increased between 2020 and 2021, in part due to the resumption of normal operation of processing factories. Appendix Section 1.11.3 provides more details on how the trading seasons were affected by the COVID-19 pandemic.

Table 1.1: Sample characteristics at baseline

	Mean	St. Dev.	25th percentile	Median	75th percentile
Individual-level characteristics:					
1 if cashews are the main source of income	0.80	0.40	1.00	1.00	1.00
Area plantation (hectares)	3.45	2.58	2.00	3.00	4.00
Total quantity produced (kg)	1526.59	1627.36	669.50	1125.00	1935.00
1 if a woman	0.06	0.23	0.00	0.00	0.00
Age	42.91	14.53	32.00	41.00	52.00
Household size	13.66	8.16	8.00	12.00	17.00
Years of education	3.95	3.85	0.00	3.00	6.00
Years of experience in cashew sector	10.41	7.11	5.00	9.00	15.00
1 if can read	0.29	0.45	0.00	0.00	1.00
1 if faces storage limitations	0.31	0.46	0.00	0.00	1.00
Number of sales	1.64	0.79	0.00	1.00	2.00
Number of potential buyers that made offers	2.82	1.77	1.00	2.00	4.00
Average price (XOF/kg)	390.92	100.58	332.18	370.94	450.00
1 if thinks that the reference price is important	0.38	0.48	0.00	0.00	1.00
1 if sold cashews in their own village	0.97	0.17	1.00	1.00	1.00
1 if sold most to local buyer	0.39	0.49	0.00	0.00	1.00
1 if sold most to itinerant buyer	0.54	0.50	0.00	1.00	1.00
1 if sourced market information from family	0.57	0.50	0.00	1.00	1.00
1 if sourced information on best time to sell	0.54	0.50	0.00	1.00	1.00
1 if heard of mobile MIS	0.06	0.24	0.00	0.00	0.00
Village-level characteristics:					
Number of cashew producers in the village	32.66	24.26	17.00	24.00	41.00
Number of cashew buyers in the village	3.57	3.41	1.00	3.00	5.00
Road distance in km to nearest sector capital	27.86	22.84	10.88	21.67	37.39
Road distance in km to nearest region capital	51.51	33.28	26.43	46.45	67.19
Road distance in km to the capital	160.63	68.37	102.08	155.56	215.86
Observations	1988				

Our sample comprises 1,988 producers across 290 villages in the country and is fairly representative of the majority of cashew producers in Guinea-Bissau.[12]

Villages in our study are in rural, often remote, areas. The median village has 24 cashew producers and three buyers that reside in the village. They are distant from markets. The median village is 22 km away from the closest sector capital, 46 km from the closest regional capital, and 155 km away from the capital, Bissau, where most of

[12] Appendix 1.8 provides details of our sampling strategy.

the raw cashew nuts are exported from.[13]

Producers in our study own relatively small plots and rely heavily on cashew production as a source of income. Our descriptive statistics, in Table 1.1, show that cashew sales constitute the primary source of income for 80% of our sample. The size of the median cashew field was around 3.4 hectares. In 2019, the average producer had reported producing about 1,500 kg of cashews. In terms of demographic characteristics, almost all producers in our sample are men (94%) with a median age of 41. Household sizes are large: the median household had twelve individuals. The median producer completed three years of education and had been working in the cashew sector for the previous nine years. About 30% of respondents were unable to read a basic sentence, hence developing a service that takes into account low literacy skills is important for many producers.

In general, producers sell their entire annual harvest during the same trading season between March and June. Most producers do not store cashews across seasons because their quality would deteriorate significantly during the rainy season, which starts in July. About two thirds of our respondents reported having no difficulties storing their cashews within the same trading season. Most producers reported that they concentrate their sales into a single transaction, as shown in Table 1.1. The median number of potential buyers is two, but the median number of sales is just one. This finding implies that deciding when to sell has high stakes for the majority of producers, as it will determine the largest share of their annual income. The reported farmgate price across all sales averages 390 XOF per kg. This price is significantly below the reference price of 500 XOF per kg that the government had originally proposed at the beginning of the 2019 trading season. In part due the discrepancy between the reference and the farmgate price, 38% of the respondents stated they believed the

[13] Guinea-Bissau is administratively divided into 9 regions, including a semi-urban region for the capital, Bissau. Regions are sub-divided into sectors, which are a smaller administrative unit. There are on average four sectors per region.

government reference price to be an important factor in their sales decisions.

Spatial arbitrage is rare in this market, as 97% of sales occur at the producers' home or somewhere else in the producer's village. Producers rarely transport their output to other markets, as doing so poses more risks and costs for them. Most producers do not have a network of potential buyers beyond the ones in their own village.[14] 54% of buyers are traders that temporarily visit the villages, whilst 39% are intermediaries that live regularly in the same village as the producers and that usually act as agents for an exporter.

Over half of the respondents mainly rely on family and friends for sales advice, while only a few use formal channels like producer associations. Among those seeking advice, the timing of the sale is one of the most sought-after types of information, along with the right price. Only 6% of producers in our sample had heard of any market information systems.

1.3 Conceptualising producers' sales decisions

To motivate our intervention, we outline a simple framework that examines the impact of increased information access on timing, sales frequency, and average prices. We borrow the model from Mitra et al. (2018) and incorporate our insights to explain specific aspects of the market in our experiment. The section aims to demonstrate that a straightforward framework capturing essential elements of our environment can lead to uncertain predictions about the effect of providing market information, contingent on behavioural assumptions.

[14] Based on qualitative interviews implemented in villages in our study.

1.3.1 A two-period model of producer sales

A producer negotiates with an intermediary buyer. There are two periods, $t = 1$ and $t = 2$, representing the first and second half of the trading season, respectively. Intermediaries resell cashews at export price x, which producers cannot directly access. Producers can sell to intermediaries in the village or go to wholesale markets, where they receive the reservation price $M(x_t)$, as an alternative if they cannot agree on a price with the intermediary. The total quantity for sale is normalised to 1, and producers choose proportions q_1 and q_2 to sell in each period. They have a prior belief about x following distribution G, with support $[\overline{x}, \underline{x}]$. Producers maximise revenue $W(y_1) + \delta W(y_2)$, where y_t is revenue at time t, $W(.)$ is a strictly concave and strictly increasing function with $W'(0) = \infty$, and $\delta \in (0, 1)$ is the discount rate. Producers cannot borrow across periods, and traders are risk-neutral.

To solve the producers' problem, we use backward induction. In period 2, the producer takes q_1 and p_1 as given. The equilibrium is non-revealing, characterized by the farmgate price offered $p_2^* = E(M(x_2)|p_1)$ and a quantity sold $(1 - q_1)$, meaning the producer accepts the trader's offer without gaining information about export prices. In the first period, the non-fully revealing equilibrium offer, as in Mitra et al. (2018), is accepted if $p_1 \geq E[M(x_1)]$. To maximise revenue $W(p_1q_1) + \delta W(p_2(1 - q_1))$, the producer chooses q_1^* with the associated first-order condition given by Equation (1.1).

$$p_1 W'(p_1 q_1) = \delta p_2^* W'(p_2(1 - q_1)) \tag{1.1}$$

The relationship between the first-period price and quantity sold depends on wealth and substitution effects. The concavity of $W(.)$ represents the wealth effect, where an increase in p_1 reduces the marginal value of revenue in that period. The substitution effect is represented by p_1 pre-multiplying the left-hand side of Equation (1.1). The net effect depends on the curvature of $W(.)$. For instance, assuming $W(y) = \frac{y^{1-\theta}}{1-\theta}$, if θ

is greater than 1, a price increase in the first period decreases q_1. Otherwise, if θ is smaller than 1, the substitution effect dominates, leading to an increase in q_1.

1.3.2 How would an information intervention change sales and prices?

The intervention is modeled as a binary signal σ_t in each period, updating the producer's beliefs about the price distribution. If $\sigma_t = L$ and the signal indicates a low price, the producer expects $x_t \in [\underline{x}, \hat{x}]$. Alternatively, if $\sigma_t = H$ and the signal indicates a high price, the producer expects $x_t \in [\underline{x}, \overline{x}]$.

The farmgate price in each period depends on the signal, $p_{t(\sigma_t)}$ such that $p_{t(L)} < \tilde{p}_t \leq p_{t(H)}$, where $\sigma_t \in \{L, H\}$ and $\tilde{p}_t$ is the pre-intervention price. The proportion of output sold at $t = 1$ satisfies the following first order condition:

$$p_{1(\sigma_1)}W'(p_{1(\sigma_1)}q_1) = \delta[\alpha_{(\sigma_2)}p_{2(H)}W'(p_{2(H)}(1-q_1))+(1-\alpha_{(\sigma_2)})(p_{2(L)}W'(p_{2(L)}(1-q_1)))] \quad (1.2)$$

where the producer believes that the price will be high with probability $\alpha_{(\sigma_t)}$.

In both the 2020 and 2021 trading seasons, the information provided to producers indicated negatively correlated export price shocks across the two periods, with $\alpha_{(H)} \leq \alpha_{(L)}$. In this scenario, producers who received a low signal in the first period could infer a higher probability of increased prices in the second period. Given the concavity of $W(.)$, if the producer observed a low signal in the first period, they would sell more in that period to smooth revenue across periods. Treated producers might receive lower prices in the first period but higher prices in the second period if wealth effects dominate.[15]

The model yields an ambiguous prediction on the effect of information on the average sale price of treated producers, depending on how the following two scenarios balance

[15] A similar result is true if we assume the export prices to be independent across the two periods, i.e., $\alpha_{(H)} = \alpha_{(L)}$. In this case, the right-hand side of Equation (1.2) is independent of the signal, but a low signal in the first period would still induce more sales by concavity of $W(.)$.

each other. First, receiving a low price signal in the first period, may actually reduce the sale price of informed producers in that period. This is in part due to the fact that these producers may be more likely to accept price offers that uninformed producers will not accept because uninformed producers' reservation price may be higher, as shown also in the theoretical frameworks proposed by Courtois and Subervie (2015) and Albuquerque et al. (2022). Second, since treated producers received a signal about higher prices in the latter half of the trading season, they may extract more of the trade surplus during that period compared to uninformed producers. If wealth effects dominate, we would expect an increase in the frequency of sales among treated producers, as they are likely to sell both earlier and later in the trading season. Our framework serves to highlight some less intuitive features of the decision producers face, but we do not claim it is the only way to model our intervention.[16]

1.4 Study design, data, and empirical strategy

1.4.1 Intervention: a Mobile Market Information System

Our intervention consists in providing market information to cashew producers in Guinea-Bissau during the trading season, via text or audio messages.

The weekly market information messages included three components: (i) current cashew-nut prices in different regions of Guinea-Bissau, (ii) important market news, and (iii) sales advice based on expected market trends. These messages were developed on a weekly basis by a network of market analysts known as the *n'kalô* service, a market information system already operating in other West African countries that

[16] For example, an extension of our model could include different types of buyers, such as risk-averse itinerant traders and risk-neutral local buyers. The negotiation process between a risk-averse buyer and producer would yield predictions akin to the risk-sharing model with full commitment, also discussed by Mitra et al. (2018). These predictions would be qualitatively consistent with those of our simple framework. Since we do not have data on itinerant traders preferences, we prefer to leave this extension for future work.

we introduced to Guinea-Bissau. The content was generated by combining analysis from national market analysts, who gather farmgate prices and local market updates, with insights from neighboring or international markets collected and summarised by an international analyst. Appendix Table 1.25 provides the complete content of the messages sent during both trading seasons.

As a part of the intervention, treated producers received a one-hour training on the main factors determining the farmgate price and how receiving our weekly messages could improve their sales decisions. The research team provided the training on the same day of the baseline survey, shortly after the baseline interviews were completed. The messages were delivered during the trading seasons through a combination of text messages, robocalls, and through an Interactive Voice Response (IVR) service that replayed the latest robocall sent. Specifically, in 2020 our weekly messages were delivered between April and August 2020, whereas in 2021 we delivered messages between March and July 2021, as we detail in Appendix Section 1.12.

1.4.2 Experimental design

We implemented a two-stage sampling and randomisation to conduct our study. In Appendix Section 1.8, we provide the details of how we conducted the sampling and randomisation, which we briefly summarise here. In the first sampling stage, we selected a random sample of 290 villages with the aim of picking villages that were as far from each other as possible. Second, upon visiting each village, the research team randomly sampled seven producers to be interviewed, with the support of the local village authority, to obtain an overall sample at baseline of 1988 producers.

In the first stage of randomisation, we randomly allocated two thirds of villages to receive our intervention, and one third to act as our control group. We stratified random assignment by creating triplets of villages that were most similar along a set

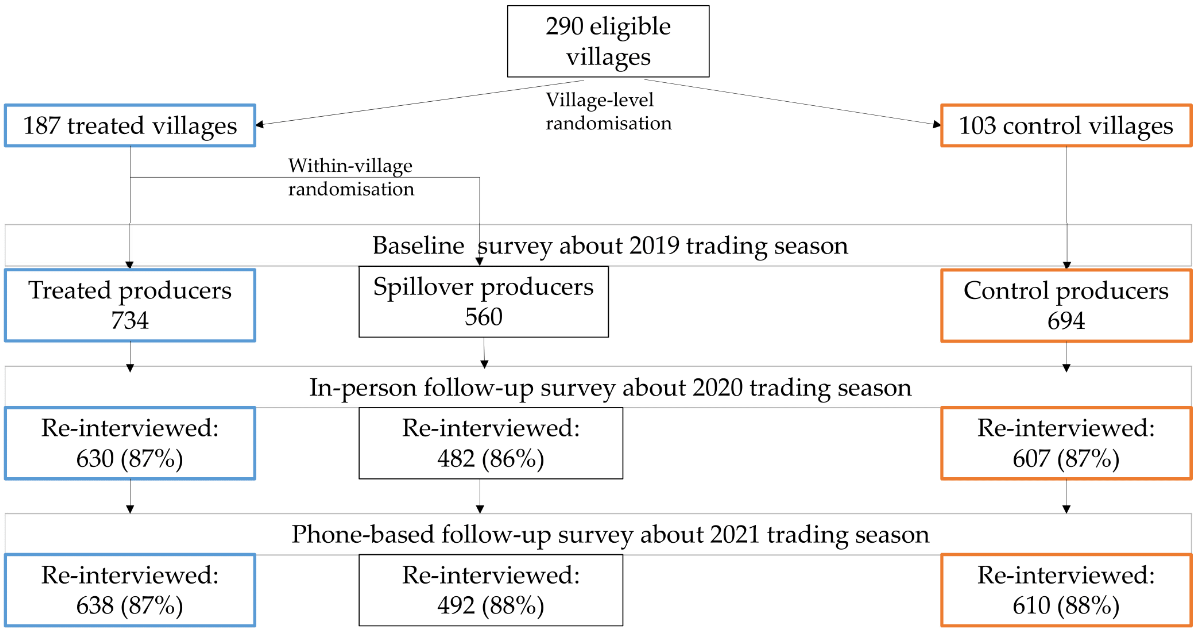


Figure 1.3: Flowchart of participants through the study[a]

[a] Notes: Each box denotes the number of participants (or villages) in the study, by treatment group and survey rounds. Percentages in parentheses indicate the re-interview rate relative to the baseline.

of village-level characteristics. In the second stage, in treated villages, the research team randomly allocated the intervention only to four out of the seven interviewed producers through a public lottery. Figure 1.3 shows how we allocated participants across treatment, spillover, and control groups.

We collected three rounds of interviews with cashew producers to evaluate the effects of our market information system. First, we conducted an in-person baseline survey between October and December 2019. Second, between April and May 2021 we conducted an in-person follow-up survey to ask about the trading season in 2020, once it was safe to resume face-to-face contact with sanitary precautions. During this survey we interviewed 86% of producers that had taken part in the baseline in 2019. We also recovered a transaction diary that we had left with producers during the baseline survey for them to fill with details of each sale during the 2020 trading season. After completing the in-person follow-up survey, producers in both the treatment and control villages were provided with information on the *n'kalô* service and were shown how to subscribe as fee-paying users to receive information during the remainder of the 2021 trading season. Treated producers continued to receive free messages in 2021. Third, to assess the effects of our information in 2021, we conducted a phone-based follow-up on a key set of outcome variables between December 2021 and March 2022, during which we reached almost 88% of the baseline sample.[17]

1.4.3 Empirical strategy

Our main specification is:

$$y_{iv} = \beta \cdot treatment_{iv} + \delta \cdot spillover_{iv} + \gamma \cdot y_{0iv} + \alpha_{t(v)} + \epsilon_{iv} \tag{1.3}$$

where: i and v index individuals and villages, respectively; y_{iv} denotes the outcome of interest measured in the follow-up; y_{0iv} denotes the outcome of interest measured in the baseline; $treatment_{iv}$ denotes individual-level assignment to the treatment group; $spillover_{iv}$ denotes individual-level assignment to the spillover group in treated villages; $\alpha_{t(v)}$ denotes the randomisation triplet fixed effect (as described in Section 1.8.3) and $t(v)$ indexes the triplet of villages; ϵ_{iv} is the unobserved variation in the outcome. We present estimates from Equation 1.3 separately for the 2020 (Year 1) and 2021 (Year 2) trading seasons.

We cluster standard errors at the village-level, the unit of our first level of randomisation. Our coefficient of interest is β, the intent-to-treat (ITT) effect. In Appendix

[17] Between July and October 2020 we had also conducted a shorter phone-based follow-up to check how our study participants were coping with the COVID-19 pandemic. Though in our first pre-analysis plan we had intended to use the data from this interim phone-based follow-up to assess the impact of the service, we decided against using it for two reasons. First, we had mistakenly contacted treated producers and later control producers making the timing of the call correlated with treatment status. Second, we only reached 66% of the baseline sample during this survey, partly because we detected fraud by one member of the research team which meant we had to discard about 10% of the interviews. Results from this interim survey are available upon request and are broadly consistent with the data we collected in person in 2021.

Section 1.10 we also describe alternative empirical specifications (e.g. top-winsorizing continuous variables, including additional controls to account for baseline imbalances, clustering at the triplet-level) to test the robustness of our main specification.

1.4.4 Inference and multiple-hypothesis testing adjustments

For each of our pre-specified outcomes, we test the following hypothesis:

(i). $H_0 : \beta = 0$: The intervention had no effect;

(ii). $H_0 : \delta = 0$: The intervention had no spillover effect;

(iii). $H_0 : \beta = \delta$: The intervention had no effect relative to the spillover group;

For each of these hypothesis tests, we report the False Discovery Rate (FDR) adjusted q-values, taken across the family of outcomes (Benjamini et al., 2006). For each type of test, we construct a q-value for that test across outcomes.[18]

For clarity, we link the statistical hypotheses above with the underlying economic predictions. A rejection of the null hypothesis (i) implies that the random allocation to receive market information did affect producers' outcomes, such as prices and revenue. Whereas a rejection of the null hypothesis (ii) implies that the intervention has affected the outcomes of untreated producers living *within* treated villages. These spillovers may be due, for example, to communication between treated producers and untreated producers living within the same village. Rejecting null hypothesis (iii) implies that producers that were allocated to the intervention (and also participated in the training) are affected by the treatment relative to untreated producers living in the same village.

[18] Specifically, we construct a set of q-values using all p-values for the null hypothesis 'The intervention had no effect'; we construct a set of q-values using all p-values for the null hypothesis 'The intervention had no spillover effect', and a set of q-values using all p-values for the null hypothesis 'The intervention had no effect relative to the spillover group'.

1.4.5 Experimental integrity

Appendix Table 1.7 show that our baseline sample is relatively balanced across treatment arms for our main outcome variables of interest, except for two of the nine variables. The differences between the mean of the quantity (and value) of cashews exchanged for rice at baseline is larger in the treatment group compared to the spillover and control groups, but these differences are not very large, accounting for about 4% of the quantity sold in monetary transactions. Our empirical strategy includes the baseline value of the outcomes as control variables, which would allay concerns that imbalance in the outcome variables of interest would affect our estimates of the treatment effects.

Appendix Tables 1.8 and 1.10 show that our baseline sample is also relatively balanced across treatment arms across producer-level and village-level baseline characteristics. The differences between the mean of 9 of the 25 producer-level characteristics are statistically significant across treatment groups. However, these differences are not large, as none of the pairwise standardised differences in means are larger than 0.21 standard deviations. Out of the eight village characteristics considered, we find that treated villages have one more village trader relative to control group villages on average. As a robustness check for our main outcomes of interest, we control for the producer and village characteristics unbalanced at baseline, and report these results in Appendix Section 1.10.2.

Overall attrition in both our follow-up surveys was small and uncorrelated with treatment assignment. In the in-person follow-up survey attrition was 14% and in the subsequent phone-based survey attrition was less than 13% overall, as shown in Appendix Table 1.11.[19] We have at least one follow-up with about 78% of our baseline

[19] These estimates are unchanged if we predict attrition indicators via post-double least absolute shrinkage and selection operator (Belloni et al., 2014), adding as potential controls all variables we analyse in our balance tests. None of these variables are retained as controls once we partial out the triplet fixed effects and the treatment and spillover indicators.

sample, though we focus our analysis on the sample present in either round.

Our measures of compliance confirm that the information was successfully delivered to the allocated group of producers. To measure compliance, we use both self-reported and administrative data. The measure constructed using self-reported data is an indicator based on the question "Did you use the service in [2020/2021]?".

In Table 1.2, we find a large positive difference in compliance between treated producers and the other producers. Using survey data, in 2020 (Year 1), 25% of treated producers reported using the service, while 2% and 3% of control and spillover producers, respectively, also reported using it. After completing the in-person survey, all study participants were informed of the existence of the service and shown how to subscribe to it. In 2021 (Year 2), after all participants had been informed about the service, 43% of the control group producers reported using it. Treated producers were still 18 percentage points more likely to use the service, as they continued receiving messages at no cost during the study. The administrative data, collected by our partner mobile network operator, supported the patterns observed in the self-reported data. According to this data, 53% of treated producers received at least eight messages in 2020 (a third of the total), while no messages were sent to other producers. In 2021, 28% of treated producers still received at least five messages (a third of the total), and 8% of other producers received a similar amount. Despite contamination across groups, we still find a large take-up difference between treated and non-treated producers.

We prefer not to estimate treatment-on-the-treated (TOT) effects but focus our attention to the intention-to-treat estimates. Given the potential presence of spillovers within the cluster, assuming that non-compliers might have had no effect of treatment is too strong an assumption to entertain. A naive estimate of the treatment-on-the-treated would overestimate the effect of the treatment on compliers, inflating the intention-to-treat estimates by the inverse probability of take-up.

Table 1.2: Compliance — Take-up of the intervention

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treatment-Spillover	Year 2 (5) Mean (SD) Total obs.	(6) Treatment	(7) Spillover	(8) Treatment-Spillover
Self-reported data:								
1 if heard of mobile MIS	0.14	0.30***	0.10***	0.20***	0.68	0.12***	0.04	0.08***
	(0.35)	(0.02)	(0.03)	(0.03)	(0.47)	(0.03)	(0.03)	(0.02)
	1714	[0.00]***	[0.00]***	[0.00]***	1724	[0.00]***	[0.13]	[0.00]***
1 if used mobile MIS	0.02	0.25***	0.03	0.22***	0.43	0.18***	0.05*	0.13***
	(0.16)	(0.02)	(0.02)	(0.02)	(0.49)	(0.03)	(0.03)	(0.03)
	1714	[0.00]***	[0.10]	[0.00]***	1724	[0.00]***	[0.13]	[0.00]***
Administrative data:								
Take-up rate	0.00	0.53***	0.00	0.53***	0.08	0.28***	0.01	0.27***
	(0.00)	(0.02)	(0.01)	(0.02)	(0.27)	(0.02)	(0.02)	(0.02)
	1988	[0.00]***	[0.86]	[0.00]***	1988	[0.00]***	[0.52]	[0.00]***

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 and 8 test for differences in parameters obtained in previous two columns. The top panel presents self-reported measures of engagement and take-up of the mobile market information system. The bottom panel presents a measure of take-up derived from administrative data from the mobile partner operator, MTN, that managed the technology of the service. The measure of take-up is equal to one if at least a third of the messages sent reached the respondent, where we define "reached" as an SMS being received or the robocall being listened for at least 45 seconds. In 2020, we sent 24 messages, so we codetake-up equal to 1 if respondents were reached by eight messages. In 2021, we sent 15 messages, so we code take-up equal to one if the respondents were reached by five messages. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

1.5 Results

In this section, we present the main effects of the intervention, estimated using the data collected in our two follow-up surveys. We first show results on the primary outcomes of interest (price and revenue), before turning our analysis to quantities of raw cashew nuts sold for cash or exchanged for rice. The analysis follows our pre-analysis plan, and we note where we may have deviated from it.

1.5.1 Prices and revenue

We find that producers earn higher prices once they receive better market information. Table 1.3 shows our results from both follow-up surveys on prices and revenue.

We find a positive treatment effect on prices that corresponds to about 7 XOF per kg, equivalent to 2% of the control group mean (q-value: 0.06). This effect comes from the 2021 trading season (column 6), whereas in 2020 (column 2) the effect is also positive but smaller (less than 1% of the control mean) and not statistically significant. For 2021, we also find a difference between the treatment group and the spillover group in the average prices though not robust to multiple hypothesis testing. There is no difference between treatment and spillover group in 2020.[20]

We also find noisy increase in revenue from cashews overall. The second row in Table 1.3 shows that the total value of cashews that were either sold or exchanged (directly for rice) increased by about 21-23% of the control group mean in both years for the treatment group, by 61,650 XOF (q-value: 0.05) in 2020 and by 105,533 XOF (q-value:

[20] The estimation sample includes all the producers that we reached in the respective follow-up survey. In order to not condition our outcome on the decision to sell, we recode the average price across sales to be equal to zero for producers that did not conduct any sales and add as a control an indicator equal to one for observations for which we impute this zero price. Our results are not affected by this imputation strategy once we control for it in the regression model.

0.00) in 2021.[21] Relative to the spillover group, producers in the treatment group earn a marginally higher revenue from sales and exchanges, on average, but this difference is not statistically significant in 2020 and only significant at the 10% level in 2021 (q-value: 0.10). The spillover effect corresponds to around 13-15% of the control group mean, but is only statistically significant in 2021 according to naive p-values. The coefficient decreases and loses statistical significance in our robustness checks, but it remains always positive and of a similar magnitude.

In the bottom rows of Table 1.3, we separately analyse the effects of the intervention on the value of monetary sales and the value of cashews exchanged for rice. The results in both years show that an increase in the value of monetary sales accounts for the treatment effects on the sum of these two components. In 2021, treated producers obtain on average 95,366 XOF (q-value: 0.00) more earnings from sales of cashew nuts relative to the control group, a 26% increase relative to control group mean, also significant at the 1% level. In 2020, treated producers obtain on average 51,336 XOF (q-value: 0.05) more earnings from sales of cashew nuts relative to the control group, a 22% increase relative to control group mean, also significant at the 5% level.

We find that producers in the spillover group report higher value of cashews exchanged for rice relative to the control group across both years, though this effect is not statistically significant once we account for multiple hypothesis testing. We do not find significant differences between the treatment and the spillover group, which are small in magnitude and negative.

Overall, we find that price increase for the treated producers in 2021 and that the intervention increased the value of sales and exchanges, mostly through an increase

[21] While the average price of sales is only computed for monetary transactions, the total value of all sales and exchanges includes also the total revenue from cashew trades that were repaid in rice. We value the rice received in exchange of cashews using elicited hypothetical valuations from producers. We obtained these valuations through a specific module in our questionnaire designed to ask how much producers think the rice individuals in their village receive in exchange for cashews is worth in monetary terms.

in the value of sales across both years. These results remain broadly robust to a number of alternative specifications. The effect on prices are not driven by large outliers, as our conclusions do not change once we top-winsorize the data at the 99th or 95th percentile (Appendix Table 1.12 and 1.13). The inclusion of unbalanced controls reduces the magnitude of the effect on prices by 1-3 XOF per kg in either year, which remain positive but not statistically significant (Appendix Table 1.14). Re-clustering reduces the statistical significance of the 2021 difference between the treatment and the spillover group. We still find a positive and significant difference in prices between treatment and control group once we pool across years (Appendix Table 1.17).

We note that our effects on revenue are noisy, judging from the standard error of the coefficients. When we winzorize the total value of sales and exchanges at the 99th and 95th percentiles, respectively in Appendix Table 1.12 and 1.13, the treatment effect remain positive but smaller and not statistically significant in 2020, corresponding to about 8-10% of the control group mean, and of a similar magnitude in 2021, but not statistically different from the spillover effect. Our estimates of the effect on revenue remain broadly similar across other robustness checks, though not always robust to multiple hypothesis testing in 2020.

To better understand the impact of treatment on average producer prices, we plot the cumulative distribution of prices across treatment arms in Figure 1.4 for 2020 and 2021. We can see that price effects are hard to detect in 2020. However, in 2021, our effects appear to be driven by the top half of the distribution of prices. In particular, treated producers are more likely to sell at the median price of 350 XOF/kg relative to other treatment groups. There are no differences across groups in the likelihood of reporting no sales at all, as hinted by the left-most vertical bar in the figure.

Table 1.3: Results — Prices and revenue

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treatment-Spillover	Year 2 (5) Mean (SD) Total obs.	(6) Treatment	(7) Spillover	(8) Treatment-Spillover
Average price (XOF/kg)	295.71	2.07	2.33	-0.26	355.77	6.89**	1.32	5.57*
	(104.36)	(3.46)	(3.35)	(3.30)	(83.13)	(3.41)	(3.37)	(3.05)
	1587	[0.55]	[0.49]	[0.94]	1704	[0.06]*	[0.69]	[0.10]
Value of all sales and exchanges (XOF)	287786	61650**	44747	16904	440724	105533***	58183**	47351*
	(315353)	(25296)	(26607)	(22543)	(461069)	(30721)	(28816)	(26518)
	1520	[0.05]**	[0.19]	[0.61]	1686	[0.00]***	[0.13]	[0.10]
Value of all sales (XOF)	233495	51336**	24129	27207	364477	95366***	41746*	53620**
	(292722)	(22678)	(23274)	(19951)	(428369)	(26633)	(25166)	(24217)
	1591	[0.05]**	[0.40]	[0.61]	1705	[0.00]**	[0.13]	[0.10]
Value of exchanges (XOF)	52326	10969	17252*	-6282	75662	9775	15941*	-6165
	(88038)	(8146)	(9739)	(7314)	(111393)	(8464)	(8616)	(7953)
	1635	[0.24]	[0.19]	[0.61]	1710	[0.25]	[0.13]	[0.44]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Column 4 and column 8 test for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

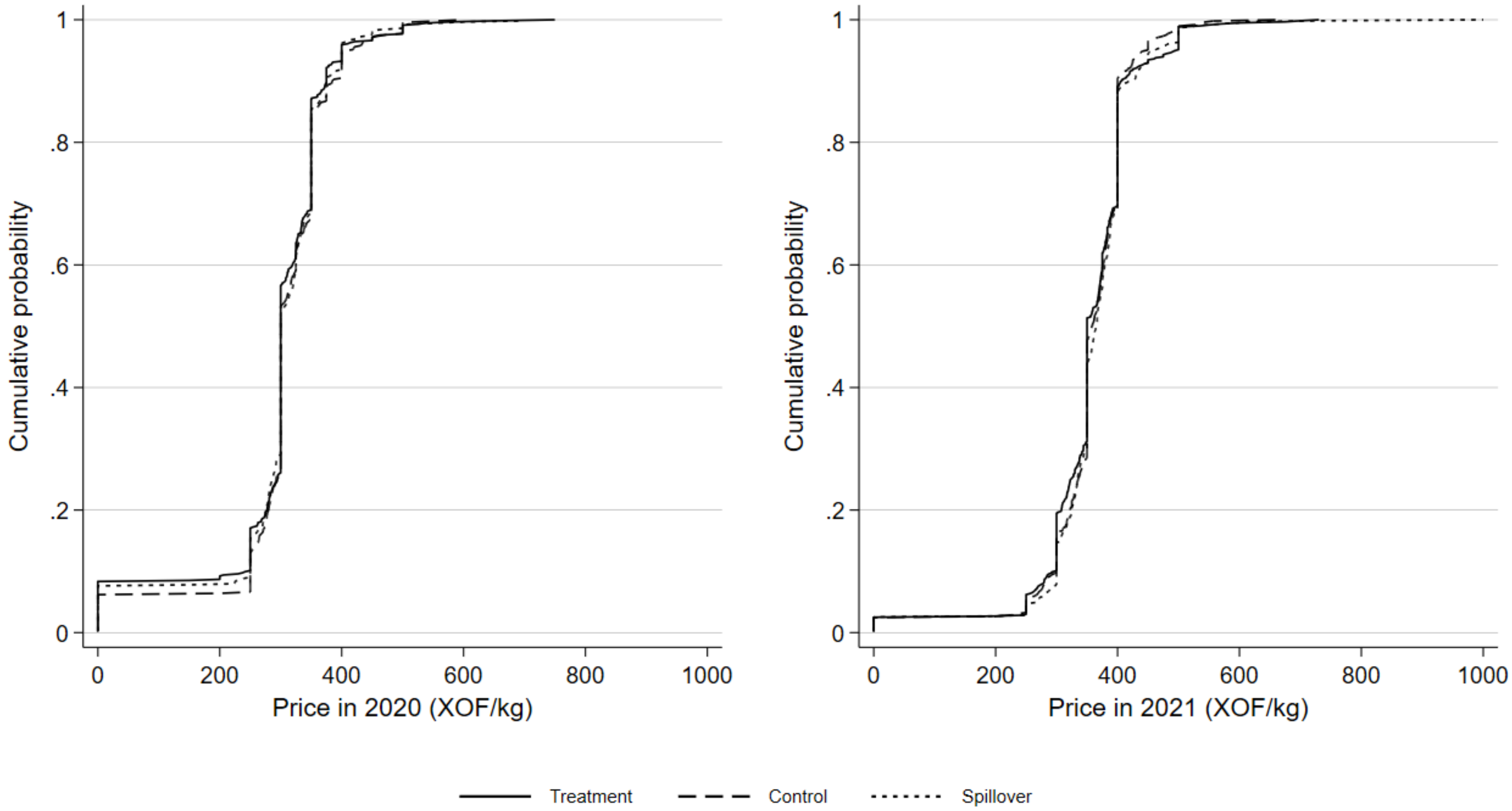


Figure 1.4: Distribution of average producer prices by treatment arm[a]

[a] Source: In-person (2020) and phone-based follow-up (2021) surveys. Prices (average across all sales for each producer) are reported in nominal West African CFA francs (XOF). We recode the price to be equal to zero for producers that did not conduct any sales.

1.5.2 Quantity sold or exchanged

To unpack our results on price and revenue, we next analyse the producers' quantities of cashews sold or exchanged, on the extensive and intensive margins. Changes in the amounts of cashews sold or exchanged can help us understand our effects on revenue, given the modest effects on average prices. Indeed, in Table 1.4 we see that treatment group report selling on average 142 kg (q-value: 0.05) in 2020 and 211 kg (q-value: 0.00) in 2021 more than those in the control group, about a 20% increase relative to the control group mean.[22] We also find statistically significant differences between the control and spillover producers in the amount of cashews sold in 2021, but not in 2020. The spillover effects are positive and lower than our treatment effects.

Our intervention induced producers to sell more frequently their cashews relative to the control and spillover groups, consistently across both years. Treated producers sell their cashews more times during the season, with the average number of sales being 0.20 higher (q-value: 0.00) relative to the control group, which represents 12-14% of the control group mean. This treatment effect is statistically significant at the 1% level and is robust to our alternative specifications. The difference between the treatment and the spillover group is also robust to our alternative specifications and remains statistically significant (q-value: 0.08-0.10). This behaviour is consistent with the advice provided by our intervention, which advised producers to sell cashews in multiple sales, given an expected increase in prices in the latter half of the trading season. We only find an increase in the number of sales among producers in the spillover group in 2021. The estimated spillover effect is positive, smaller than for the treatment group but not statistically significant after controlling for multiple hypothesis testing.

Our intervention does not have a large effect on the quantity of cashew exchanged

[22] In our robustness specifications the coefficients on quantity sold remain smaller but broadly robust in 2021, whereas in 2020 they also decrease and remain statistically significant at the 10% according to naive *p*-values.

for rice relative for treated producers to control producers, and the same is true for producers in the spillover group. The total quantity of cashew exchanged for rice is somewhat larger for both treated and spillover producers, more so for the latter. Spillover producers report an extra 70 kg of cashews exchanged for rice relative to the control group in 2021 (q-value: 0.07) and 26 kg more relative to the treated producers (q-value: 0.53). We find no treatment or spillover effects of the intervention on the share of the quantity sold over the quantity exchanged and sold.

1.6 Mechanisms

In this section, we explore mechanisms that can explain our results. First, we show some evidence consistent with our intervention having increased the bargaining power of producers. Second, we test for *between*-cluster spillovers to see whether our effects on prices may have been attenuated by the information affecting control producers, although we had not pre-specified this analysis. Third, we try to understand whether producers changed the timing of their sales or their knowledge about relevant information on the cashew market as a result of our messages. Finally, we explore alternative behavioural explanations to try to account for the increase in the quantity sold and to try to rule out whether our results may be explained by changes in outcomes where we would not expect to see any results *a priori* — a set of placebo outcomes.

1.6.1 Bargaining power

We find evidence that an increase in bargaining power can explain why prices increased. Four findings support this mechanism.

First, we find that producers, especially in 2021, report having been able to negotiate a higher price than the one they were originally offered (top panel, Table 1.5). We

Table 1.4: Results — Quantities sold and exchanged

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treatment-Spillover	Year 2 (5) Mean (SD) Total obs.	(6) Treatment	(7) Spillover	(8) Treatment-Spillover
Total quantity sold (kg)	708.21	141.75**	63.53	78.21	981.13	211.49***	93.73	117.77**
	(855.14)	(61.11)	(62.11)	(54.72)	(1163.13)	(64.24)	(61.88)	(54.62)
	1622	[0.05]*	[0.45]	[0.38]	1709	[0.00]***	[0.22]	[0.08]*
Number of sales	1.49	0.21***	0.07	0.14**	1.66	0.20***	0.08*	0.11**
	(1.01)	(0.06)	(0.06)	(0.06)	(0.92)	(0.05)	(0.05)	(0.05)
	1693	[0.00]***	[0.45]	[0.10]*	1730	[0.00]***	[0.21]	[0.08]*
1 if exchanged cashew for rice	0.56	-0.01	-0.03	0.03	0.56	-0.01	-0.00	-0.01
	(0.50)	(0.03)	(0.03)	(0.03)	(0.50)	(0.03)	(0.03)	(0.03)
	1706	[0.85]	[0.45]	[0.59]	1724	[0.60]	[0.99]	[0.58]
Total quantity exchanged (kg)	217.70	19.62	20.91	-1.30	231.72	44.46	70.75**	-26.28
	(323.48)	(24.38)	(25.70)	(23.10)	(346.31)	(27.70)	(28.41)	(26.38)
	1681	[0.53]	[0.45]	[0.96]	1707	[0.18]	[0.07]*	[0.53]
Share of quantity sold over quantity exchanged and sold	0.73	0.02	0.02	0.00	0.79	0.01	-0.00	0.01
	(0.32)	(0.02)	(0.02)	(0.02)	(0.25)	(0.02)	(0.02)	(0.01)
	1604	[0.53]	[0.45]	[0.96]	1691	[0.60]	[0.99]	[0.57]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Column 4 and column 8 test for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

see that treated producers are 7% more likely to report having gained a higher price than the one originally offered relative to control producers, though this effect is only significant at the 10% using naive *p*-values. We find no evidence that the treatment induced producers to act as intermediary buyers. In our control group, 7% of producers report having bought cashews from other producers for resale. However, we find that producers in the spillover group reported receiving offers from more buyers relative to the control group in 2020. The difference between the treatment and spillover group is negative but not statistically significant after accounting for multiple-hypothesis testing. A possible explanation for the latter pattern may be that the service induced some traders to purchase cashew nuts from other producers in the same village, that were not as informed about market developments as the treated ones. On average the control group reported receiving offers from about 3 potential buyers. The spillover producers received offers from an extra 0.3 buyers (q-value: 0.01) relative to the control group, an increase of 10% of the control group mean.

Second, treated producers change the allocation of their sales to different buyers (middle panel, Table 1.5). Treated producers sell less to local buyers and more to itinerant buyers, especially in 2021. Treated producers sell, on average, 6 percentage points more of their output to itinerant buyers. Those type of buyers are more likely to self-finance their operations and have more room to negotiate with producers over a price. In 2020, we also see that treated producers sold 4 percentage points more of their output to itinerant buyers, but this effect is not statistically significant.[23]

Third, we find that producers that receive our information change the source of information about market conditions. They move away from intermediaries as a source of information and they start to rely more on the messages that we have sent them through

[23] We deviate slightly from our analysis plan in reporting the shares sold to different buyers, as opposed to the pre-specified indicator variables equal to one if the producer sold the majority of the stock to a specific type of buyer. In Appendix Table 1.18 we present our pre-specified outcomes within this family of variables. Our conclusions remain qualitatively similar, though not robust to multiple hypothesis testing with the pre-specified measures.

their mobile phones (bottom panel, Table 1.5). In 2020, treated producers report being 4-6 percentage points (q-value: 0.00) less likely to rely on intermediaries as a source of market advice. We interpret this switch in source of information alongside the change in the type of buyer as evidence that the messages changed producers' negotiation strategy. We had not pre-specified this family of outcomes and only recorded this data in the in-person follow-up, so we can only speculate that the change in this behaviour might have persisted also in 2021.

Fourth, we find that our positive effects are larger for producers with more bargaining power to begin with. In Figure 1.5 we plot heterogeneous treatment effects on our two focal outcomes, average price and quantity sold across the two years. We find suggestive evidence that our effects on prices are larger for producers with cashew plots that are in the top tercile of the distribution (row 2, column 3, Figure 1.5). Producer with larger plots are those with larger amounts to sell, which may be able to negotiate harder for deals than those with fewer bags of cashew nuts to sell. There is no clear difference in treatment effects or consistent pattern based on whether producers were able to read, their level of assets, or the distance with the closest urban market (in the administrative capital of the sector), or the number of other producers in the village.

1.6.2 Between-cluster spillovers

To further unpack our main results, we explore whether spillover effects may have occurred across villages. While our sampling design aimed to minimize between-cluster spillovers by maximising distances between sampled villages, market information could still flow between closely connected villages, potentially influencing how itinerant traders negotiate prices across nearby areas (Soldani et al., 2023; Falcao Bergquist et al., 2021). We test for the presence of *between*-village using a similar approach to Miguel and Kremer (2004) and Egger et al. (2022), which we describe in detail in Ap-

Table 1.5: Mechanisms — Increased bargaining power

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treatment-Spillover	Year 2 (5) Mean (SD) Total obs.	(6) Treatment	(7) Spillover	(8) Treatment-Spillover
Panel A: Bargaining power								
Number of succesful bargains	0.14	-0.00	-0.02	0.01	0.46	0.07*	0.07	0.00
	(0.37)	(0.02)	(0.02)	(0.02)	(0.71)	(0.04)	(0.05)	(0.05)
	1690	[0.89]	[0.44]	[0.57]	1729	[0.29]	[0.41]	[0.95]
1 if acted as intermediary	0.07	0.01	0.01	-0.01	0.06	0.01	0.02	-0.01
	(0.26)	(0.01)	(0.02)	(0.02)	(0.23)	(0.01)	(0.02)	(0.02)
	1713	[0.89]	[0.44]	[0.57]	1725	[0.89]	[0.41]	[0.95]
Number of potential buyers that made offers	2.91	0.08	0.30***	-0.21**	3.18	0.01	0.05	-0.04
	(1.82)	(0.09)	(0.10)	(0.10)	(2.02)	(0.10)	(0.11)	(0.12)
	1671	[0.89]	[0.01]***	[0.12]	1703	[0.89]	[0.63]	[0.95]
Panel B: Share of cashews sold to...								
...a local buyer	0.47	-0.02	0.02	-0.05	0.57	-0.05*	-0.03	-0.02
	(0.49)	(0.03)	(0.03)	(0.03)	(0.48)	(0.03)	(0.03)	(0.03)
	1494	[0.47]	[0.69]	[0.25]	1651	[0.08]*	[0.56]	[0.54]
...an itinerant buyer	0.47	0.04	-0.01	0.04	0.41	0.06**	0.03	0.03
	(0.49)	(0.03)	(0.03)	(0.03)	(0.48)	(0.03)	(0.03)	(0.03)
	1494	[0.37]	[0.83]	[0.25]	1651	[0.08]*	[0.56]	[0.43]
...the wholesale market	0.05	-0.03**	-0.02	-0.01	0.01	-0.01	0.00	-0.01
	(0.20)	(0.01)	(0.01)	(0.01)	(0.11)	(0.01)	(0.01)	(0.01)
	1494	[0.06]*	[0.55]	[0.31]	1651	[0.42]	[0.62]	[0.43]
Panel C: Received market advice from...								
...family and friends	0.49	0.02	-0.02	0.04				
	(0.50)	(0.03)	(0.03)	(0.03)				
	1713	[0.46]	[0.57]	[0.20]				
...intermediaries	0.17	-0.06***	-0.02	-0.04**				
	(0.37)	(0.02)	(0.02)	(0.02)				
	1713	[0.00]***	[0.57]	[0.05]**				
...mobile messages	0.04	0.15***	0.01	0.14***				
	(0.19)	(0.02)	(0.01)	(0.02)				
	1713	[0.00]***	[0.57]	[0.00]***				
...other sources	0.12	-0.04***	-0.02	-0.02				
	(0.33)	(0.02)	(0.02)	(0.01)				
	1713	[0.01]***	[0.57]	[0.20]				

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Column 4 and column 8 test for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it is available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

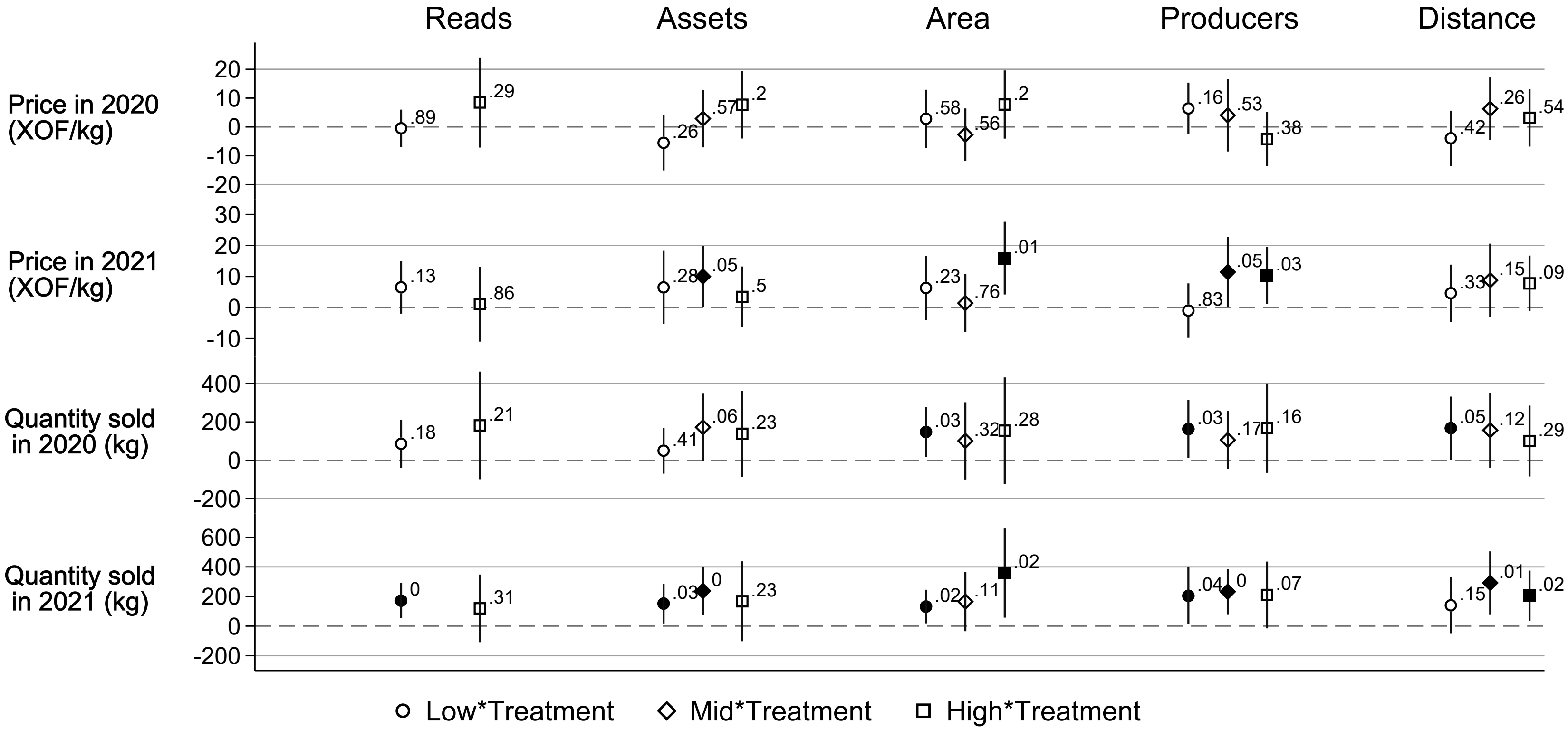


Figure 1.5: Treatment heterogeneity in price and quantity sold[a]

[a] Source: In-person (2020) and phone-based follow-up (2021) surveys. Each cell represents coefficients from a separate OLS regression. For continuous dimensions of heterogeneity, we trichotomize the variable — interacting in each case our treatment indicator with (i) an indicator equal to 1 if the baseline value lies at or below the lower tercile, (ii) an indicator equal to 1 if the baseline value lies strictly between the two terciles, and (iii) an indicator equal to 1 if the baseline values lies at or above the upper tercile, following Bari ⓡ, Faisal and Malik ⓡ, Kashif and Meki ⓡ, Muhammad and Quinn ⓡ, Simon (2021). For the first column (Reads), we interact treatment with an indicator variable equal to one if the respondent could read a sentence in Bissau-Guinean Kriol. Assets heterogeneity is based on the respondent's principal component of assets (Filmer and Pritchett, 2001). Area heterogeneity is based on the respondent's size of the cashew tree plot (in hectares). Producers heterogeneity is based on the number of producers in the respondent's village. Distance heterogeneity is based on the road distance to the nearest sector capital (in km). All regressions control for randomisation triplet fixed effects and indicators for the middle and upper tercile of the heterogeneity dimension. Standard errors are clustered at the village-level. Bars represent 95% confidence interval based on standard p-values. Filled markers imply statistical significance at the 5% level relative to the control group in the bottom tercile of the heterogeneity dimension, the omitted category.

pendix Section 1.10.6. Specifically, we test whether the number of treated individuals within a radius of 5 km, conditional on the number of villages in our sample within that radius, significantly changes prices and quantities sold beyond the producer-level treatment assignment. Identification relies on our village-level randomisation— the number of treated producers within a 5 km radius should be plausibly exogenous after controlling for the number of villages in our study within a 5 km radius. Appendix Figure 1.8 shows how treatment intensity varies spatially in our sample.[24]

We find evidence of *between*-village spillovers, with prices being higher in more intensely treated areas than those without any treated producers (white markers in Figure 1.6). In 2020, our estimate of the effect of an additional treated producer within a 5 km is similar to our direct estimate of treatment effects, and is statistically significant at the 10% level. The magnitude of these spillovers on prices corresponds to a 0.7% increase for every treated producer within a radius of 5 km. To unpack these effect across our experimental arms, we interact the number of treated producers within 5 km and the indicators for whether producers are in either the treatment or spillover group (black markers in Figure 1.6). The *between*-cluster spillovers are driven by producers in control villages that neighbour treatment villages. These producers report an increase of about 2-4 XOF (1.3% of the control group mean) for every additional individual treated within a 5 km radius, relative to control group producers without any treated individual within a 5 km radius. The direct effects in treated villages that are not close to other treated villages are positive and statistically significant at the 10% level. Assuming a linear relationship between the number of treated individuals within a 5 km radius and the price, these results imply an increase of around 4-8% in control villages that were more intensely treated. Our positive *between*-village spillover estimates

[24] Specifically, 40% of villages have another sampled village within a 5 km radius, 9% of villages have two sampled villages, 1% had three sampled villages, and the remaining 50% had none. 27% of villages have another treated village within a 5 km radius, 8% of villages have two treated villages, and the remaining 65% had none.

imply that our intention-to-treat estimates may be a lower-bound of the overall effects of the intervention. Our estimates of *between*-village effects in 2021 are smaller than in 2020, but going in the same direction as the previous year. Finally, in the bottom half of Figure 1.6, we do not find effects of treatment intensity on the quantity sold. Our findings are consistent with buyers not being able to distinguish between treated and untreated producers and hence adapting their behaviour towards all producers, for instance by making better price offers across villages that are close to each other, as suggested by Fafchamps and Minten (2012).[25]

Finally, we suggest three plausible reasons to reconcile our estimates in 2021 and 2020. First, the direct effects may have been greater in 2021 as a result of learning among treated producers, who use the information better in the second year. Second, it is plausible that since in 2021 all producers in the study had heard about the service, and some of the control group producers had also subscribed, our indirect effects may be attenuated. Third, if the indirect effects are influenced by buyer behaviour, there may also have been a learning process among buyers, and they could have started to differentiate between informed and uninformed producers better, leading to reduced spillover effects. In aggregate, these three reasons may help explain some of the small differences across years, which are nonetheless remarkably similar.

1.6.3 Timing of sales

A mechanism that could explain our results is that producers tried to change the timing of their sales in order to sell when prices were highest. Instead, we find that treated producers were more likely to sell during the first half of the trading season (in, or before, April and May) relative to the control group (Table 1.6). This period coincided with the time when we initially started sending messages to treated producers, though

[25] In line with this hypothesis, in Appendix Table 1.28, we found that control villages reported receiving more itinerant traders than treated villages.

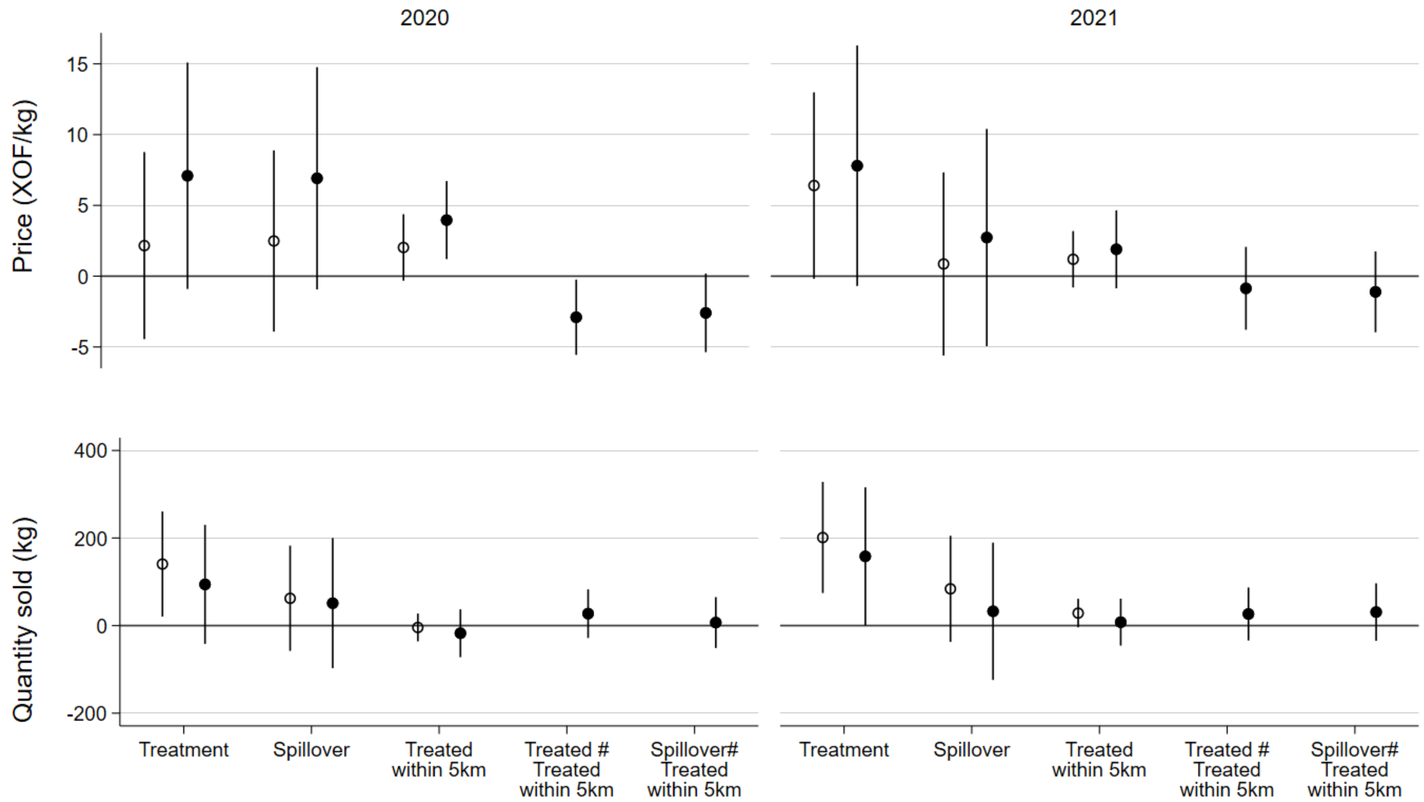


Figure 1.6: Spatial spillovers on prices and quantity sold[a]

[a] Source: In-person (2020) and phone-based follow-up (2021) surveys. Each sub-plot represents coefficients from two separate OLS regression where the unit of observation is the producer. The white coefficients are estimate from a regression that control for randomisation tripled fixed effects, the number of villages within a 5 km radius that are in our sample, and the baseline value of the outcomes. The black coefficients are estimated from a model where we add to the previous regression the interaction terms between the number of treated producers within 5 km and the treatment (and spillover) indicators. The radius of 5 km 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. Bars represent 95% confidence interval based on Conley (1999) standard errors, accounting for spatial correlation within a 5 km radius.

the messages advised producers to wait to sell, as prices were expected to increase. In 2020, treated producers are 4 percentage points more likely to sell in April relative to the control group, though this effect is not robust to multiple hypothesis testing. Whereas in 2021, they increased the probability of recording a sale in April or May by 7 percentage points (q-value: 0.02-0.03). In 2020, treated producers were 3 percentage points less likely to report exchanging cashew for rice towards the beginning of trading season and more likely to do so in July or later (when the terms of exchange had improved). Qualitatively, we see a similar pattern in 2021, though it is not statistically significant and effect sizes are smaller. The difference in timing of sales is consistent with treated producers selling more than once during the trading season.

Producers do not sell more frequently in high price periods, but they do report higher prices in those periods in 2021. To convey this finding descriptively, Figure 1.7 plots the interaction between our treatment indicator and an indicator for the month of sales, running the regression at the transaction-level. While there is no difference in the average effect by month of sales in 2020, prices for treated producers relative to control producers are almost 20 XOF per kg higher in June 2021, which was the month when prices peaked. In the bottom right quadrant, we also see that treated producers sold more on the intensive margin in June 2021 relative to control producers in the same month, though this effect is not statistically significant.

To understand what motivated producers to sell more than once, we used an open-ended question to elicit their reasons for adopting this sale strategy. In the first three rows of Appendix Table 1.19, we report the producers' responses by treatment status.[26] Treated producers were 3 percentage points more likely than the control group to directly refer to the market information system as a reason for selling more than once. Treated producers were also 7 percentage points more likely than the control group

[26] We had not pre-specified the analysis of the outcomes in this table, which we only collected during the in-person survey.

to say they sold more than once because they thought that prices would increase in the future and because they thought this strategy would reduce their risk of price uncertainty. This reasoning is consistent with the information delivered by the market information system, which suggested that prices would increase in the second half of the trading seasons. Moreover, we see that treated producers are also 5 percentage points more likely to sell more than once because they report that this enabled them to smooth consumption (i.e. to avoid spending all the revenue at once or because they lacked liquidity) relative to the control group. In the second panel of the same table, we also see that treated producers report being 5-8 percentage points more likely to have sought advice on market timing, significantly so relative to the spillover group.

1.6.4 Alternative behavioural mechanisms

We do not find evidence consistent with several other behavioural mechanisms. Firstly, we find no evidence that treated producers changed their beliefs about the market (Appendix Table 1.20). Producers, on average, expect high prices for the upcoming trading season. Treated producers had lower and more realistic expectations about their sale price. However, these differences were not statistically significant and noisy. Secondly, we did not observe an increase in information sharing among producers as a result of the intervention. Specifically, we did not find that treated producers were more likely to share or receive information about offers. Thirdly, we investigated various outcomes where we did not anticipate any effects (Appendix Table 1.21). In 2020, we found no effects on those particular outcomes, such as producers being more likely to record their sales using a transaction diary or their (hypothetical) risk aversion.[27] However, we did observe a slight increase in the number of treated producers reporting greater trust, which may be a secondary effect of the intervention for the producers

[27] We elicited risk aversion using an hypothetical choice between two lotteries, as in (Binswanger, 1980).

who benefited. Fourthly, we provide evidence from producers' transaction diaries and interviews conducted with a subset of intermediaries (Appendix Section 1.13 and 1.14), which aligns qualitatively with the producers' interview responses.

Additionally, we explored alternative uses of cashews to determine whether producers used fewer cashews for other purposes. While we found lower amounts of cashews used by treated producers to repay loans (Appendix Table 1.22) or for small exchanges, these reductions were not significant enough to fully explain the increases in quantity sold (Appendix Table 1.23). Furthermore, in our phone-based follow-up survey, we examined whether the quantity increases were due to changes in production decisions made by producers but found no effects on the measures we collected. We remain cautious in interpreting our effects on the quantity sold, though they are consistent with a positive correlation between prices and quantity sold (Appendix Table 1.24).

1.7 Conclusion

This paper estimated the effects of introducing a new market information system among cashew producers in Guinea-Bissau. The market information system provided free weekly text and voice-messages to treated producers during the 2020 and 2021 trading season. The information sent to producers contained up-to-date farmgate prices, market news, and sales advice on when to sell. On the whole, we see that treated producers benefited from these messages and were able to negotiate higher prices on average. These effects are larger for producers that had more output to sell. We find that *between*-village spillovers of the intervention could attenuate the intention-to-treat effects on average prices among treated producers relative to control producers. We speculate that these spillovers may occur through itinerant traders changing their negotiation strategy, as in Soldani et al. (2023).

Our findings yield several policy implications. First, providing up-to-date and reliable

Table 1.6: Mechanisms — Changes in timing of sales

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treatment-Spillover	Year 2 (5) Mean (SD) Total obs.	(6) Treatment	(7) Spillover	(8) Treatment-Spillover
1 if sold cashews in April or before	0.15	0.04**	0.02	0.02	0.17	0.07***	0.05*	0.03
	(0.35)	(0.02)	(0.02)	(0.03)	(0.38)	(0.02)	(0.03)	(0.03)
	1361	[0.17]	[0.80]	[0.99]	1477	[0.02]**	[0.25]	[0.89]
1 if sold cashews in May	0.51	0.02	-0.01	0.03	0.49	0.07***	0.06*	0.02
	(0.50)	(0.03)	(0.04)	(0.03)	(0.50)	(0.03)	(0.03)	(0.03)
	1361	[0.71]	[0.88]	[0.99]	1477	[0.03]**	[0.25]	[0.89]
1 if sold cashews in June	0.49	0.05	0.05	-0.00	0.64	-0.02	-0.03	0.01
	(0.50)	(0.03)	(0.03)	(0.03)	(0.48)	(0.03)	(0.03)	(0.03)
	1361	[0.29]	[0.33]	[0.99]	1477	[0.84]	[0.84]	[0.89]
1 if sold cashews in July or later	0.07	0.01	0.01	0.00	0.12	0.02	0.00	0.02
	(0.25)	(0.02)	(0.02)	(0.02)	(0.33)	(0.02)	(0.02)	(0.02)
	1361	[0.71]	[0.81]	[0.99]	1477	[0.74]	[0.84]	[0.89]
1 if exchanged rice in April or before	0.09	-0.03*	-0.03*	0.00	0.10	-0.01	-0.00	-0.00
	(0.28)	(0.01)	(0.02)	(0.02)	(0.29)	(0.02)	(0.02)	(0.02)
	1591	[0.17]	[0.21]	[0.99]	1662	[0.84]	[0.84]	[0.93]
1 if exchanged rice in May	0.23	-0.01	-0.00	-0.00	0.22	0.00	0.02	-0.01
	(0.42)	(0.02)	(0.03)	(0.02)	(0.42)	(0.03)	(0.03)	(0.03)
	1591	[0.81]	[0.90]	[0.99]	1662	[0.95]	[0.84]	[0.89]
1 if exchanged rice in June	0.26	-0.01	-0.02	0.01	0.27	-0.01	-0.02	0.01
	(0.44)	(0.03)	(0.03)	(0.02)	(0.44)	(0.02)	(0.03)	(0.03)
	1591	[0.76]	[0.80]	[0.99]	1662	[0.84]	[0.84]	[0.89]
1 if exchanged rice in July or later	0.01	0.03***	0.03***	0.00	0.05	0.01	0.01	0.00
	(0.11)	(0.01)	(0.01)	(0.01)	(0.21)	(0.01)	(0.01)	(0.01)
	1591	[0.01]**	[0.05]*	[0.99]	1662	[0.84]	[0.84]	[0.89]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

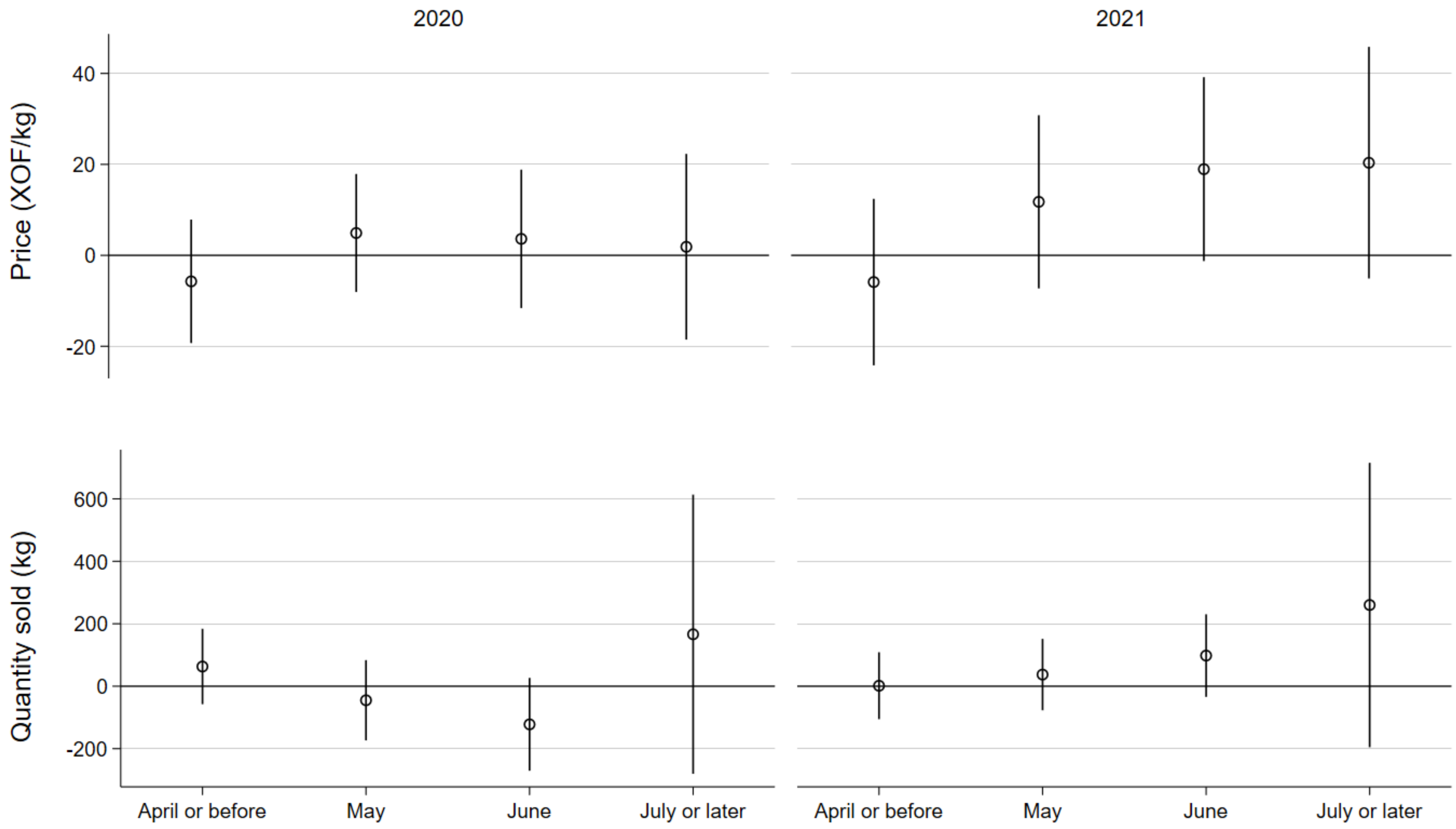


Figure 1.7: Treatment by month of transactions in price and quantity sold[a]

[a] Source: In-person (2020) and phone-based follow-up (2021) surveys. Each sub-plot represents coefficients from a separate OLS regression where the unit of observation is a transaction. All regressions control for randomisation triplet fixed effects. Bars represent 95% confidence intervals.

information on market conditions has improved market outcomes for producers in this context. Subscribers to the service could quickly repay the user-fee if they earned similar additional returns as those we observed among treated producers. Second, the government policy of providing a reference price only at the beginning of the trading season can be improved upon, by communicating price updates more frequently.

There are several open questions that remain unanswered. As the service continues to be rolled-out and its user-base grows, the effects we estimated may change, as more producers become better informed about the market dynamics. We plan to explore some of these issues in future research.

Appendix

This appendix has six sections. In Section 1.8 we provide details of the sampling and randomisation procedures we implemented. In Section 1.9 we provide balance tables and attrition. Section 1.10 shows robustness of our main results to alternative econometric specifications. It also illustrates effects of the intervention on additional outcomes not reported in the main text. In Section 1.11, we provide additional details on the cashew nut production and commercialisation. In Section 1.13 and Section 1.14 we show complementary results based on a sub-sample of producers that kept written records of their transactions and from intermediaries based in a sub-sample of villages.

1.8 Sampling and randomisation

1.8.1 Village-level sampling

Our sampling strategy has two main goals: (i) making the treatment and control groups statistically balanced across a set of baseline characteristics, and (ii) minimising spillovers of the treatment to untreated villages. There is a trade-off between these two goals: minimising spillovers requires that treatment and control groups be sufficiently far apart geographically; while statistical balance requires that treatment and control groups be similar to each other, which in turn often requires geographical proximity (Duflo et al., 2008). We address this trade-off as follows.

In order to select villages that would be part of our study, we created a grid containing $2.7km^2$ cells, covering the entirety of the country. We sample 290 villages from about 1,800 villages, using geo-coded census data, such that one village is chosen from each of those cells.[28] We compute for each village the closest distance to another sampled village. We store the value of the minimum distance in the sample of villages drawn. We iterate these steps 999 times and choose the sample that has the largest minimum distance.

We excluded from our sampling frame villages that have the following characteristics:

- with fewer than ten households according to the 2009 census (to allow a sufficiently large sample of cashew producers in every village),
- located on the islands (due to budget constraints),
- located in a circle of 15km in radius in the region of Tombali, which are closest to the river estuary (due to budget constraints, as reaching those villages is logistically difficult),

[28] We constructed our sampling frame by merging the 2009 census data, the latest census available to-date, with GPS coordinates obtained from a geocoded administrative dataset maintained by the United Nations Office for the Coordination of Humanitarian Affairs.

- located in the sector of Boé, in the East of the country (due to budget constraints and poor phone signal in the area),
- for which we did not have reliable GPS and population data (so those unmatched from the fuzzy merging of the census and the GPS data),
- located in Biombo, the region to the West of the capital Bissau (as it was used for the pilot).

These exclusions account for at most 40% of producers in the country, mostly driven by exclusion of the smaller villages. We posit that since most of the villages that were not included in the sampling frame are more remote and smaller, producers living in those excluded villages are likely to be most affected by lack of information and could benefit more from the intervention. Once the final sample of villages was selected, we sought the contact details of the village leaders of our final sample of villages. Whenever possible, we contacted the village leaders ahead of the data-collection baseline visit to inform them of the study and seek their collaboration.

1.8.2 Producer-level sampling

After the data-collection team reached the sampled village and the village leaders granted them permission to work, they asked the village leader for a list of all producers living in the village owning a cashew tree plot. In eliciting this list, the data-collection team stressed that every producer with a cashew tree plot should be included, including small ones. On the day of the visit, the data-collection team used a random number generator to sample seven producers from this list.

1.8.3 Village-level randomisation

We follow the recommendations of Bruhn and McKenzie (2009), stratifying our randomisation to increase efficiency. Randomisation of treatment across villages was implemented constructing, in each region, triplets of villages that are as similar as possible along a number of dimensions that are likely to affect the impact of treatment. We combine the population data from the 2009 census and other geo-coded databases to construct the following variables:

(i). road distance to the nearest sectoral capital (to proxy the closest location where small intermediaries are based)

(ii). road distance to the nearest regional capital (to proxy the closest location where wholesale intermediaries are based)

(iii). road distance to the port in Bissau (to proxy for distance from the main export route, and inversely from distance to the closest borders)

(iv). linear distance to the nearest (MTN) mobile network tower, and

(v). number of households for each village, according to the 2009 census.

Because these characteristics are likely to be correlated, we use the Mahalanobis distance as a metric of similarity across villages that takes into account the correlation across these characteristics, as in Fafchamps and Minten (2012). The Mahalanobis distance between them is then defined as:

$$||z_l - z_j|| = ((z_l - z_j)' S^{-1} (z_l - z_j))^{1/2} \tag{1.4}$$

where z_l and z_j denote the vector or relevant characteristics from villages l and j, respectively, and S is the covariance matrix of characteristics z. Pairs of villages with a smaller Mahalanobis distance are more similar along these dimensions. Since these characteristics are weighted by the inverse of the covariance matrix S, correlation between characteristics — e.g. between our various distance measures — is given less weight.

We select, within each region, the allocation of villages into triplets that minimises the sum, over all triplets, of the Mahalanobis distances within each triplet. The search is conducted using an algorithm that randomly tries different combinations of villages into triplets. Within each triplet, one village was then randomly assigned to control and two villages to treatment. This ratio of treated to control villages yields a similar number of individual producers that are either treated, spillover, or control, in each triplet.

1.8.4 Within-village randomisation

Once the seven randomly sampled producers in treatment villages completed the baseline interview, an on-the-spot within-village lottery determined producer-level assignment to either the treatment or spillover groups. The supervisor of the data-collection team administered the lottery. Each producer drew one of two kinds of goodies from a bag. Using this lottery, four producers were assigned to the treatment group and three to the spillover one. After the producer-level randomisation, the treated producers took part in a training session about the intervention and the determinants of farmgate raw cashew-nut prices.

1.9 Experimental integrity

1.9.1 Balance

Table 1.7: Baseline balance — Primary outcome variables

	Control (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treat. vs. spillover	(5) Max pairwise st. diff.
Average price (XOF/kg)	395.89	-3.66	1.35	-5.01	0.05
	(96.24)	(5.51)	(5.60)	(4.75)	1968
	1968	[0.74]	[0.81]	[0.50]	
Value of all sales and exchanges (XOF)	527100.44	35669.54	24404.24	11265.30	0.03
	(921775.27)	(44414.13)	(43399.62)	(29831.03)	1978
	1978	[0.74]	[0.80]	[0.88]	
Value of all sales (XOF)	486696.38	14815.85	15806.13	-990.28	0.01
	(912168.16)	(43992.64)	(43245.56)	(28614.91)	1984
	1984	[0.74]	[0.80]	[0.97]	
Value of exchanges (XOF)	39662.45	21556.26***	10321.69*	11234.56**	0.22
	(85677.82)	(6797.60)	(6024.35)	(5681.82)	1982
	1982	[0.01]**	[0.46]	[0.22]	
Total quantity sold (kg)	1206.35	54.39	36.52	17.87	0.01
	(1855.65)	(96.73)	(94.05)	(66.13)	1984
	1984	[0.74]	[0.80]	[0.88]	
Number of sales	1.62	0.01	0.07	-0.05	0.08
	(0.81)	(0.04)	(0.05)	(0.04)	1988
	1988	[0.74]	[0.46]	[0.50]	
1 if exchanged cashew for rice	0.30	0.04	0.02	0.02	0.15
	(0.46)	(0.03)	(0.03)	(0.02)	1985
	1985	[0.30]	[0.80]	[0.50]	
Total quantity exchanged (kg)	113.66	57.26***	25.75	31.51**	0.21
	(247.43)	(19.08)	(17.06)	(15.94)	1982
	1982	[0.01]**	[0.46]	[0.22]	
Share of quantity sold over quantity exchanged and sold	0.90	-0.02	-0.01	-0.02	0.13
	(0.20)	(0.01)	(0.01)	(0.01)	1979
	1979	[0.30]	[0.80]	[0.44]	

Notes: Coefficient of treatment and spillover at baseline (columns 2-3). Column 4 tests for differences in parameters obtained in previous two columns. Column 5 reports the standardised pairwise maximum difference between mean across all study groups. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. q-values reported in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 displays the control mean; standard deviation; and total number of observations.

Table 1.8: Baseline balance — Producer characteristics

	(1) Control Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treat. vs. spillover	(5) Max pairwise st. diff.
1 if female	0.05	0.00	-0.00	0.00	0.02
	(0.23)	(0.01)	(0.01)	(0.01)	1985
	1985	[0.95]	[1.00]	[0.97]	
Age	43.47	-1.06	-1.11	0.05	0.06
	(14.96)	(0.79)	(0.77)	(0.75)	1988
	1988	[0.37]	[0.60]	[0.97]	
Household size	13.40	0.57	1.02*	-0.45	0.08
	(7.88)	(0.54)	(0.57)	(0.41)	1988
	1988	[0.45]	[0.38]	[0.92]	
Years of education	3.89	0.01	-0.10	0.11	0.04
	(3.77)	(0.23)	(0.23)	(0.18)	1988
	1988	[0.95]	[1.00]	[0.92]	
1 if Kriol is the most spoken language at home	0.21	-0.01	-0.01	0.00	0.05
	(0.41)	(0.02)	(0.02)	(0.02)	1988
	1988	[0.87]	[1.00]	[0.97]	
1 if cashew is the main source of income	0.79	0.02	0.00	0.02	0.06
	(0.41)	(0.02)	(0.02)	(0.02)	1988
	1988	[0.45]	[1.00]	[0.92]	
1 if faces storage limitations	0.27	0.06**	0.02	0.04*	0.15
	(0.45)	(0.03)	(0.02)	(0.02)	1988
	1988	[0.08]*	[0.99]	[0.92]	
Minimum age of trees (years)	3.32	-0.06	0.08	-0.14	0.06
	(3.43)	(0.17)	(0.20)	(0.20)	1966
	1966	[0.87]	[1.00]	[0.92]	
Max age of trees (years)	17.23	-0.20	0.04	-0.25	0.04
	(8.41)	(0.47)	(0.48)	(0.48)	1868
	1868	[0.87]	[1.00]	[0.93]	
1 if trees were plagued	0.69	-0.03	-0.01	-0.03	0.08
	(0.46)	(0.03)	(0.03)	(0.03)	1980
	1980	[0.37]	[1.00]	[0.92]	
1 if sells other crops	0.65	0.09***	0.10***	-0.01	0.21
	(0.48)	(0.02)	(0.02)	(0.02)	1988
	1988	[0.00]***	[0.00]***	[0.97]	

Notes: Coefficient of treatment and spillover at baseline (columns 2-3). Column 4 tests for differences in parameters obtained in previous two columns. Column 5 reports the standardised pairwise maximum difference between mean across all study groups. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. q-values reported in brackets and are calculated across all variables reported in Appendix Table 1.8 and 1.9. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 displays the control mean; standard deviation; and total number of observations.

Table 1.9: Baseline balance — Producer characteristics (continued)

	(1) Control Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treat. vs. spillover	(5) Max pairwise st. diff.
Index of trust (Anderson, 2008)	0.00	0.07	0.07	0.00	0.06
	(1.00)	(0.05)	(0.05)	(0.05)	1988
	1988	[0.29]	[0.60]	[0.97]	
Index of food security (Anderson, 2008)	0.00	-0.02	0.01	-0.03	0.03
	(1.00)	(0.05)	(0.06)	(0.05)	1983
	1983	[0.87]	[1.00]	[0.92]	
1 if has a fixed buyer	0.30	0.05*	0.03	0.02	0.10
	(0.46)	(0.03)	(0.03)	(0.03)	1987
	1987	[0.23]	[0.83]	[0.92]	
Area plantation (hectares)	3.26	0.28**	0.27*	0.01	0.11
	(2.29)	(0.13)	(0.14)	(0.15)	1988
	1988	[0.19]	[0.38]	[0.97]	
Years of experience in cashew sector	10.17	0.65*	0.25	0.41	0.08
	(7.05)	(0.39)	(0.38)	(0.38)	1988
	1988	[0.23]	[1.00]	[0.92]	
Standardised principal component of wealth (Filmer and Pritchett, 2001)	-0.19	0.30***	0.37***	-0.07	0.20
	(1.63)	(0.10)	(0.10)	(0.09)	1979
	1979	[0.03]**	[0.00]***	[0.92]	
Index of numeracy (Anderson, 2008)	0.00	0.10*	0.02	0.08	0.11
	(1.00)	(0.06)	(0.06)	(0.05)	1988
	1988	[0.23]	[1.00]	[0.92]	
1 if can read	0.29	0.00	0.00	-0.00	0.02
	(0.45)	(0.03)	(0.03)	(0.02)	1988
	1988	[0.95]	[1.00]	[0.97]	
1 if present-biased	0.17	0.03*	0.00	0.03	0.07
	(0.37)	(0.02)	(0.02)	(0.02)	1985
	1985	[0.23]	[1.00]	[0.92]	

Notes: Coefficient of treatment and spillover at baseline (columns 2-3). Column 4 tests for differences in parameters obtained in previous two columns. Column 5 reports the standardised pairwise maximum difference between mean across all study groups. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. q-values reported in brackets are calcualted over all variables reported in Appendix Table 1.8 and 1.9.. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 displays the control mean; standard deviation; and total number of observations.

Table 1.10: Baseline balance — Village characteristics

	Control (1) Mean (SD) Total obs.	(2) Treatment
Number of cashew producers in the village	33.87	-0.68
	(23.49)	(2.83)
	288	[0.81]
Number of cashew buyers in the village	2.90	1.13***
	(2.58)	(0.41)
	287	[0.04]**
Road distance in km to nearest sector capital	25.59	1.02
	(23.70)	(2.65)
	290	[0.81]
Road distance in km to nearest region capital	51.48	-0.85
	(33.51)	(2.96)
	290	[0.81]
Road distance in km to the capital	162.30	4.15
	(69.26)	(4.60)
	290	[0.81]

Notes: Coefficient of village-level treatment at baseline (columns 2). Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the village. All models control for the randomisation triplet fixed-effects. Standard errors are in parentheses and are robust. Stars on the coefficient estimates reflect unadjusted p-values. q-values reported in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 displays the control mean; standard deviation; and total number of observations.

1.9.2 Attrition

Table 1.11: Attrition rates at either follow-up

	(1)	(2)	(3)	(4)	(5)
	Control Mean (SD) Total obs.	Treatment	Spillover	Treat. vs. spillover	Max pairwise st. diff.
1 if attrited at...					
...the in-person follow-up	0.13	0.01	0.00	0.00	0.05
	(0.33)	(0.02)	(0.02)	(0.02)	1988
	1988	[0.81]	[0.88]	[0.92]	
...the phone-based follow-up	0.12	0.00	-0.01	0.01	0.03
	(0.33)	(0.02)	(0.02)	(0.02)	1988
	1988	[0.81]	[0.88]	[0.92]	
...either follow-up	0.22	0.01	0.00	0.00	0.04
	(0.42)	(0.02)	(0.02)	(0.02)	1988
	1988	[0.81]	[0.88]	[0.92]	

Notes: Coefficient of treatment and spillover at baseline (columns 2-3). Column 3 tests for differences in parameters obtained in previous two columns. Column 4 reports the standardised pairwise maximum difference between mean across all study groups. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. q-values reported in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 displays the control mean; standard deviation; and total number of observations.

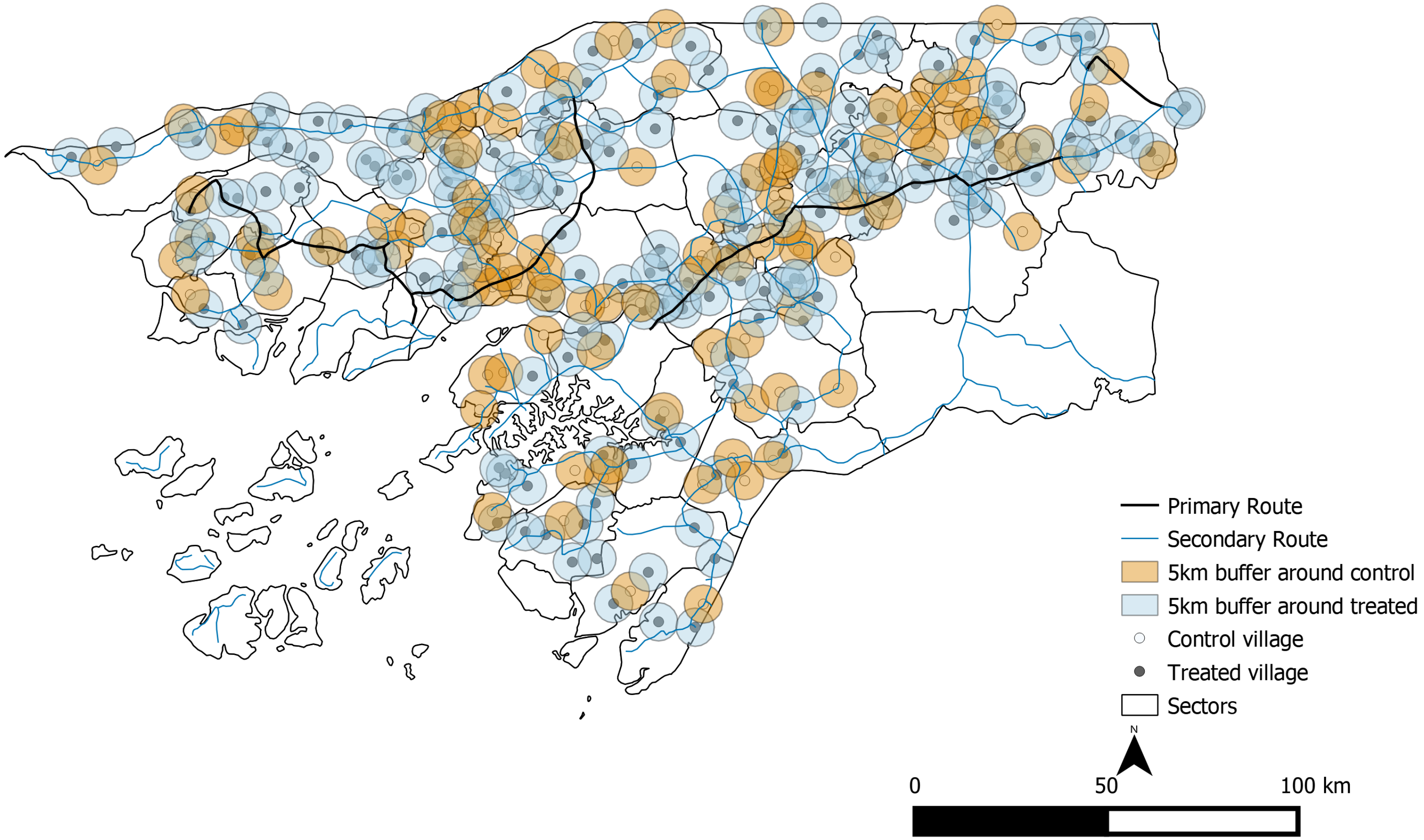


Figure 1.8: Geographic distribution of treatment with 5 km buffer around villages

1.10 Robustness checks

1.10.1 Winsorized outcomes

We winsorized at the 1% level our continuous main outcomes as a robustness check, as we had pre-specified. This affects the significance of our results on the quantity and revenue from sales, though the coefficients remain positive and relatively large. We also look at outcomes winsorized at the 5% level, which we had not pre-specified, which shows that increase in quantity sold and value from sales remains marginally significant using naive *p*-values.

1.10.2 Controlling for unbalanced baseline characteristics

We ran Equation (1.3) including as controls any producer or village characteristic that was found to be substantially unbalanced at baseline, in Appendix Tables 1.8 and 1.10. We include the following covariates:

- Household size
- 1 if faces difficulties storing
- 1 if sells other agricultural products
- Area of plantation
- Years of experience in the sector
- Standardized principal component of assets (Filmer and Pritchett, 2001)
- Numeracy index (Anderson, 2008)
- 1 if extremely present-biased (Ashraf et al., 2006)
- Number of traders in the village

When the baseline value of one of these covariates was missing, we impute to be equal to the sample mean and control for the imputations with an indicator variable.

1.10.3 Estimation using the post-double least absolute shrinkage and selection operator (PDSLASSO)

We estimated our intention-to-treat effects using the post-double least absolute shrinkage and selection operator (PDSLASSO), as described by Belloni et al. (2014). Instead of Equation (1.3), we increase precision of our estimates by letting this algorithm pick additional control variables from baseline values of the producer and village characteristics, and values of the main outcome variables.

1.10.4 Clustering standard errors at the randomisation triplet-level

We estimated our intention-to-treat effects clustering standard errors at the randomisation triplet-level following the recommendation of de Chaisemartin and Ramirez-Cuellar (2023), though we had not pre-specified this alternative strategy.

1.10.5 Pooling across trading seasons

We estimated our intention-to-treat effects pooling data from the 2020 and the 2021 trading seasons.

1.10.6 Between-village spillover specifications

We estimate the effects of treatment intensity on our main outcomes using the following specification:

$$
\begin{aligned}
y_{iv} =& \beta_1 \cdot treatment_{iv} + \delta_1 \cdot spillover_{iv} + \gamma_1 \cdot y_{0iv} \\
& \pi_1^{5km} \cdot \mathrm{V}_{iv\neg v}^{5km} + \pi_2^{5km} \cdot \mathrm{TP}_{iv\neg v}^{5km} + \alpha_v + \epsilon_{iv}
\end{aligned} \tag{1.5}
$$

where variables are defined as in Equation 1.3, $\mathrm{TP}_{iv\neg v}^{5km}$ is the total amount of producers *assigned* to treatment within 5 km of producer i (excluding treated producers in village v), and $\mathrm{V}_{iv\neg v}^{5km}$ is the number of villages within a 5 km radius (excluding village v). Conditional on the number of villages within a given radius, the total amount of treated individuals that are within this radius is exogenous.[29] This equation allows us to estimate the following effects: β_1 gives us the (direct) intent-to-treat effect of the intervention on producers in treated villages. δ_1 gives us the spillover effect of the intervention on producers *within* treated villages. π_2^{5km} gives us the (indirect) *between*-village effect of the total amount of treated individuals within a 5 km radius.

For analyses using spatial treatment intensity, we account for spatial dependence using Conley standard errors with a uniform kernel up to the boundary of the largest doughnut (Conley, 1999).

In order to pick the radius distance of 5 km, we estimate the effect of treatment intensity within a series of non-overlapping doughnuts, $d = 1, ..., D$, each with inner radius r and outer radius $r + c$ kilometres, as in Egger et al. (2022), with $c \in \{1, 2, 3, 4, 5\}$. We estimate 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 $r \in \{0, 1, . . . , 9\}$, and then iterate this procedure for doughnuts with a larger outer radius for each c. For each specification and outcome, we then select the model which

[29] Conditional exogeneity of the number of producers treated within a doughnut comes from the fact that the number of treated producers is a fixed discrete number depending on the random assignment at the village-level. That is, treated villages had 4 treated producers, whereas the pure control villages had none.

minimises the Bayesian Information Criterion (BIC).[30] Across all the fifty specifications we estimated for price, the specification with the lowest BIC for which we observed positive variation in treatment intensity in at least 10% of the sample is the doughnut with $r = 0$ and $r + c = 5$.[31] We therefore focus on presenting results from this specification.

[30] These specifications implicitly assume a linear relationship between the intensity of treatment within a given radius and the outcome variable of interest, but allows this effect to vary between the different 'doughnuts' specified in the regression.

[31] For example, the specification with doughnuts with $r = 0$ and $r + c = 2$ had a marginally lower BIC, but only 5% of the sample had any non-zero variation in this measure of spatial treatment intensity. The lack of variation in treatment intensity within smaller radii is due to our sampling strategy, which aimed to maximise the minimum distance between any village in our sample.

Table 1.12: Robustness — Main outcomes winsorized at 99^{th} percentile

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treatment-Spillover	Year 2 (5) Mean (SD) Total obs.	(6) Treatment	(7) Spillover	(8) Treatment-Spillover
Price and revenue:								
Average price (XOF/kg)	294.76	1.73	2.76	-1.03	354.64	5.66*	1.22	4.44
	(101.66)	(2.96)	(3.16)	(2.98)	(79.70)	(2.98)	(3.23)	(2.79)
	1587	[0.56]	[0.52]	[0.73]	1704	[0.08]*	[0.71]	[0.17]
Value of all sales and exchanges (XOF)	284193.51	30127.41	18143.43	11983.98	434464.65	77702.81***	43159.75*	34543.06
	(289006.62)	(18509.29)	(21135.80)	(18573.33)	(423634.36)	(23248.72)	(25442.04)	(22831.16)
	1520	[0.26]	[0.52]	[0.73]	1686	[0.00]***	[0.18]	[0.17]
Value of all sales (XOF)	230293.14	25365.33	2140.42	23224.92	356704.84	73674.76***	29882.23	43792.53**
	(269143.16)	(16662.85)	(18341.82)	(15886.19)	(385129.20)	(19610.92)	(21728.01)	(19610.05)
	1591	[0.26]	[0.91]	[0.58]	1705	[0.00]***	[0.23]	[0.10]
Value of exchanges (XOF)	52020.20	9291.95	12687.02	-3395.07	75155.05	8983.78	14674.83*	-5691.05
	(86004.73)	(7644.53)	(8269.69)	(5759.07)	(108324.67)	(8204.34)	(8239.38)	(7541.22)
	1635	[0.30]	[0.50]	[0.73]	1710	[0.27]	[0.18]	[0.45]
Quantity sold and exchanged:								
Total quantity sold (kg)	695.97	82.83*	22.28	60.55	956.20	172.65***	80.48	92.17*
	(764.74)	(48.32)	(52.72)	(45.51)	(1013.12)	(50.53)	(54.31)	(48.30)
	1622	[0.22]	[0.67]	[0.46]	1709	[0.00]***	[0.23]	[0.14]
Number of sales	1.48	0.21***	0.08	0.13**	1.63	0.21***	0.10**	0.11**
	(0.94)	(0.06)	(0.06)	(0.06)	(0.81)	(0.04)	(0.04)	(0.05)
	1693	[0.00]***	[0.64]	[0.10]*	1730	[0.00]***	[0.07]*	[0.09]*
1 if exchanged cashew for rice	0.56	-0.01	-0.03	0.03	0.56	-0.01	-0.00	-0.01
	(0.50)	(0.03)	(0.03)	(0.03)	(0.50)	(0.03)	(0.03)	(0.03)
	1706	[0.85]	[0.64]	[0.59]	1724	[0.60]	[0.99]	[0.58]
Total quantity exchanged (kg)	217.70	13.48	16.16	-2.68	230.05	42.29	64.57**	-22.27
	(323.48)	(23.44)	(24.58)	(21.08)	(335.19)	(26.95)	(26.63)	(24.70)
	1681	[0.71]	[0.64]	[0.96]	1707	[0.20]	[0.07]*	[0.57]
Share of quantity sold over quantity exchanged and sold	0.73	0.02	0.02	0.00	0.79	0.01	-0.00	0.01
	(0.32)	(0.02)	(0.02)	(0.02)	(0.25)	(0.02)	(0.02)	(0.01)
	1604	[0.70]	[0.64]	[0.96]	1691	[0.60]	[0.99]	[0.57]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

Table 1.13: Robustness — Main outcomes winsorized at 95^{th} percentile

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treatment-Spillover	Year 2 (5) Mean (SD) Total obs.	(6) Treatment	(7) Spillover	(8) Treatment-Spillover
Price and revenue:								
Average price (XOF/kg)	291.62	2.12	2.28	-0.16	352.17	5.95**	2.28	3.67
	(96.87)	(2.49)	(2.63)	(2.37)	(75.51)	(2.71)	(2.93)	(2.57)
	1587	[0.40]	[0.69]	[0.95]	1704	[0.04]**	[0.44]	[0.31]
Value of all sales and exchanges (XOF)	268381.83	22670.69	10018.81	12651.88	411529.74	67367.14***	45968.37**	21398.77
	(227328.24)	(13876.99)	(15519.40)	(12974.91)	(344012.16)	(19570.91)	(21354.65)	(18289.62)
	1520	[0.21]	[0.69]	[0.66]	1686	[0.00]***	[0.07]*	[0.32]
Value of all sales (XOF)	212156.79	19263.07*	870.66	18392.41*	334113.48	68252.59***	38016.15**	30236.44*
	(198563.75)	(11401.24)	(12720.95)	(10984.52)	(308129.59)	(16615.62)	(18100.38)	(16243.81)
	1591	[0.21]	[0.94]	[0.38]	1705	[0.00]***	[0.07]*	[0.25]
Value of exchanges (XOF)	47953.99	5369.29	6359.93	-990.63	70998.43	7404.42	11033.59	-3629.16
	(69132.58)	(5582.65)	(6042.44)	(4439.21)	(93578.96)	(6835.91)	(6711.52)	(5762.07)
	1635	[0.40]	[0.69]	[0.95]	1710	[0.28]	[0.14]	[0.53]
Quantity sold and exchanged:								
Total quantity sold (kg)	644.94	59.06*	5.31	53.75*	887.45	173.22***	107.88**	65.33
	(572.80)	(33.74)	(37.73)	(32.42)	(778.79)	(42.32)	(45.85)	(42.01)
	1622	[0.20]	[0.89]	[0.25]	1709	[0.00]***	[0.05]**	[0.30]
Number of sales	1.43	0.19***	0.08	0.11**	1.61	0.20***	0.10**	0.11**
	(0.81)	(0.05)	(0.05)	(0.05)	(0.76)	(0.04)	(0.04)	(0.05)
	1693	[0.00]***	[0.61]	[0.10]*	1730	[0.00]***	[0.05]**	[0.10]*
1 if exchanged cashew for rice	0.56	-0.01	-0.03	0.03	0.56	-0.01	-0.00	-0.01
	(0.50)	(0.03)	(0.03)	(0.03)	(0.50)	(0.03)	(0.03)	(0.03)
	1706	[0.85]	[0.61]	[0.59]	1724	[0.60]	[0.99]	[0.58]
Total quantity exchanged (kg)	204.12	4.86	14.63	-9.77	215.94	31.11	42.41**	-11.31
	(270.63)	(20.20)	(21.21)	(16.49)	(282.80)	(21.15)	(20.43)	(17.81)
	1681	[0.85]	[0.61]	[0.69]	1707	[0.24]	[0.06]*	[0.58]
Share of quantity sold over quantity exchanged and sold	0.73	0.02	0.02	0.00	0.79	0.01	-0.00	0.01
	(0.32)	(0.02)	(0.02)	(0.02)	(0.25)	(0.02)	(0.02)	(0.01)
	1604	[0.70]	[0.61]	[0.96]	1691	[0.60]	[0.99]	[0.58]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

Table 1.14: Robustness — Main outcomes controlling for unbalanced characteristics

	Year 1				Year 2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean (SD) Total obs.	Treatment	Spillover	Treatment-Spillover	Mean (SD) Total obs.	Treatment	Spillover	Treatment-Spillover
Price and revenue:								
Average price (XOF/kg)	295.71	0.43	0.65	-0.22	355.77	3.92	-1.03	4.95
	(104.36)	(3.46)	(3.40)	(3.31)	(83.13)	(3.35)	(3.27)	(3.07)
	1587	[0.90]	[0.85]	[0.95]	1704	[0.32]	[0.75]	[0.14]
Value of all sales and exchanges (XOF)	287785.22	48463.11**	29645.71	18817.40	440723.50	75783.73**	30559.25	45224.48*
	(315352.77)	(24318.03)	(25169.23)	(22167.60)	(461068.58)	(30027.99)	(27649.03)	(25103.27)
	1520	[0.17]	[0.48]	[0.68]	1686	[0.03]**	[0.54]	[0.14]
Value of all sales (XOF)	233494.84	39058.70*	13231.17	25827.52	364476.23	65761.91**	16116.88	49645.03**
	(292721.24)	(22580.75)	(22820.34)	(19575.44)	(428368.42)	(26544.51)	(24352.45)	(22738.07)
	1591	[0.17]	[0.75]	[0.68]	1705	[0.03]**	[0.68]	[0.12]
Value of exchanges (XOF)	52325.81	9586.30	14398.90	-4812.60	75661.67	7178.64	12237.10	-5058.46
	(88037.41)	(7668.80)	(9180.83)	(7242.70)	(111392.97)	(8136.67)	(8229.46)	(7713.05)
	1635	[0.28]	[0.47]	[0.68]	1710	[0.38]	[0.54]	[0.51]
Quantity sold and exchanged:								
Total quantity sold (kg)	708.21	113.11*	36.95	76.16	981.13	152.63**	39.23	113.40**
	(855.14)	(60.10)	(59.66)	(53.87)	(1163.13)	(65.32)	(62.96)	(52.43)
	1622	[0.15]	[0.67]	[0.40]	1709	[0.05]*	[0.86]	[0.08]*
Number of sales	1.49	0.18***	0.04	0.14**	1.66	0.17***	0.06	0.11**
	(1.01)	(0.06)	(0.06)	(0.06)	(0.92)	(0.05)	(0.05)	(0.05)
	1693	[0.01]**	[0.67]	[0.11]	1730	[0.00]***	[0.60]	[0.08]*
1 if exchanged cashew for rice	0.56	0.00	-0.03	0.03	0.56	-0.01	0.01	-0.01
	(0.50)	(0.03)	(0.03)	(0.03)	(0.50)	(0.03)	(0.03)	(0.03)
	1706	[0.92]	[0.67]	[0.44]	1724	[0.75]	[0.86]	[0.59]
Total quantity exchanged (kg)	217.70	14.48	10.12	4.37	231.72	34.74	56.83**	-22.09
	(323.48)	(23.18)	(24.29)	(23.09)	(346.31)	(26.69)	(27.08)	(25.54)
	1681	[0.80]	[0.68]	[0.85]	1707	[0.32]	[0.18]	[0.59]
Share of quantity sold over quantity exchanged and sold	0.73	0.01	0.01	-0.00	0.79	0.01	-0.00	0.01
	(0.32)	(0.02)	(0.02)	(0.02)	(0.25)	(0.02)	(0.02)	(0.01)
	1604	[0.80]	[0.67]	[0.85]	1691	[0.75]	[0.86]	[0.59]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

Table 1.15: Robustness — Main outcomes estimated using PDSLASSO

	Year 1				Year 2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean (SD) Total obs.	Treatment	Spillover	Treatment-Spillover	Mean (SD) Total obs.	Treatment	Spillover	Treatment-Spillover
Price and revenue:								
Average price (XOF/kg)	295.71	2.07	2.33	-0.26	355.77	6.89**	1.32	5.57*
	(104.36)	(3.36)	(3.26)	(3.21)	(83.13)	(3.32)	(3.28)	(2.97)
	1587	[0.54]	[0.62]	[0.93]	1705	[0.05]*	[0.69]	[0.08]*
Value of all sales and exchanges (XOF)	287785.22	49955.96**	30500.25	19455.72	440723.50	92352.29***	39643.84	52708.45**
	(315352.77)	(23903.98)	(24599.16)	(21429.83)	(461068.58)	(29925.93)	(27802.66)	(25310.76)
	1527	[0.14]	[0.43]	[0.50]	1696	[0.00]***	[0.31]	[0.07]*
Value of all sales (XOF)	233494.84	39165.14*	10791.39	28373.75	364476.23	80224.19***	22902.73	57321.47**
	(292721.24)	(21653.65)	(21815.02)	(19148.46)	(428368.42)	(26425.18)	(24418.94)	(23194.90)
	1594	[0.14]	[0.62]	[0.50]	1710	[0.00]***	[0.46]	[0.05]*
Value of exchanges (XOF)	52325.81	11216.27	17525.80*	-6309.53	75661.67	9382.75	13794.64*	-4411.89
	(88037.41)	(7925.40)	(9502.94)	(7093.11)	(111392.97)	(7769.93)	(7972.89)	(7354.12)
	1639	[0.21]	[0.26]	[0.50]	1716	[0.23]	[0.31]	[0.55]
Quantity sold and exchanged:								
Total quantity sold (kg)	708.21	112.08*	28.20	83.88	981.13	177.22***	50.18	127.04**
	(855.14)	(58.72)	(58.31)	(52.46)	(1163.13)	(64.75)	(61.83)	(52.70)
	1625	[0.14]	[0.63]	[0.27]	1714	[0.01]**	[0.69]	[0.04]**
Number of sales	1.49	0.21***	0.07	0.14**	1.66	0.20***	0.08*	0.11**
	(1.01)	(0.06)	(0.06)	(0.06)	(0.92)	(0.05)	(0.05)	(0.05)
	1693	[0.00]***	[0.63]	[0.09]*	1731	[0.00]***	[0.19]	[0.04]**
1 if exchanged cashew for rice	0.56	-0.00	-0.03	0.03	0.56	-0.01	-0.00	-0.01
	(0.50)	(0.03)	(0.03)	(0.03)	(0.50)	(0.03)	(0.03)	(0.03)
	1707	[0.89]	[0.63]	[0.47]	1727	[0.62]	[0.90]	[0.72]
Total quantity exchanged (kg)	217.70	20.22	14.13	6.08	231.72	45.40*	63.69**	-18.30
	(323.48)	(22.92)	(24.44)	(22.13)	(346.31)	(25.46)	(26.22)	(24.52)
	1685	[0.63]	[0.63]	[0.86]	1713	[0.12]	[0.08]*	[0.59]
Share of quantity sold over quantity exchanged and sold	0.73	0.01	0.01	-0.00	0.79	0.01	-0.00	0.01
	(0.32)	(0.02)	(0.02)	(0.02)	(0.25)	(0.02)	(0.02)	(0.01)
	1610	[0.76]	[0.63]	[0.86]	1700	[0.62]	[0.90]	[0.59]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

Table 1.16: Robustness — Main outcomes clustering at the randomisation triplet-level

	Year 1				Year 2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean (SD) Total obs.	Treatment	Spillover	Treatment-Spillover	Mean (SD) Total obs.	Treatment	Spillover	Treatment-Spillover
Price and revenue:								
Average price (XOF/kg)	295.71	2.07	2.33	-0.26	355.77	6.89*	1.32	5.57*
	(104.36)	(4.43)	(3.89)	(3.03)	(83.13)	(3.64)	(3.45)	(2.84)
	1587	[0.64]	[0.55]	[0.93]	1704	[0.08]*	[0.70]	[0.14]
Value of all sales and exchanges (XOF)	287785.22	61649.88**	44746.80	16903.08	440723.50	105532.81***	58182.34*	47350.47
	(315352.77)	(28196.67)	(30931.74)	(29016.04)	(461068.58)	(38646.95)	(31129.62)	(33431.31)
	1520	[0.06]*	[0.37]	[0.75]	1686	[0.01]**	[0.19]	[0.21]
Value of all sales (XOF)	233494.84	51335.03**	24128.93	27206.10	364476.23	95365.15***	41745.18*	53619.97*
	(292721.24)	(22663.69)	(23608.62)	(26001.84)	(428368.42)	(30726.31)	(24768.03)	(29301.24)
	1591	[0.06]*	[0.41]	[0.70]	1705	[0.01]***	[0.19]	[0.14]
Value of exchanges (XOF)	52325.81	10968.68	17251.21	-6282.54	75661.67	9774.72	15940.01	-6165.30
	(88037.41)	(9491.92)	(12859.48)	(6686.81)	(111392.97)	(10476.33)	(11222.08)	(8449.50)
	1635	[0.34]	[0.37]	[0.70]	1710	[0.35]	[0.21]	[0.47]
Quantity sold and exchanged:								
Total quantity sold (kg)	708.21	141.75**	63.53	78.21	981.13	211.49***	93.73	117.77*
	(855.14)	(67.14)	(65.52)	(71.98)	(1163.13)	(74.30)	(67.09)	(65.77)
	1622	[0.09]*	[0.55]	[0.54]	1709	[0.01]**	[0.28]	[0.19]
Number of sales	1.49	0.21***	0.07	0.14***	1.66	0.20***	0.08	0.11**
	(1.01)	(0.06)	(0.06)	(0.05)	(0.92)	(0.06)	(0.06)	(0.05)
	1693	[0.01]***	[0.55]	[0.03]**	1730	[0.00]***	[0.28]	[0.10]*
1 if exchanged cashew for rice	0.56	-0.01	-0.03	0.03	0.56	-0.01	-0.00	-0.01
	(0.50)	(0.04)	(0.04)	(0.03)	(0.50)	(0.04)	(0.04)	(0.02)
	1706	[0.88]	[0.55]	[0.54]	1724	[0.70]	[0.99]	[0.56]
Total quantity exchanged (kg)	217.70	19.62	20.91	-1.30	231.72	44.46	70.75*	-26.28
	(323.48)	(32.85)	(34.91)	(21.96)	(346.31)	(32.66)	(37.15)	(26.09)
	1681	[0.69]	[0.55]	[0.96]	1707	[0.29]	[0.28]	[0.52]
Share of quantity sold over quantity exchanged and sold	0.73	0.02	0.02	0.00	0.79	0.01	-0.00	0.01
	(0.32)	(0.02)	(0.03)	(0.02)	(0.25)	(0.02)	(0.02)	(0.01)
	1604	[0.68]	[0.55]	[0.96]	1691	[0.70]	[0.99]	[0.52]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

Table 1.17: Robustness — Main outcomes pooling across years

	(1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treat. vs. spillover
Price and revenue:				
Average price (XOF/kg)	326.60	4.79*	1.77	3.02
	(98.69)	(2.52)	(2.55)	(2.31)
	3292	[0.08]*	[0.49]	[0.26]
Value of all sales and exchanges (XOF)	367270.23	83574.12***	50786.10**	32788.02
	(404914.15)	(24422.26)	(23862.50)	(20451.26)
	3207	[0.00]***	[0.07]*	[0.22]
Value of all sales (XOF)	300792.18	72884.01***	31846.50	41037.51**
	(374321.08)	(21215.67)	(20495.86)	(18187.91)
	3297	[0.00]***	[0.16]	[0.10]*
Value of exchanges (XOF)	64140.44	9693.10	16223.28**	-6530.18
	(101175.95)	(6934.40)	(7730.56)	(6333.36)
	3346	[0.16]	[0.07]*	[0.30]
Quantity sold and exchanged:				
Total quantity sold (kg)	848.06	174.78***	77.30	97.48**
	(1033.22)	(53.00)	(51.45)	(44.18)
	3332	[0.00]***	[0.22]	[0.07]*
Number of sales	1.57	0.20***	0.08*	0.13***
	(0.97)	(0.04)	(0.04)	(0.04)
	3424	[0.00]***	[0.16]	[0.01]**
1 if exchanged cashew for rice	0.56	-0.01	-0.01	0.01
	(0.50)	(0.03)	(0.03)	(0.02)
	3431	[0.74]	[0.69]	[0.74]
Total quantity exchanged (kg)	224.72	33.01	46.87**	-13.86
	(335.03)	(21.61)	(22.78)	(18.89)
	3389	[0.21]	[0.16]	[0.74]
Share of quantity sold over quantity exchanged and sold	0.76	0.01	0.01	0.01
	(0.29)	(0.02)	(0.02)	(0.01)
	3296	[0.58]	[0.69]	[0.74]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 pooling across 2020 and 2021. Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 displays the control mean; standard deviation; and total number of observations across all groups.

Table 1.18: Mechanisms — Marketing behaviour (pre-specified)

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treatment-Spillover	Year 2 (5) Mean (SD) Total obs.	(6) Treatment	(7) Spillover	(8) Treatment-Spillover
1 if delayed sales because expected price to raise	0.79	-0.01	-0.01	0.00	0.85	-0.04*	-0.02	-0.02
	(0.41)	(0.02)	(0.02)	(0.02)	(0.36)	(0.02)	(0.02)	(0.02)
	1703	[0.63]	[0.72]	[0.96]	1717	[0.13]	[0.76]	[0.61]
1 if sped up sales because expected price to drop	0.27	0.04	0.03	0.00	0.22	0.01	0.01	-0.01
	(0.44)	(0.02)	(0.02)	(0.03)	(0.41)	(0.02)	(0.02)	(0.02)
	1694	[0.28]	[0.39]	[0.96]	1713	[0.88]	[0.76]	[0.83]
1 if sold most to local buyer	0.43	-0.02	0.04	-0.06*	0.56	-0.05*	-0.02	-0.03
	(0.50)	(0.03)	(0.03)	(0.03)	(0.50)	(0.03)	(0.03)	(0.03)
	1622	[0.63]	[0.39]	[0.38]	1709	[0.13]	[0.76]	[0.61]
1 if sold most to itinerant buyer	0.43	0.04	0.01	0.04	0.39	0.06**	0.03	0.03
	(0.50)	(0.03)	(0.03)	(0.03)	(0.49)	(0.03)	(0.03)	(0.03)
	1622	[0.28]	[0.79]	[0.64]	1709	[0.13]	[0.76]	[0.61]
1 if sold most directly to a market	0.04	-0.02**	-0.02	-0.01	0.01	-0.00	0.00	-0.00
	(0.20)	(0.01)	(0.01)	(0.01)	(0.11)	(0.01)	(0.01)	(0.01)
	1622	[0.13]	[0.39]	[0.67]	1709	[0.88]	[0.76]	[0.83]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

Table 1.19: Mechanisms — Reasons for multiple sales and type of advice sought

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treat. vs. spillover
Sold more than once because...				
...they think they get better prices	0.23	0.07***	0.02	0.04
	(0.42)	(0.02)	(0.02)	(0.03)
	1693	[0.01]***	[0.91]	[0.12]
...it was adviced by *n'kalô*	0.01	0.03***	-0.00	0.03***
	(0.09)	(0.01)	(0.01)	(0.01)
	1693	[0.00]***	[0.91]	[0.01]***
...they can smooth consumption	0.22	0.05**	-0.01	0.06**
	(0.41)	(0.02)	(0.03)	(0.03)
	1693	[0.05]**	[0.91]	[0.04]**
Sought market advice on...				
...the best time to sell	0.48	0.05	-0.04	0.08***
	(0.50)	(0.03)	(0.03)	(0.03)
	1700	[0.31]	[0.28]	[0.01]**
...the price to sell at	0.61	0.01	-0.05	0.06**
	(0.49)	(0.03)	(0.03)	(0.03)
	1706	[0.76]	[0.28]	[0.06]*
...whom to sell to	0.24	0.03	-0.01	0.04
	(0.42)	(0.02)	(0.02)	(0.02)
	1632	[0.42]	[0.67]	[0.11]

Notes: Producer-level intention-to-treat (ITT) estimates in 2020 (year 1) reported in columns 2 and 3. Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 displays the control mean; standard deviation; and total number of observations across all groups.

Table 1.20: Mechanisms — Beliefs about prices and sharing of offers

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treatment-Spillover	Year 2 (5) Mean (SD) Total obs.	(6) Treatment	(7) Spillover	(8) Treatment-Spillover
Beliefs about prices								
Expected sale price (XOF/kg)	512.48	-15.37	-7.21	-8.16	664.05	-15.28	-8.48	-6.79
	(193.57)	(10.51)	(11.24)	(11.84)	(251.30)	(14.08)	(14.86)	(14.48)
	1580	[0.43]	[0.54]	[0.71]	1562	[0.39]	[1.00]	[0.64]
Desired reference price (XOF/kg)	697.85	-15.66	-10.37	-5.30	815.13	-21.58	0.52	-22.10
	(277.93)	(16.21)	(17.03)	(14.40)	(252.18)	(15.79)	(17.17)	(16.72)
	1701	[0.50]	[0.54]	[0.71]	1693	[0.39]	[1.00]	[0.56]
1 if thinks that the reference price is important in marketing	0.38	0.01	0.02	-0.01	0.24	-0.02	0.00	-0.02
	(0.49)	(0.03)	(0.03)	(0.03)	(0.43)	(0.02)	(0.02)	(0.03)
	1689	[0.69]	[0.54]	[0.71]	1717	[0.39]	[1.00]	[0.64]
Number of producers...								
...informed of an offer received	6.82	0.35	0.13	0.23	11.28	-0.15	-0.90	0.75
	(8.66)	(0.48)	(0.48)	(0.44)	(13.32)	(0.72)	(0.78)	(0.77)
	1596	[0.46]	[0.79]	[0.73]	1633	[0.84]	[0.25]	[0.45]
...that shared their offers	4.78	0.71*	0.59	0.13	8.93	-0.78	-1.28	0.50
	(8.62)	(0.37)	(0.37)	(0.37)	(16.05)	(0.82)	(0.84)	(0.66)
	1583	[0.11]	[0.22]	[0.73]	1614	[0.69]	[0.25]	[0.45]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

Table 1.21: Mechanisms — Placebo outcomes

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treatment-Spillover	Year 2 (5) Mean (SD) Total obs.	(6) Treatment	(7) Spillover	(8) Treatment-Spillover
1 if filled the sale diary well	0.31	-0.03	-0.04	0.01				
	(0.46)	(0.03)	(0.03)	(0.02)				
	1719	[0.54]	[0.45]	[0.53]				
1 if is extremely risk averse	0.35	-0.00	0.03	-0.03				
	(0.48)	(0.03)	(0.03)	(0.03)				
	1627	[0.89]	[0.45]	[0.53]				
1 if trusts most people	0.70	0.04*	0.02	0.02	0.76	0.04**	0.01	0.04
	(0.46)	(0.03)	(0.03)	(0.03)	(0.43)	(0.02)	(0.02)	(0.03)
	1712	[0.27]	[0.45]	[0.53]	1713	[0.05]**	[0.79]	[0.15]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

Table 1.22: Mechanisms — Cashews used to repay loans

	Year 1				Year 2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean (SD) Total obs.	Treatment	Spillover	Treatment-Spillover	Mean (SD) Total obs.	Treatment	Spillover	Treatment-Spillover
1 if borrowed money or rice	0.31	0.02	0.03	-0.01	0.42	-0.04	0.02	-0.06**
	(0.46)	(0.03)	(0.03)	(0.03)	(0.49)	(0.03)	(0.03)	(0.03)
	1702	[0.65]	[0.57]	[0.75]	1724	[0.15]	[0.62]	[0.08]*
Implicit price of loans repaid with cashews (XOF/kg)	58.96	12.62**	3.30	9.33	91.35	8.26*	-2.81	11.07*
	(142.79)	(5.49)	(6.13)	(5.79)	(147.69)	(4.61)	(3.74)	(6.61)
	1645	[0.09]*	[0.59]	[0.41]	1603	[0.15]	[0.62]	[0.13]
Total quantity paid for loans (kg)	44.09	-3.23	7.00	-10.24	57.17	-4.86	12.45	-17.30**
	(106.88)	(7.08)	(8.99)	(8.11)	(121.42)	(7.60)	(8.75)	(8.49)
	1686	[0.65]	[0.58]	[0.41]	1716	[0.52]	[0.62]	[0.08]*
Share of quantity sold over quantity exchanged and paid in loans	0.86	0.01	0.03	-0.02	0.91	0.02	0.00	0.01
	(0.29)	(0.02)	(0.02)	(0.02)	(0.20)	(0.01)	(0.01)	(0.01)
	1599	[0.65]	[0.49]	[0.44]	1696	[0.15]	[0.75]	[0.23]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

Table 1.23: Mechanisms — Other uses of cashews

	Year 1 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treatment-Spillover	Year 2 (5) Mean (SD) Total obs.	(6) Treatment	(7) Spillover	(8) Treatment-Spillover
Total quantity produced with imputations	1056.06	177.39**	113.68	63.71	1369.04	238.62***	193.34**	45.28
	(1046.24)	(71.32)	(72.66)	(64.28)	(1413.66)	(89.25)	(85.41)	(67.39)
	1718	[0.07]*	[0.47]	[0.55]	1739	[0.04]**	[0.12]	[0.63]
1 if lost cashew post-harvest	0.05	0.00	0.02	-0.02	0.08	-0.02	-0.00	-0.01
	(0.22)	(0.01)	(0.01)	(0.01)	(0.27)	(0.01)	(0.02)	(0.02)
	1708	[0.91]	[0.47]	[0.55]	1719	[0.39]	[0.88]	[0.63]
Quantity lost post-harvest (kg)	0.74	-0.03	0.23	-0.26	1.86	-0.43	-0.12	-0.31
	(5.05)	(0.27)	(0.37)	(0.36)	(14.77)	(0.52)	(0.75)	(0.74)
	1692	[0.91]	[0.67]	[0.59]	1704	[0.41]	[0.88]	[0.68]
1 if processed cashew nuts	0.00	0.00	0.00	-0.00	0.01	-0.00	0.00	-0.01
	(0.04)	(0.00)	(0.00)	(0.00)	(0.10)	(0.00)	(0.00)	(0.00)
	1699	[0.70]	[0.60]	[0.91]	1719	[0.41]	[0.88]	[0.61]
Quantity of processed raw cashews (kg)	0.15	0.59	0.08	0.52	2.07	-0.39	0.06	-0.45
	(3.68)	(0.52)	(0.34)	(0.53)	(37.27)	(0.27)	(0.40)	(0.38)
	1699	[0.63]	[0.82]	[0.55]	1719	[0.37]	[0.88]	[0.61]
Additional uses not pre-specified:								
Quantity traded in small weekly exchanges (kg)	145.40	-9.38	-9.47	0.09	83.73	2.52	-2.88	5.40
	(289.09)	(16.48)	(17.19)	(17.03)	(145.90)	(7.64)	(7.55)	(7.46)
	1604	[0.79]	[0.99]	[1.00]	1590	[0.74]	[0.88]	[0.85]
Total quantity of cashews stored (kg)	0.60	-0.35	-0.43	0.08	0.04	0.54	-0.07	0.61
	(10.70)	(0.31)	(0.39)	(0.14)	(1.02)	(0.58)	(0.76)	(1.30)
	1679	[0.79]	[0.99]	[1.00]	1699	[0.68]	[0.93]	[0.85]
Total quantity of cashews paid to labourers (kg)	162.72	8.85	-0.25	9.10	147.26	18.16	20.79	-2.64
	(261.04)	(17.21)	(16.03)	(14.24)	(245.84)	(15.92)	(15.97)	(13.55)
	1421	[0.79]	[0.99]	[1.00]	1392	[0.68]	[0.63]	[0.85]
1 if perceived production increased in 2021 relative to 2020	0.33	-0.01	0.00	-0.01	0.64	0.02	0.03	-0.01
	(0.47)	(0.03)	(0.03)	(0.03)	(0.48)	(0.03)	(0.03)	(0.03)
	1707	[0.79]	[0.99]	[1.00]	1723	[0.68]	[0.63]	[0.85]
1 if perceived production decreased in 2021 relative to 2020	0.56	0.01	0.01	0.00	0.33	-0.02	-0.03	0.01
	(0.50)	(0.03)	(0.03)	(0.03)	(0.47)	(0.03)	(0.03)	(0.03)
	1707	[0.79]	[0.99]	[1.00]	1723	[0.68]	[0.63]	[0.85]

Notes: Producer-level intention-to-treat (ITT) estimates reported in columns 2 and 3 across year 1 (2020) and year 2 (2021). Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 and column 5 display the control mean; standard deviation; and total number of observations across all groups in the estimation sample.

1.11 Production and commercialisation details

1.11.1 Production costs

The production costs relate to the payment of labour for two main activities: cleaning of the field and collection of the nuts. The cleaning of the field under the cashew trees is necessary in order to have easier access to the nuts that have fallen from the trees. Family or hired seasonal labour is hired to clean the fields before fruits are ripe and start falling together with the nuts. The second activity involves actually picking up nuts that have fallen from the trees. For the product to be ripe and of good quality, the nut should be picked up from the floor after it has fallen from the tree. Occasionally, credit constrained producers may collect the nuts directly from the trees, but this means that the nut has not fully absorbed all nutrients from the tree and ends up being smaller and lower quality. Labour costs are noisy to measure, especially since they are often paid in-kind using a rule-of-thumb type of payment (e.g. every third day of work, the amount collected goes to the labourers), or through shared meals.

1.11.2 Rice exchanges and interlinked contracts

Many producers exchange their cashews in return for rice or to obtain loans to pay for their immediate needs. While technically illegal (according to national legislation), in-kind exchanges and loans of rice are common and represent an important margin on the commercialisation decisions of producers in this market. In our baseline sample, 35% of the producers have bartered cashews for rice and 34% have borrowed money or rice in exchange for cashews. The quantities involved in these transactions are generally a relatively small fraction of the overall production for most producers. In 2019, we estimate that on average 13% of marketable production was used for barters with rice or loans. The majority of these informal loans are reported to take place in

Table 1.24: Mechanisms — Perceived productivity changes

	Year 2 (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treat. vs. spillover
1 if exerted very high effort in collection	0.30	-0.03	-0.05*	0.02
	(0.46)	(0.03)	(0.03)	(0.03)
	1719	[0.58]	[0.32]	[0.81]
1 if exerted very high effort in monitoring	0.34	-0.02	-0.04	0.02
	(0.47)	(0.03)	(0.03)	(0.03)
	1722	[0.75]	[0.49]	[0.81]
1 if exerted high effort in collection	0.70	-0.01	-0.02	0.00
	(0.46)	(0.03)	(0.03)	(0.03)
	1719	[0.75]	[0.65]	[0.87]
1 if exerted high effort in monitoring	0.72	-0.03	-0.00	-0.03
	(0.45)	(0.03)	(0.03)	(0.03)
	1722	[0.58]	[0.86]	[0.81]
1 if increased area of plantation	0.51	0.01	0.02	-0.01
	(0.50)	(0.03)	(0.03)	(0.03)
	1725	[0.83]	[0.65]	[0.81]
1 if new trees yielded	0.44	0.02	0.03	-0.01
	(0.50)	(0.03)	(0.03)	(0.03)
	1725	[0.75]	[0.63]	[0.81]

Notes: Producer-level intention-to-treat (ITT) estimates in 2021 (year 2) reported in columns 2 and 3. Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 displays the control mean; standard deviation; and total number of observations across all groups.

the months preceding the trading season (November to February), when producers are most liquidity constrained. These interlinked contracts reduce the ability of producers to pick a better time during the trading season to sell their production. However, even those who take up a loan (in-kind or in-cash) often still make sales during the trading season.

1.11.3 Cashew trading seasons during the COVID-19 pandemic

The global surge of COVID-19 in the first quarter of 2020 occurred during the run-up to the cashew trading season in Guinea-Bissau, substantially affecting the market conditions due to a combination of international and local factors. The 2020 trading season started on the 27^{th} of May, two months after what would have been the usual start of the season. The delay was a result of COVID-19 and a contested presidential election held in December 2019. However, small trades of cashews were recorded from a number of buyers since March 2020.
Internationally, the two largest importers of raw cashew nuts, India and Vietnam, had already closed their borders and cashew processing plants when Guinea-Bissau confirmed its first two positive cases of COVID-19 on March 25^{th}, 2020. Even though demand for processed cashew remained high in the first half of 2020, this disruption in the supply chain caused a substantial slump in the international demand for raw cashew nuts. Sales in other cashew-producing countries in West Africa almost came to a complete halt, as border closures all over the world brought additional market uncertainty. Guinea-Bissau's air and land borders were closed, as a preventive measures to reduce the spread of COVID-19. These measures prevented international and regional cashew value-chain agents, mostly buyers and intermediaries, from entering Guinea-Bissau, also reducing the influx of capital needed to buy raw cashew nuts from producers. National travel restrictions between regions also limited the movement of

seasonal workers that support the labour-intensive collection of cashew nuts.
Locally, the government delayed the start of the official trading season, typically starting by the end of March, until May 27^{th}. While trade between producers and intermediaries generally takes place before the official start of the season, this is technically illegal and demand for raw cashew nuts generally increases substantially once the government announces the official start of the season.
In 2021, the national market operated more regularly, as a result of both the government and other stakeholders in the cashew value chain agreeing on a set of measures to reduce risk of spreading COVID-19 during the trading season. The start of the official trading season was still delayed until April 7^{th}. However, the number of foreign exporters buying in the local market in 2021 was smaller than pre-pandemic years, due to reduced international travel.

1.12 Weekly messages

Our intervention sent weekly messages using a combination of short messages, robocalls, and an interactive voice response (IVR) system during the 2020 and 2021 trading seasons. Short text messages were sent to those randomly selected producers between April 1^{st} and May 8^{th}, 2020. We also delivered a short text message as our first message of the 2021 trading season. In 2020, from May 16^{th}, the weekly messages were sent via robocalls. Finally, since June 26^{th}, the information was also available on demand through an IVR service. The IVR service was active for only two months prior to the end of the 2020 season, due to delays associated with COVID-19 and the implementing mobile operator, MTN. Our original plan was to develop the IVR service before the start of the 2020 season. According to our original plan, we would have had two separate treatment arms, with the first one receiving the service for free and the other one only after subscribing to it, paying a small monthly price. Given the delay in setting up the IVR service, we decided to merge these two treatment arms into one, providing the service for free to every treated producer. In 2021, while we also experienced some delays with our mobile operator in the sending of robocalls, almost all of the messages were sent as a robocall and the IVR service was active from April 23^{rd}.

Table 1.25: Messages sent during the 2020 and 2021 cashew trading seasons

Date sent	Format	English translation
		2020 trading season
01/04/2020	SMS 1	BELAB and the *n'kalô* service will send you information each week on the cashew-nut market. Since the voice-messages are not ready yet, we will send you an SMS.
01/04/2020	SMS 2	Opening of the season was delayed because of coronavirus. We advise to wait to sell until the confusion is over. Dry your nuts well to keep quality
16/04/2020	SMS 3	Opening of the season was delayed still because of coronavirus. Wait to sell until it opens and until the price goes above 300 XOF/kg. Dry your nuts well to keep quality
28/04/2020	SMS 4 (placebo)	BELAB/*n'kalô* that interviewed you in 2019 wishes you a good cashew trading season. Remember to fill in your sales diary in order to take part in the lottery at the end of the season.
28/04/2020	SMS 5	Reference price was announced to be 375XOF/KG. Official opening is still delayed because of the state of emergency. Dry your nuts well to keep the quality.
08/05/2020	SMS 6	Port is still closed because of coronavirus. We think the price will increase in the next few weeks, so wait until the price reaches 375XOF/KG or more to sell.
16/05/2020	Robocall 1	Dear stakeholders of the cashew sector, this is André NANQUE, from nkalô, to talk about the commercialization of cashew nuts this year. Each week, we will send messages to keep you informed about the market situation and prices. The official launch of the cashew trading season is still delayed due to the Coronavirus epidemic. But the season is due to launch in a few weeks. In neighboring countries like Senegal, Gambia and Guinea-Conakry, prices have increased in recent weeks. In fact, cashew processors in India and Vietnam are short of cashews and need it in the coming months. Prices in other countries have risen above 350 XOF/kg. We believe that as soon as the season really starts, prices will go up a lot; therefore, we recommend that you dry your nuts well and maintain your stocks until prices rise in June.
21/05/2020	Robocall 2	Unfortunately, the official launch of the cashew trade season has been delayed. It shouldn't be too late, but no date is known yet. Despite this, prices started to rise this week due to the high demand in the international market. In Biombo, Cacheu and Oio, prices range from 250 to 300 XOF/kg, in Gabu the prices fixed at 250 XOF/kg and remain the same as last week, in Bafata, Bolama, Quinará and Tombali, where the price is also at 250 XOF/kg. In neighboring countries, Senegal, Gambia and Guinea Conakry, prices have exceeded 350 XOF/kg. We recommend that you wait and start selling your product only if the price offered to you reaches 375 XOF/kg or more.
29/05/2020	Robocall 3	"Since the season has started late, it is also likely to end late. It will take a few more weeks for exporters from the port of Bissau to start shipping cashews on ships to India and Vietnam. The demand for Guinea-Bissau cashews is very strong there and, therefore, we believe prices will increase in the coming weeks. Currently, in the regions of Bafata, Gabu, Bolama, Quinara and Tombalí, prices have remained at 250 XOF/kg and in Oio, prices still vary between 250 - 300 XOF/kg. On the other hand, in Cacheu, prices started to rise and are between 300-350 XOF/kg, while in Biombo, the price increase is even stronger and sales are made between 350-375 XOF/kg. With competition that will rise, we believe that prices will still rise but in the coming weeks. In order not to take too much risk, but to take advantage of the price increase, we advise all producers to wait until 375 XOF/kg is offered before selling the first half of their production and keep the other half selling later when prices can still be higher. We remind you that in Senegal and The Gambia, prices exceed 400 XOF/kg."
11/06/2020	SMS 7	Prices are still rising: we advise to sell half of your stock if prices reach 375XOF/KG or more, and sell the rest later.
17/06/2020	Robocall 4	The commercialisation of cashew nuts is intensifying across the country and, in recent days, representatives of Indian and Vietnamese buyers have started to arrive in the country to check the quality. This week's prices have not changed much from last week. The purchase prices of producers remain between 300 and 350 XOF/kg in the regions of Bafata, Bolama, Gabu, Oio, Quinara and Tombali. Higher prices are practiced in Biombo regions, where sales of 375 XOF/kg are made and in Cacheu, where prices reach up to 400 XOF/kg in locations close to Senegal. At the port of Bissau, cashew nuts trucks are paid between 380 and 420 XOF/kg, depending on the quality of the nuts. When the nuts are very dry, they are more expensive. With increasing competition between buyers, we continue to think that prices will increase slightly in the coming weeks and we always advise to store the cashews until the price of 375 XOF/kg is offered, and to sell half of your stocks when that price is offered.
25/06/2020	Robocall 5	"Large quantities have arrived in Bissau since the season was launched. This slightly reduced demand at the port and caused a slight drop in prices for wholesale prices at the port of Bissau. As a result, producer prices have fallen slightly in some areas of production and are stable in others. Currently, prices are between 300 and 325 XOF/kg in the regions of Bafata and Gabu and between 300 and 350 XOF/kg in all other regions of the country. With the coronavirus epidemic still a problem in many countries around the world, cashew buyers have reduced their orders. Unlike the past few weeks, we are no longer sure that prices will rise. That is why we recommend that you sell most of your cashews if the prices offered are higher than 325 XOF/kg.

01/07/2020	Robocall 6	This week, the first shipment has already left the ports of Bissau, bound for India. Despite these first exports, there are still a lot of stocks in the port of Bissau and, therefore, exporters are not in a hurry to buy. Above all, they want to export the nuts they have in their stores before placing further orders with traders in the production areas. That is why prices practically do not change, always with sales between 300 and 350 XOF/kg in the areas of cashew production. We always recommend selling most of your nuts if a price of 325 XOF/kg or more offers you, because with the situation of the coronavirus, we do not know how the prices will evolve in the coming weeks.
14/07/2020	Robocall 7	This week, there was a certain slowdown in terms of cashew nut transactions in Bissau. Traders have reduced deliveries and are demanding higher prices from exporters. But at the same time, demand from the Vietnam and India factories remains low and cashew-nut prices have fallen in Senegal, Gambia and the Ivory Coast. In certain regions of Guinea Bissau, prices also drop a little. In the regions of Bolama, Quinara and Tombali, prices remained at 300 XOF/kg; in the regions of Bafata and Gabu, prices will vary between 250-350XOF/kg; in Biombo and Oio, prices are between 300-350 XOF/kg and in Cacheu, there were slight increases of 25 XOF in prices, with variations between 325-375 XOF/kg. We recommend selling most of the production when a price of 325 XOF/kg or more is offered to you"
16/07/2020	Robocall 8	This week, new exporters started to buy cashew nuts which increased demand at the port. Prices rose marginally at the port of Bissau. The prices at the weybridge averaged between 375 and 390 XOF/kg, whereas last week they were between 370 et 380 XOF/kg. This increase in demand also increased producer prices. In the regions of Bafata, Gabu, Oio and Bolama the producer price is around 300 XOF/kg. In Quinará and Tombali, producer prices are between 300 et 325 XOF/kg, slightly better than the previous week. In Biombo, certain producers can sell up to 330 XOF/kg, however in Cacheu prices are between 350 et 375 XOF/kg. As last week, we recommend to sell most or all production when a price of 325 XOF/kg or more is offered to you.
22/07/2020	Robocall 9	This week, the competition between cashew exporters in Bissau port further increased and prices reached 400 XOF/kg. The producer prices also slightly increased. Higher prices continue to be paid in Cacheu, where producer prices are between 350 and 375 XOF/kg. In the regions of Bafata, Bolama, Gabu, Biombo and Oio prices are between 300 and 350 XOF/kg i.e. an increase of between 10 and 50 XOF/kg. The lower prices are paid in Quinara and Tombali where cashews are purchased for prices between 300 and 325 XOF/kg. This increased demand is an excellent opportunity to try to negotiate a good price for the cashew nuts you still have with you. We recommend negotiating a price of 350 XOF/kg to sell all remaining stocks.
31/07/2020	Robocall 10	This week, the competition between nut exporters in the port of Bissau continues to be great and the prices also increased slightly. The producer prices also increased slightly in the production zones. The highest prices in Cacheu productor reached between 350 and 400 XOF/kg. In the regions of Bafata, Biombo, Bolama, Oio and Gabu prices were 350 XOF/kg, which was the same price of last week. In the regions of Quinara and Tombali the price increase was strong. In these two regions the prices went from 300 and 325 XOF/kg last week to 350 and 375 XOF/kg this week. The season will end in a few weeks. We therefore recommend negotiating a minimum price of 350 XOF/kg to sell all remaining stocks. You can get a good price without waiting too much more to sell.
06/08/2020	Robocall 11	This week, the cashew trading season is near the end in the regions of Bafata, Biombo, Bolama, and Oio where the last prices paid to producers were 350 XOF/kg. Quinara and Tombali persist in high demand for cashew nuts and prices paid are between 350 and 375 XOF/kg , like last week. In Cacheu, cashews nuts are still purchased between 375 and 400 XOF/kg but there is limited quantity available. In the port of Bissau, the competition fell slightly and prices decreased slightly. International demand is limited at this point so it is very unlikely that prices will keep increasing. Our advice is to sell the nuts that you still have as fast as possible.
13/08/2020	Robocall 12	This week, the commercialization season cashews finished in the regions of Bafata, Biombo, Bolama and Oio, where heavy rain began and the few available stocks made the trade very difficult. In Quinara and Tombali, there is still demand for cashews with stable prices between 350 and 375 XOF/kg. In Cacheu, nuts are still purchased between 375 and 400 XOF/kg. In the port of Bissau, the competition increased slightly and prices also marginally increased. International demand is limited at this point what makes very unlikely rising prices. Our advice and sell the nuts that are with you as fast as possible.
22/08/2020	Robocall 13	This week the trading season ended in all cashew producing regions. Almost all seasonal trading points of intermediaries from producers have closed, with no purchase and sales of cashews recorded inside the country, throughout the week. The last sales are made in the capital, Bissau, between traders and exporters with a sale price between 400 and 430 XOF/kg. As such, we'll stop sending weekly messages from next week. We hope you have enjoyed our information and we will work to continue to inform you from the beginning of the next cashew trading season. Until the next year!
09/09/2020	SMS 8 (placebo)	BELAB/*n'kalô* that interviewed you in 2019 reminds you to keep your sales diary in order to take part in the lottery before the next trading season.

2021 trading season		
SMS 1	24/03/2021	We are restarting the *n'kalô* messages for 2021. Producer prices are between 250-300 XOF/kg. We think they will increase in the coming months so we advise you not to sell yet. Have a good season!
Robocall 1	23/04/2021	This week, prices rose slightly in all Guinea-Bissau cashew producing areas. However, prices have not yet reached peak, as not all buyers and exporters received their permits and, therefore, the product cannot yet go out to the harbour. This week, the highest prices are observed in the regions of Biombo, Cacheu and Oio, where producer prices are between 350 and 360 XOF/kg. In the regions of Bafata, Quinara and Tombali prices range from 300 to 350 XOF/kg. The regions where the smallest prices are observed are Bolama and Gabu where prices are between 250 and 350 XOF/kg. Even in these regions, prices tend to increase. We continue to recommend waiting for prices to increase before selling. A price of 450 XOF/kg seems to us easily reach before sale.
Robocall 2	29/04/2021	Again this week, prices are slightly rising across the country. The competition between buyers begins to emerge but did not reach its maximum because deliveries in the port of Bissau are not yet possible. The producer prices recorded are: in the Gabu region, prices have been varied between 250 and 350 XOF/kg; In the Bolama region, the minimum price has grown by 50 XOF/kg, with prices between 300 and 350 XOF/kg, similar to the regions of Bafata, Quinara and Tombali; In the Biombo region, minimum price reductions were recorded at 50 XOF/kg and the maximum price at 10 XOF/kg compared to last week. In this region the prices ranged between 300 and 350 XOF/kg; In the Oio region, prices remained relatively stable throughout the week, with prices ranging between 350and 360 XOF/kg; In the Cacheu region, along with the borders with Senegal, maximum prices grew by 15 XOF/kg compared to last week, reaching values higher than the reference price announced by the government, prices are ranging between 350 and 375 XOF/kg along the border. In neighboring countries, prices continue to go up. They are between 400 and 525 XOF/kg in Senegal and most areas of West African cashews production, where the season started earlier. We are therefore convinced that prices will continue to increase in the coming weeks and therefore always recommend that you avoid selling at a price below 450 XOF/kg.
Robocall 3	13/05/2021	This week, despite the official opening of the scale in Bissau, the trucks did not begin to deliver the cashew nuts to the harbour. The reason stems from negotiations between cashew buyers and the government on taxes. Once the situation is resolved, sales can finally accelerate and prices should rapidly increase. In recent days, prices have continued to increase despite everything, exceeding 350 XOF/kg in most production zones prices of 400 XOF/kg that is already available in the regions of Biombo and Cacheu. Some buyers in the Cacheu region began offering 450 XOF/kg. We continue to recommend that you wait until this price of 450 XOF/kg is offered to you before selling most of your cashews stocks.
Robocall 4	21/05/2021	The cashew market in Guinea-Bissau, throughout the week, registered positive evolutions in terms of negotiations, which had been taking place between the government and the confederation of cashews associations, and also in relation to the price to the producer. From the information collected from the Cashew National Agency, soon the new fees and taxes resulting from the above mentioned negotiations should be announced. In recent days, the following variations have been recorded: in the Bolama region, prices remained relatively stable compared to last week, ranging between 300-350 XOF/kg; In the Tombali region, the prices were relatively stable throughout the week, with oscillation of around 350 XOF/kg, in the regions of Bafata and Biombo, this week, similar variations were registered to those last week, that is, between 350-400 XOF/kg; In Gabu and Oio regions, prices remained relatively stable throughout the week, with variations between 350-375 XOF/kg, similar variations for Quinara region, where prices grew by 25 XOF/kg compared to last week. Finally, in the Cacheu region, near the borders with Senegal, prices have been ranging between 400 - 450 XOF/kg like last week. Prices will increase in the next days after the Government announcements. We still advise to wait 450 XOF/kg to sell most of your production.
Robocall 5	28/05/2021	This week, cashew traders and exporters finally reached an agreement with the government about the level of export taxes and deliveries to the port of Bissau have began. The first sale prices of cashews that arrive in Bissau are between 450 and 490 XOF/kg. In the vast majority of production areas, prices increased to 400 XOF/kg, although some 350 XOF/kg sales are still observed in the Oio and Bolama regions. In the regions of Cacheu and Bafata some producers can negotiate 450 XOF/kg. Now that the campaign is fully launched, our recommendation is as follows: If you receive an offer of 400 XOF/kg, sell half of your cashews stocks and if 450 XOF/kg are offered, sell all your cashews stocks. Even if the demand is good, it is quite uncertain that prices rise above 450 XOF/kg, so we should avoid looking for higher prices because it is also possible for prices to fall at the beginning of the rainy season.
Robocall 6	09/06/2021	This week, in Bissau, the prices oscillated between 460-475 XOF/kg, ie there was a minimum price growth at 10 XOF/kg and a decrease in the maximum price at 15 XOF/kg, compared to last week. This decrease in the maximum price, next to the scales, is due to the mass arrival of significant amount (between 30,000 - 40,000 tons) of nuts from the regions, resulting from the recent opening of Bissau scales and also, due to the need to free space in the warehouses of the regions, so that there is continuity of the purchases of the intermediaries with the producers. This week, in the regions, overall, producer price variations continued between 350 and 450 XOF/kg with most purchases made at 400 XOF/kg. Like last week, our recommendation is as follows: If you receive an offer of 400 XOF/kg, sell at least half of your cashew stocks and if 450 XOF/kg are offered, sell all your cashews stocks.

Robocall 7	11/06/2021	This week prices are around 475 XOF/kg in Bissau, a slight increase over last week. The demand is strong in the port of Bissau, as in other cashew producing countries. Like the last week, in the regions, producer price variations remain between 350 and 450 XOF/kg with most purchases at 400 XOF/kg. In parts of the Cacheu region the prices have increased up to 475 XOF/kg. Like our recommendation last week: If you receive an offer of 400 XOF/kg, sell at least half of your cashew stocks and if 450 XOF/kg is offered, sell all your cashews stocks.
Robocall 8	18/06/2021	This week, prices are relatively stable in Bissau as in production areas. The demand remains good, but the supply is strong and more than 130,000 tons have already been delivered to the port of Bissau. Bissau prices are still between 465 and 475 XOF/kg and prices in production areas are between 350 and 475 XOF/kg. More precisely, in Bolama and Oio regions, price variations remain between 350 - 400 XOF/kg. Bafata and Biombo have prices of around 400 XOF/kg, i.e. a minimum price rise of 50 XOF/kg compared to last week; In the regions of Gabu, Quinara and Tombali, price variations are between 375 - 400 XOF/kg, i.e. minimum price increases of 25 XOF/kg compared to last week; In the Cacheu region, price variations remain between 350 - 475 XOF/kg. The rains are about to start, which is why our recommendation is now: If you are offerred 400 XOF/kg, sell all your cashew stocks.
Robocall 9	25/06/2021	This week, the shipments of cashews from the harbour finally could start and this caused a slight increase in Bissau prices as in most areas of cashews production. The prices of cashews that arrive in Bissau increased, ranging between 470 and 480 XOF/kg. At the level of production zones, the price of 400 XOF/kg to the producer tends to be most common in all regions. Only the Bolama and Oio regions still record some purchases at 350 XOF/kg, but even in these regions the price of 400 XOF/kg is offered by some buyers. In all other regions, prices reached 400 XOF/kg and in the Cacheu region prices reached 475 XOF/kg in some villages. As the rains are starting, we recommend that you enjoy this situation to negotiate the sale of your latest stocks at 400 XOF/kg and finish your season.
Robocall 10	02/07/2021	This week, the competition among cashew exporters to access the latest remaining stock increased markedly. Prices in the port of Bissau rose between 495 and 500 XOF/kg. This increase in port prices resulted in a slight increase in producer prices in production areas. The price of 400 XOF/kg spread virtually to all zones and prices even increase between 425 and 500 XOF/kg in the Cacheu region. As the campaign approaches the end, we continue to advise to take advantage of this good demand to negotiate at least 400 XOF/kg and sell your remaining stocks.
Robocall 11	09/07/2021	At the end of the past week, the quotations of cashew prices climbed in the international market. Processors in India and Vietnam have been informed that the harvest is coming to an end in West Africa and made large requests from the latest fresh cashew stocks available in Guinea-Bissau. With this increase in demand, Bissau prices have increased even more and are now between 500 and 510 XOF/kg for the cargo arriving in the port. In production areas, producers' purchase prices have increased even more and now range from 400 to 450 XOF/kg, with most purchases at 450 XOF/kg. This unexpected increase is an unexpected opportunity to sell your remaining cashews. Therefore, we recommend that you negotiate 450 XOF/kg to sell your latest fresh cashews stocks. Note that the season will end soon.
Robocall 12	17/07/2021	Now the commercialization of cashews ended in other major producing countries, such as Ivory Coast, Nigeria and Ghana, so Indian buyers and Vietnamese are increasing their requests for Guinea-Bissau cashews. Under these conditions, wholesale prices increased slightly in the port of Bissau and are between 515 and 520 XOF/kg. Producers prices remain between 400 and 450 XOF/kg, with more and more purchases at 450 XOF/kg. As last week, we recommend that you negotiate a price of 450 XOF/kg to sell your latest nuts and finish the season before rainfall decreases quality.
Robocall 13	23/07/2021	This week, the demand remains very strong in the port of Bissau to capture the latest cashew nuts available in Guinea-Bissau. The prices of trucks loaded with cashews delivered in the port of Bissau have increased even more and now between 520 and 540 XOF/kg. In production zones, the latest major sales are between 400 and 500 XOF/kg, with most sales at 450 XOF/kg. We always recommend negotiating at least a price of 450 XOF/kg to sell your latest stocks before rainfall degrade the quality. The season is coming to an end and we will soon be completing the spread of our cashew market information messages.
Robocall 14	29/07/2021	This week, the demand remains very high in the port of Bissau to buy the latest cashews available in Guinea-Bissau. Wholesale cashew prices delivered to the port of Bissau grew a lot and range from 535 to 550 XOF/kg. In production zones, the last major sales were between 400 and 500 XOF/kg, with most sales at 450 XOF/kg. We recommend negotiating at least this price of 450 XOF/kg to sell all your stock, as the campaign is at the end. Despite the end of the cashew trading season, N'Kalô will continue its services, informing you of various issues, which may help you make better decisions.

1.13 Producers' transaction diaries

During our baseline survey, producers in our sample were provided with a simple paper-based template to record their transactions during the 2020 trading season. We refer to this template as a transaction diary.[32] The data-collection team showed producers how to correctly fill in this diary to record the date, quantity sold and price of a transaction.

Importantly, producers recorded the exact day in which they made a transaction. Instead, in our follow-up interviews, we only asked producer to recall in which month they had made a transaction.

During the in-person follow-up interview, the survey team collected the diaries that had been left after the baseline. After data-collection, two operators digitised the data from the sales diaries to check whether they had been correctly used to record the date, quantity and price of sales made during the 2020 seasons. In the rest of this sub-section we use the data recorded in these transaction diaries to provide further evidence on the impact of the intervention.

The information provided by the intervention matched the price dynamic observed in our sample. Appendix Figure 1.9 shows the daily average prices recorded by the sample of producers that filled in the diaries.[33] The vertical lines show the dates in which producers received a message from *n'kalô*. Blue lines show a "bearish" message, that suggested that prices were low and it was better to wait. Orange lines provided a "bullish" message, that suggested that prices were high and it was a good time to sell. The price movements recorded in the transactions are consistent with the information

[32] To encourage the use of the sales diary, the data-collection team promised that producers that kept records could participate in a lottery to win a 50kg bag of rice. Two placebo messages sent to all producers in 2020 reminded producers to use the sales diary.

[33] Appendix Table 1.21 shows that having filled in the diary is not correlated with treatment. Not all producers filled in the diaries. Other producers could not find the diary during the in-person follow-up interview.

that producers received.

The transaction diaries provide further evidence that the effects of our intervention may be due to the messages content, rather than just due to a salience effect, as in Bettinger et al. (2021). For example, we can rule out that producers mechanically sold more frequently when they received messages. As Appendix Figure 1.10 illustrates, there is not a consistent pattern between days in which the messages were delivered and days in which producers recorded making a sale. The message content matters.

In Appendix Table 1.26, we estimate treatment effects on outcomes constructed using the data from the transaction diaries. Several of these effects point in the same direction as those we had estimated for the full sample. Average prices are marginally higher for treated producers, but not significantly so, relative to the control group. The number of sales is not greater for the treatment group relative to the control group, but this may be because we are selecting a group of producers that is already making at least one transaction.

The treatment effect on quantity sold and the value of sales point in the same direction as those estimated for the full sample using data from the in-person follow-up survey. As Appendix Figure 1.10 had alluded, we find no evidence that treated producers sell within days of receiving a message. Instead, we find that treated producers recorded making more transaction in June relative to the control group, a time when prices were at their highest and when the intervention messages had advised producers to sell. We do not find that they sold earlier in the trading season, but this lack of finding may also be because the sub-sample of control producers had recorded more sales in this period relative to the full sample of control producers.

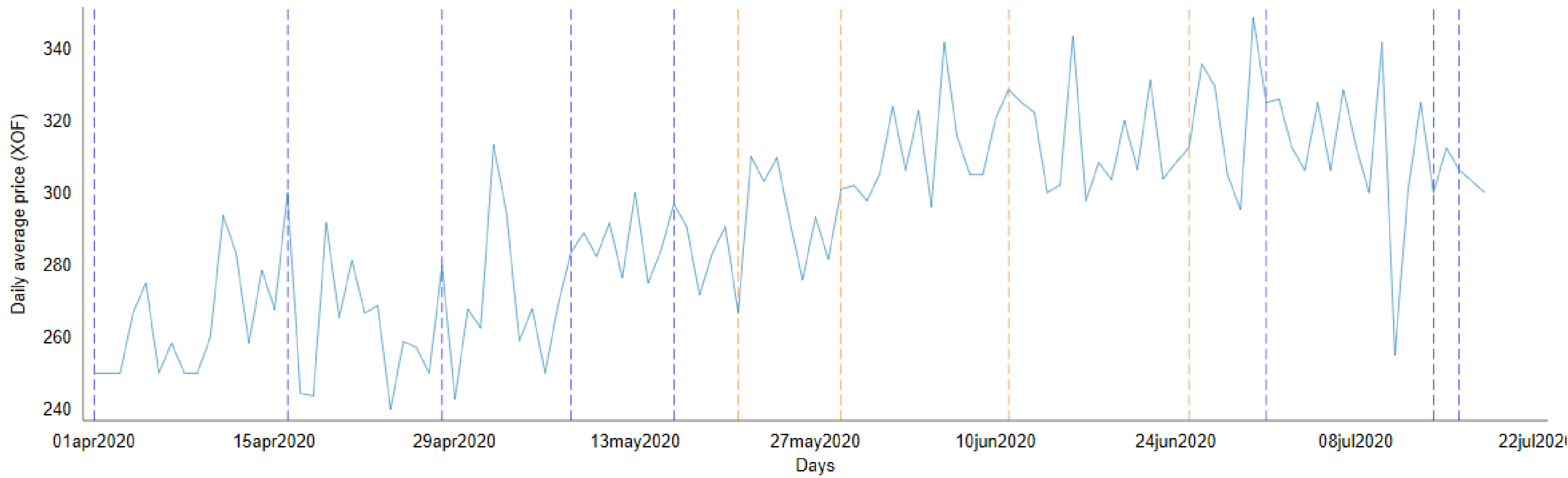


Figure 1.9: Daily average prices recorded in transaction diaries[a]

[a] Daily average price recorded in transaction diaries left with producers after baseline. The vertical lines show the dates in which producers received a message from *n'kalô*. Blue lines show a "bearish" message that suggested that prices were low and it was better to wait. Orange lines show a "bullish" message that suggested that prices were high and it was a good time to sell.

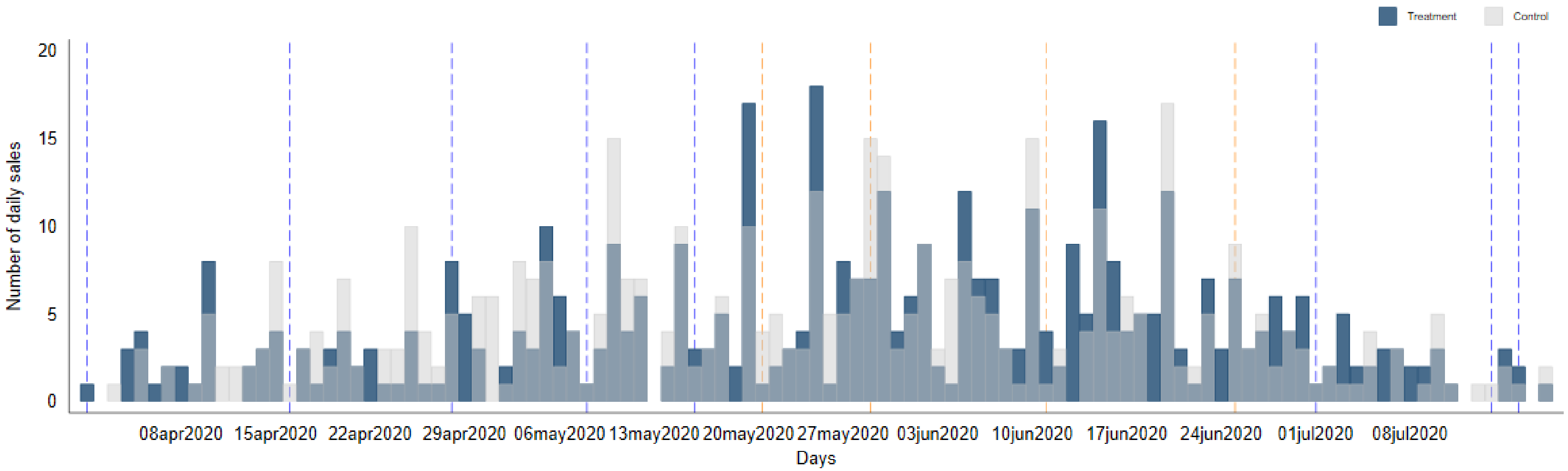


Figure 1.10: Number of daily sales recorded in transaction diaries by treatment[a]

[a] Number of daily sales recorded in transaction diaries left with producers after baseline, between treatment and control producers (spillover group is omitted from the graph). The vertical lines show the dates in which producers received a message from *n'kalô*. Blue lines show a "bearish" message that suggested that prices were low and it was better to wait. Orange lines show a "bullish" message that suggested that prices were high and it was a good time to sell.

1.14 Village-trader outcomes

During our in-person follow-up, we tried to interview an intermediary present in the village at the time the data-collection team surveyed producers. We were not able to survey an intermediary in every village, and almost all of the intermediaries that we interviewed were those that lived permanently in the village. Table 1.27 shows the main outcomes measured at the intermediary-level across treated and control villages. No outcome is statistically different between the two groups of villages. Consistent with the information provided by treated producers, village traders in treated villages record purchasing a larger amount of raw cashew nuts relative to those in control villages, though the difference is not statistically significant, it represents about 10% of the control mean. We also see a reduction in the margins earned by the sample of surveyed intermediaries in treated villages relative to control villages. We are cautious in not over-interpreting this difference because the sample for whom we can observe the resale price is small: many intermediaries reported not knowing the resale price but buying on commission for an exporter in exchange of a fixed income or quantity-based fee.

We also sought information on the number of intermediaries that visited the villages. We asked the village leader, or whom they had delegated, how many itinerant intermediaries had visited the village during the trading season. In line with our findings of positive *between*-village spillovers, we find in Appendix Table 1.28 that control villages received a higher number of visits by itinerant intermediaries.

Table 1.26: Mechanisms — Salience and diary-based measures

	Control (1) Mean (SD) Total obs.	(2) Treatment	(3) Spillover	(4) Treat. vs. spillover
Average price from the diary (XOF/kg)	310.96	3.92	-8.94	12.85**
	(59.06)	(6.14)	(5.97)	(5.81)
	448	[0.52]	[0.54]	[0.11]
Number of sales from the diary	3.00	-0.26	-0.16	-0.10
	(2.45)	(0.27)	(0.26)	(0.20)
	523	[0.44]	[0.67]	[0.62]
Total quantity sold from the diary (kg)	1022.76	259.82	-89.52	349.34
	(1856.30)	(235.80)	(203.95)	(235.91)
	462	[0.44]	[0.67]	[0.19]
Total value of sales from the diary (XOF)	338145.95	85771.11	-29425.56	115196.68
	(775014.25)	(81123.21)	(68545.01)	(74244.75)
	463	[0.44]	[0.67]	[0.19]
Timing:				
1 if sold cashews within two days of a message	0.46	-0.02	-0.00	-0.02
	(0.50)	(0.05)	(0.06)	(0.06)
	514	[0.87]	[1.00]	[0.89]
1 if sold cashews in April 2020 or before	0.27	-0.03	0.01	-0.04
	(0.44)	(0.05)	(0.05)	(0.05)
	514	[0.87]	[1.00]	[0.89]
1 if sold cashews in May 2020	0.58	0.01	-0.01	0.02
	(0.50)	(0.06)	(0.06)	(0.06)
	514	[0.87]	[1.00]	[0.89]
1 if sold cashews in June 2020	0.54	0.10*	0.04	0.06
	(0.50)	(0.06)	(0.06)	(0.05)
	514	[0.47]	[1.00]	[0.89]
1 if sold cashews in July 2020 or later	0.15	0.01	0.02	-0.01
	(0.35)	(0.04)	(0.04)	(0.05)
	514	[0.87]	[1.00]	[0.89]

Notes: Producer-level intention-to-treat (ITT) estimates in 2020 (year 1) from transaction diaries reported in columns 2 and 3. Columns 4 tests for differences in parameters obtained in previous two columns. Outcome variables are listed on the left and described in detail in the pre-analysis plan. The unit of observation is the individual producer. All models control for the randomisation triplet fixed-effects and the baseline value of the outcome when it was available. Standard errors are in parentheses and are clustered at the village-level. Stars on the coefficient estimates reflect unadjusted p-values. Sharpened q-values controlling the false discovery rate across outcomes within each family are shown in brackets. * denotes significance at 10 pct.; ** at 5 pct.; and *** at 1 pct. level. Column 1 displays the control mean; standard deviation; and total number of observations across all groups.

Table 1.27: Village intermediaries outcomes

	(1) Control	(2) Treatment	(3) (1) - (2) [p-value]	(4) N
1 if intermediary was interviewed in 2021	0.42	0.47	-0.05	
	(0.50)	(0.50)	[0.39]	304.00
1 if made any purchase	0.85	0.83	0.02	
	(0.36)	(0.38)	[0.71]	136.00
Number of producers with whom they exchanged	4.23	6.32	-2.09	
	(6.47)	(10.21)	[0.16]	131.00
Quantity purchased (kg)	10626.15	11823.76	-1197.62	
	(17849.26)	(16001.22)	[0.70]	136.00
Purchase price (XOF/kg)	314.63	316.10	-1.46	
	(51.53)	(39.39)	[0.88]	114.00
Resale price (XOF/kg)	395.00	382.71	12.29	
	(71.06)	(81.00)	[0.57]	53.00
Margin (XOF/kg)	74.17	65.57	8.60	
	(67.37)	(78.32)	[0.68]	53.00
1 if made any exchange	0.56	0.60	-0.04	
	(0.50)	(0.49)	[0.66]	136.00
Quantity of cashews obtained from exchanges (kg)	2749.68	3145.35	-395.67	
	(4103.98)	(5085.62)	[0.63]	132.00
Quantity of rice exchanged for cashews (kg)	2545.21	2967.53	-422.32	
	(3942.35)	(4702.57)	[0.58]	133.00

Notes: Column 1 reports the mean in control villages for outcomes related to the 2020 trading season. Column 3 reports the mean in treatment villages. Outcome variables are listed on the left. The data comes from a survey of village-based intermediaries. Due to logistical issues it was not possible to interview an intermediary in every village. Standard deviations are reported in parentheses. Columns 3 reports the difference between columns 1 and 2 and the p-value from a t-test of equality between the means of intermediaries in treatment and control villages in brackets.

Table 1.28: Village intermediaries visits

	(1) Control	(2) Treatment	(3) (1) - (2) [p-value]	(4) N
Number of itinerant traders that visited the village	10.13	6.75	3.38	
	(14.01)	(9.29)	[0.03]	284.00
Number of seasonal traders that stayed in the village	2.81	2.16	0.65	
	(4.54)	(3.32)	[0.20]	284.00
Number of traders based in the village	4.42	4.03	0.39	
	(5.00)	(4.71)	[0.52]	282.00
1 if itinerant traders visited in March	0.07	0.09	-0.02	
	(0.25)	(0.29)	[0.46]	284.00
1 if itinerant traders visited in April	0.62	0.51	0.11	
	(0.49)	(0.50)	[0.07]	284.00
1 if itinerant traders visited in May	0.78	0.68	0.11	
	(0.41)	(0.47)	[0.04]	284.00
1 if itinerant traders visited in June	0.72	0.57	0.15	
	(0.45)	(0.50)	[0.01]	284.00
1 if itinerant traders visited in July	0.10	0.08	0.02	
	(0.30)	(0.28)	[0.66]	284.00

Notes: Column 1 reports the mean in control villages for outcomes related to the 2020 trading season. Column 3 reports the mean in treatment villages. Outcome variables are listed on the left. The data comes from a survey of village leaders. Standard deviations are reported in parentheses. Columns 3 reports the difference between columns 1 and 2 and the p-value from a t-test of equality between the means between treatment and control villages in brackets.

Chapter 2

Labour Market Effects of Ethiopia's Social Safety Net

* *Acknowledgements:* This chapter builds on my MPhil thesis. This research was supported by the CGIAR Standing Panel on Impact Assessment (SPIA) under its program of work 2019-2021. Balliol College and St. Antony's College also provided financial support towards data-collection. I am extremely grateful to James Fenske and Simon Franklin for steering me throughout the early stages of this project; to Pramila Krishnan, Stefan Dercon, Binta Zahra Diop, Clement Imbert, John Hoddinott, Steve Bond, Douglas Gollin, Lukas Hensel and participants of the IFAD's Jobs, Innovation and Value Chains in the age of Climate Change 2022 Conference, SPIA Small Grant Meetings, CSAE Conference 2018, EEA Conference 2018 for stimulating and useful discussions. Many thanks to two anonymous MPhil examiners for their useful comments. My gratitude extends to the unmentioned individuals that supported me during the project. All errors remain my own.

2.1 Equilibrium effects of social protection programmes

Policies aiming to increase the welfare of individuals living in poverty can also affect non-participants by shifting the labour market equilibrium (Bandiera et al., 2017; Mobarak and Rosenzweig, 2014; Bryan et al., 2014; Egger et al., 2022; Imbert and Papp, 2015; Muralidharan et al., 2023). For example, programmes offering skill and asset transfers, rainfall insurance, resettlement, cash transfers, or guaranteed employment schemes may affect labour decisions and wages of individuals not directly targeted by the interventions. A common theme across these studies is that landless workers bear most of the general equilibrium effects of these interventions. However, there is limited evidence on the labour market effects of transfer programmes in markets where the agricultural workforce comprises mostly small landowners and few landless workers. In markets with relatively few landless workers, the general equilibrium effects of these programmes may be diminished.

This paper examines how a large social protection programme can affect non-beneficiaries through changes in local labour markets. I focus on Ethiopia's Productive Safety Net Programme (PSNP), which provides cash or food transfers conditional on public works participation to over ten million beneficiaries annually. It is one of the largest rural social protection systems in Africa reaching almost 10% of the population (Gilligan et al., 2009). Its impressive scale has contributed to making it a frequently used reference in international comparisons of similar programmes in policy circles.[1] As such, rigorous evaluations of this programme can provide insights that are of significance both within and outside the Ethiopian context. By analysing the district-level exposure to the programme, I aim to provide a first assessment of how this programme affects labour markets, moving beyond the individual-level effects that have so far been the

[1] See, for example, Alderman and Yemtsov (2012), Grosh et al. (2008), McCord (2013), and Subbarao et al. (2013).

focus of previous evaluations (Subbarao et al., 2013).

To identify the main effects of the programme on local labour markets, I estimate a difference-in-differences model. I investigate whether the programme has affected employment participation, occupational categories, hours worked and wages in the targeted districts, relative to those that did not receive the programme. I use a unique geo-referenced dataset combining three cross-sections of the National Labour Force Survey, observing over 400,000 individuals in all regions of Ethiopia and spanning from 1999 to 2013.[2] I complement this main source of information with other geo-referenced datasets: village-level census data, climatic variables and the district-level historical frequency of aid receipts. To disentangle the effects within treated districts, I also employ the Ethiopian Socio-Economic Surveys (ESS), which include household and community data on PSNP participation.[3]

I have two main results on how the programme affects labour market outcomes. First, I find a reallocation of the workforce towards non-agricultural self-employment (5 percentage point increase) in targeted districts.[4] The reallocation towards non-agricultural self-employment is driven by women in my sample. To unpack this result, I descriptively compare the labour market outcomes of non-beneficiaries in targeted districts with individuals in untargeted communities within those districts. Using the ESS, I describe that non-beneficiaries in untargeted communities in PSNP districts experienced a larger shift away from agriculture towards other forms of self-employment. Second, I find no impact on the extensive and intensive margins of labour supply or wages in rural districts targeted by the programme. These results holds across several robustness checks. Conducting a placebo test with pre-programme data shows parallel trends

[2] Since I only focus on two periods for my main analysis, my identification strategy is not affected by the potential biases of staggered difference-in-differences models with more periods (de Chaisemartin and D'Haultfoeuille, 2022).

[3] The ESS allows me to identify the communities targeted by the programme and combine it with district-level data, but it covers fewer districts and was not collected before the programme started.

[4] Throughout the article I interchangeably refer to districts as *woredas*.

in outcomes between targeted and untargeted districts prior to the program's start, supporting the validity of the findings. Moreover, including additional demographic controls does not alter the results. Overall, these results are consistent with the idea that the PSNP stimulated demand for local goods and market access.

This paper contributes to the literature on the impacts of public works programmes on rural economies.[5] My study is most closely related to the work of Imbert and Papp (2015), who also use a difference-in-differences model to estimate the effect of India's Mahatma Gandhi National Rural Employment Guarantee Act (NREGA) on wages and employment.[6] More recently, Muralidharan et al. (2023) document the substantial general equilibrium effects of NREGA. In contrast with the evidence from India, my findings suggest that wages of private sector labourers do not seem to respond significantly to the presence of public works programmes. The difference is likely due to factors such as programme design or structural differences in the labour markets analysed. Importantly, unlike NREGA, the PSNP transfers were set below the prevailing market wage. This decision was made in order to minimise the risk of creating a disincentive for participation in other productive activities. Wages in the Indian employment guarantee scheme are generally above the private sector wage for casual labourers (Subbarao et al., 2013).

This paper also contributes to our understanding of the broader impacts of the PSNP. Previous studies on the PSNP focused on estimating the impact of the programme only on individual beneficiaries, collecting information exclusively in targeted districts (Berhane et al., 2011, 2014; Gilligan et al., 2009, 2011; Hoddinott et al., 2011, 2012). But, as McCord and Slater (2013) point out, enlarging the unit of analysis beyond the beneficiary-level is of particular relevance to a programme like the PNSP, which aims

[5] See Bagga et al. (2023) for a review on the effects of workfare programmes. Besley and Coate (1992), Ravallion (1991), and Basu (2013) also provide theoretical treatments of workfare programmes.

[6] Other recent examples of papers estimating the labour market impacts of NREGA are Berg et al. (2018), Zimmermann (2020), Fetzer (2020) and Santangelo (2019).

to benefit the whole community. For example, one of the aims of the programme is to increase resilience and agricultural productivity within the whole community targeted so as to stimulate production and local market activities of food and non-food products (World Bank, 2014). Studies that explore effects beyond the individual-level, to analyse the district-level impacts of the programme, remain scant. Filipski et al. (2016) documented that "new income created by PSNP benefits households that do not receive cash transfers; these non-beneficiaries benefit [...] through local and national markets", which is consistent with how I interpret my results. In contrast to earlier work, this paper unpacks some of those general-equilibrium patterns to further document the potential micro-level spillovers of the programme, focusing on labour markets.

Two recent studies on the wider effects of the PSNP are complementary to my analysis. First, Gazeaud and Stephane (2023) find little evidence on the effectiveness of the PSNP public works in improving the agricultural productivity in districts targeted by the PSNP. Abebe et al. (2021) study the effects of Ethiopia's Urban Productive Safety Net Program, which provides employment on local public works to the urban poor, and was rolled out randomly across neighbourhoods of Addis Ababa. They find that the programme increased public employment, reduced private labour supply among beneficiaries, and increased overall private wages by 18%. My work complements both of these studies since I focus on the *rural* PSNP, rather than its *urban* counterpart, and because I focus on the labour market and household-level decisions, rather than aggregate yields data.

In a broader sense, this paper contributes to the literature studying the functioning of rural labour markets in low- and middle-income countries (Behrman, 1999). It aims to enhance our understanding of how labour markets in low- and middle-income countries respond to in-kind and cash transfer programmes. Recent papers have shown that households change their labour supply decisions in response to the provision of different in-kind assets, such as land-titles, better housing conditions, roads, electrification, and agricultural inputs (Field, 2007; Dinkelman, 2011; Franklin, 2020; Asher and Novosad, 2020; Moneke, 2020; Diop, 2023). Aid or cash transfers have been found to have null or positive effects on the labour supply of recipients.[7]

This paper is organised as follows. Section 2.2 provides details on the targeting of the programme, the main dataset and outcome variables analysed, presents summary statistics, and outlines the main empirical strategy. Section 2.3 presents and discusses the main results, along with several robustness checks. Section 2.4 concludes.

2.2 Background, data, and empirical strategy

This section briefly describes how Productive Safety Net Programme (PSNP) participants are targeted, the data employed, the main empirical strategy, and descriptive statistics. My primary analysis leverages repeated cross-sections of the Ethiopian National Labour Force Survey (LFS). I match district identifiers across rounds of the LFS to construct my main outcome and control variables. Further, I combine additional data sources to construct district-level covariates on: (i) the geographical assignment of the programme, (ii) the frequency of relief assistance received prior to the PSNP, (iii) rainfall, (iv) temperature, and (v) population density. In a secondary complementary analysis, I employ a panel dataset from the Ethiopian Socio-Economic Surveys (ESS). Appendix 2.6 provides more institutional details about the programme. Appendix 2.7 further describes all data sources and the strategy used to match districts across waves of the LFS.

[7] For example, Egger et al. (2022) find that an unconditional cash transfer programme in Western Kenya that injected about 15% of GDP had a positive effect on labour demand. Banerjee et al. (2017) find no evidence of disincentive effects among transfer recipients by combining datasets from seven randomised controlled trials from different countries. In Ethiopia previous food aid programmes were found not to disincentive work among recipients (Abdulai et al., 2005; Quisumbing and Yohannes, 2005).

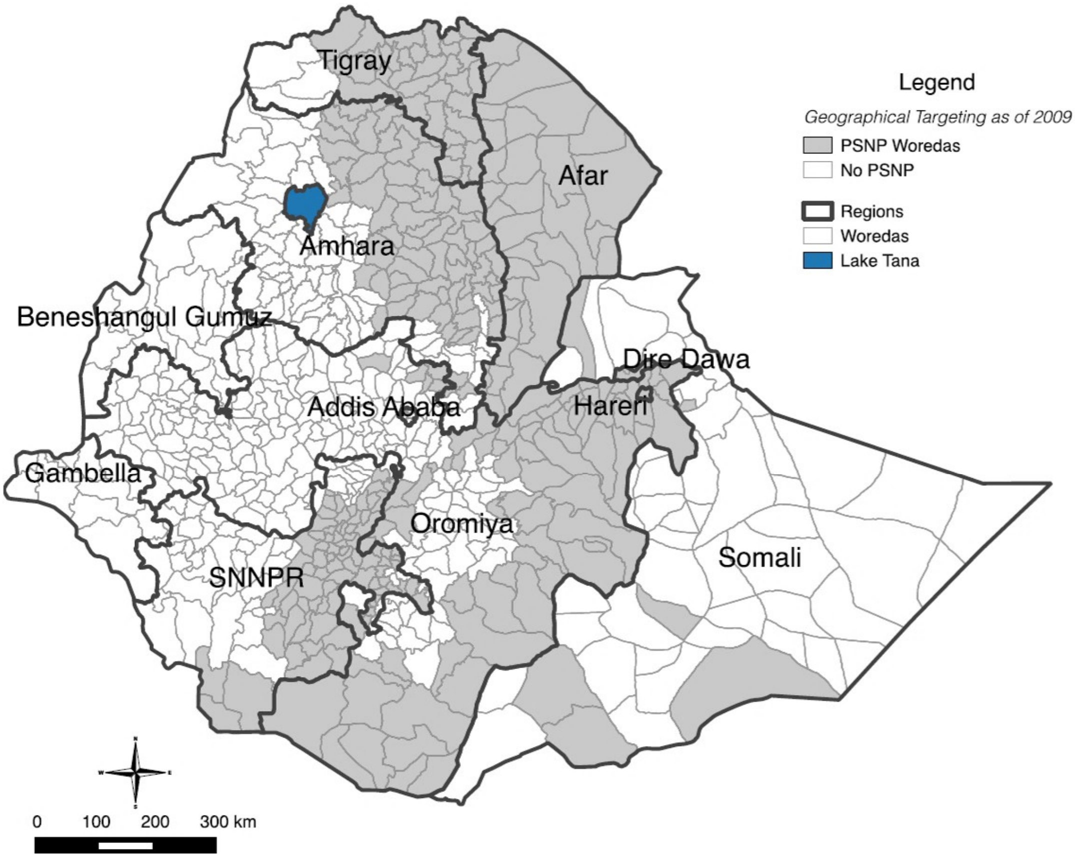


Figure 2.1: Productive Safety Net geographic targeting[a]

[a] *Notes:* PSNP assignment of 290 districts (*woredas*), as of the end of PSNP Phase II (2006-2009).

2.2.1 PSNP targeting

The PSNP targets districts based on the historical allocation of food aid prior to 2005.[8] Within targeted districts, local officials and community leaders select beneficiaries by constructing lists of eligible households for each community within the district. The main eligible beneficiaries of the programme are chronically food insecure households. Targeted beneficiaries participate in public works in exchange for cash or food transfers, while some receive unconditional support based on their circumstances. In 2009, cash or food transfers conditional on public works participation comprised 84% of the total transfer to beneficiaries (World Bank, 2010b).

The PSNP was initially introduced in 192 rural districts (*woredas*) in 2005 and expanded to 290 districts by the end of 2009, forming the treated sample for my analysis. Figure 2.1 displays the geographical distribution of these targeted districts, sourced from the programme reports (World Bank, 2010a).

2.2.2 Data and main outcomes

For the main analysis, I combine three nationally representative cross-sections of the Labour Force Surveys, conducted in 1999, 2005, and 2013. I restrict my sample to individuals living in rural regions, since the PSNP only targeted rural areas in this period.[9] Figure 2.2 provides a timeline of the programme phases during the years I analyse. The primary sample analysed consists of individual-level observations from a balanced panel of 453 rural districts from the 2005 and 2013 rounds, illustrated in Appendix Figure 2.3.

The main outcomes are measures of employment, on the intensive and extensive margins, wages, and occupational categories. First, I categorise individuals aged 17 to 65

[8] Specifically, the 2006 Project Implementation Manual states that a *woreda* was eligible for the programme if it was: '[i] in one of 8 regions (Tigray, Amhara, Oromiya, SNNPR, Afar, Somali, rural Harari and Dire Dawa), and [ii] has been a recipient of food aid for a significant period, generally for at least each of the last three years' (GFDRE, 2006, pp.3). The same criterion is reiterated in the 2010 revised version of the PIM, which also adds that in 2004 eligibility was defined more broadly, but was later revised. The previous broader eligibility criteria would have deemed *woredas* eligible based on 'the frequency with which they required food assistance in the ten years preceding the design of the PSNP (the ten years up to 2004)' (GFDRE, 2010, pp.7). It is not clear how many years were deemed enough in the broader criterion, and to what extent the revised one was followed.

[9] The CSA defines as non-rural all (enumeration) areas with a population of more 1000 individuals, and any administrative capitals (regional, zonal or district capitals) regardless of population. More information on the survey design is available on `http://tinyurl.com/csa-nlfs2013`, visited on the 14/04/2016.

PSNP Phases	Phase I	Phase II				Phase III			
PSNP geographic expansion	PSNP launched in 192 woredas	262 PSNP woredas, as district split	290 PSNP woredas, as Afar added			319 PSNP woredas, as Somali added			
Year	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
Main data sources	Labour Force Survey 2005 Round		2007 Census						Labour Force Survey 2013 Round

Figure 2.2: Timeline of the PSNP and data sources[a]

[a] *Notes:* The Productive Safety Net Programme (PSNP) was launched in 2005, in a testing phase in 192 districts (*woredas*). In the second phase (2006-2009), the number of districts targeted reached 290, which I refer as the main sample of treated districts. In 2010, 49 districts from the Somali region were added to the programme.

as currently employed, unemployed, or inactive.[10] Currently employed individuals are those who reported engaging in productive activities for at least one hour in the week before the interview, following the ILO definition (Hussmanns, 2007). For unemployment, I consider individuals as unemployed if they are not currently employed but are available for work and willing to take up a job opportunity, even if they have not actively searched for work in the last three months, as in Franklin (2014) and Broussar and Tekleselassie (2012). Second, I construct individual indicators for those employed, such as hours worked in the last seven days, engagement in additional working activities, and willingness to work more hours. Third, I also create a measure of real monthly wages for manual labourers, using regional deflators from Headey et al. (2012). Fourth, I create main occupation categories grouping the International Organization for Standardization (ISO) codes available in the LFS.[11] These occupational categories allow me to estimate transitions related to the primary source of livelihood, but they do not fully capture the diverse range of activities that individuals in rural Ethiopia may be engaged in beyond their main employment activity (Dercon and Krishnan, 1996).

2.2.3 Summary statistics

Table 2.1 shows the means of the main covariates used in the analysis for both PSNP districts (Column 1) and non-targeted districts (Column 2). Column 3 displays the *p*-value from a *t*-test of equality of means between the two groups.

Panel A shows some differences in labour market conditions between PSNP and non-PSNP districts in the baseline year of comparison (2005). PSNP districts show lower fractions of workers engaged in agriculture and seasonally unemployed individuals, but higher fractions of manual labourers and public sector workers. Observable demographic characteristics are balanced between the two groups of districts. Measures of human capital and household demographics show no significant differences. The frequency of relief assistance prior to 2005 is higher in PSNP districts, as expected, since this variable was used in targeting the programme at the district-level. Weather conditions also differ, with PSNP districts experiencing less rainfall on average.

Table 2.2 shows that despite differences in control variables, the outcomes of interest are balanced between districts targeted by the PSNP and other districts in 2005. Most individuals (around 82%) are employed, with the majority being self-employed in agriculture (crop, livestock, mixed-farming, or forestry). Of those employed, 10-13% are in self-employment outside of agriculture, usually working in trade or crafts work.[12] Public and private sector labourers constitute a relatively small category of employment.

[10] The age cutoffs were chosen based on the Programme Implementation Manual specification that individuals below 17 years of age should not participate in public works, which is in line with findings from the recent programme evaluation Berhane et al. (2011). The manual also specifies that elderly should not participate in the programme, without specifying an age. Thus, the upper age cutoff is chosen so as to follow previous studies of the labour supply responses of food aid programmes in Ethiopia (e.g. Abdulai et al. (2005) and Quisumbing and Yohannes (2005)).

[11] The ISO codes can be found on `http://tinyurl.com/csa-isco08`, accessed on 09/05/2016.

[12] This occupation is more common among women, with 22% of working women engaged in non-agricultural activities, while the proportion of working men in this category is 6%.

Labourers undertake relatively low-skill tasks usually in agriculture or construction work. Public labourers may include PSNP participants, as well as labourers in other publicly funded projects. The additional outcome variables related to the intensive margin of labour supply are also balanced between the two groups of districts.

2.2.4 Empirical strategy

My main identification strategy compares changes over time in targeted districts with changes in other districts. To improve identification, I include as controls the variables in Panel A of Table 2.1 to account for differential dynamics across districts. These controls encompass the frequency of aid receipts and district-level labour market conditions, which are interacted with a time varying indicator, and time-varying district controls related to rainfall and temperature.[13] As an additional specification, I also include individual-specific controls to improve efficiency, though since treatment is a the district-level those controls are not needed for identification. District-level averages of the individual-level controls are presented in Panel B of Table 2.1.[14]

For my main analysis, I estimate a difference-in-differences specification across two periods, with the 2005 wave being the pre-treatment period and the 2013 wave being the post-treatment period. Since I only focus on two periods, I am not concerned with potential biases of staggered difference-in-differences models with more periods (de Chaisemartin and D'Haultfoeuille, 2022). Hence, I estimate the following linear specification:

[13] The district-level controls of labour market conditions are obtained averaging the individual-level observations from the Labour Force Survey at the round-specific district-level. Temperature, rainfall, and frequency of aid between 1995 and 2004 are only observed at the district-level. More details on these variables are in Appendix Section 2.7.

[14] Individual-level controls comprise: age; indicators for whether the individual is female, their level of education (omitting no schooling), married, or the household head; indicators for whether the individual's household has someone aged five years of age or below, someone aged 70 or above, someone with a disability, a female household head, or a household head with any schooling.

Table 2.1: Summary statistics — Mean balance of district controls in 2005

	PSNP (1)	Control (2)	p-value (3)	Source (4)	Time-Varying? (5)
Panel A: District-level controls					
Female labour force participation rate	0.77	0.78	0.491	2005 LFS	No
Male labour force participation rate	0.92	0.92	0.964	2005 LFS	No
Literacy rate	0.27	0.27	0.820	2005 LFS	No
Fraction in-migrants	0.04	0.04	0.482	2005 LFS	No
Fraction disabled	0.02	0.03	0.348	2005 LFS	No
Fraction female headed household	0.16	0.16	0.561	2005 LFS	No
Fraction working in agriculture	0.73	0.77	0.018	2005 LFS	No
Fraction of workers seasonally not at work	0.02	0.03	0.001	2005 LFS	No
Fraction public employees	0.03	0.01	0.003	2005 LFS	No
Fraction private employees	0.02	0.03	0.483	2005 LFS	No
Fraction labourers	0.03	0.01	0.057	2005 LFS	No
Cumulative Belg season rainfall (standardized)	0.11	0.46	0.000	GPCC	Yes
Cumulative Meher season rainfall (standardized)	-0.44	-0.16	0.000	GPCC	Yes
Average Belg season temperature (standardized)	0.26	0.38	0.011	UDel_AirT	Yes
Average Meher season temperature (standardized)	0.07	-0.04	0.001	UDel_AirT	Yes
Years of emergency assistance (1995-2004)	7.68	1.69	0.000	NDRMC	No
Panel B: Individual-level controls					
Age	34	33	0.569	2005 LFS	Yes
Fraction female	0.52	0.53	0.941	2005 LFS	Yes
Fraction with some schooling	0.15	0.15	0.947	2005 LFS	Yes
Fraction with primary schooling	0.03	0.03	0.893	2005 LFS	Yes
Fraction with some secondary schooling	0.06	0.07	0.973	2005 LFS	Yes
Fraction with secondary schooling or more	0.01	0.02	0.821	2005 LFS	Yes
Fraction married	0.72	0.72	0.969	2005 LFS	Yes
Fraction of households with no children below age 5	0.02	0.03	0.923	2005 LFS	Yes
Fraction of households with elderly above age 70	0.05	0.05	0.936	2005 LFS	Yes
Fraction of households with individuals with a disability	0.09	0.11	0.664	2005 LFS	Yes
Fraction of household heads	0.44	0.44	0.964	2005 LFS	Yes
Fraction of female household heads	0.16	0.16	0.928	2005 LFS	Yes
Fraction of household heads with primary education, or more	0.10	0.11	0.735	2005 LFS	Yes
Fraction of household heads with some schooling, below primary	0.19	0.19	0.932	2005 LFS	Yes
District Observations	215	238			
Individual Observations	31574	26805			

Notes: Panel A presents means of the district-level controls used in the main regression model for different samples. Column 1 includes controls for districts that were targeted by the PSNP. Column 2 includes district controls for districts that were not targeted by the PSNP (which form the control group). Column 3 presents the *p*-values of the student's *t* *t*-test of equality of means. Standard errors for the student's *t*-test of equality of means are computed assuming correlation of individual observations within each district in a given year. The LFS controls are computed using the 2005 Labour Force Survey round, with sampling weights adjusted for boundary changes. The sample is restricted to individuals of ages between 17-65, using information from the usual activity reported. Cumulative rainfall is expressed as the standardized deviation from the 1979-2014 mean cumulative rainfall during the rain seasons for the *Meher* harvest (June-October) and Belg harvest season (February-May). Temperature is calculated as the standardized deviation from the 1979-2014 monthly averages for the respective pre-harvest rainy season. Years of assistance refers to the frequency in years between 1994-2004, of emergency assistance received by district.
Panel B presents means of the individual-level means. Apart from age, all controls are indicator variables. The omitted category is a male individual with no schooling, unmarried, who is not a household head, and living in a male-headed household, where the household head has no schooling, there are children aged below 5, and no member of the household is above 70 years of age, or has a disability

Table 2.2: Summary statistics — Mean balance of district controls in 2005

	PSNP (1)	Control (2)	p-value (3)
Main Outcome Variables			
Employed (%)	81.8	83.1	0.731
Self-employed in ag. (%)	81.8	86.4	0.185
Self-employed not in ag. (%)	13.1	10.2	0.338
Public sector labourers (%)	1.0	0.1	0.175
Private sector labourers (%)	0.9	1.2	0.766
Unemployed (%)	1.6	1.8	0.852
Inactive (%)	16.6	15.1	0.671
Additional Outcome Variables			
Total hours worked in main occupation in the last 7 days	27.4	26.6	0.619
Underemployed (%)	30.0	28.2	0.676
Has more than one productive activity (%)	22.3	18.9	0.386
Total hours worked in the last 7 days	30.1	28.5	0.342
Private sector labourers' monthly real wage	350.0	347.4	0.950
In-migrants (%)	5.6	7.6	0.403
Household size	5	5	0.700
District observations	215	238	
Individual observations	31574	26805	

Notes: This table presents means of the outcome variables for different samples. All samples are restricted to persons aged 17 to 65. Column 1 only includes districts that were targeted by the PSNP. Column 2 only includes districts that were not targeted by the PSNP (which form the control group). Column 3 presents the p-values of the student's t-test of equality of means in columns 1 and 2. Standard errors for the student's t-test are computed assuming correlation of individual observations within each district.

$$
\begin{aligned}
Y_{idt} =& \beta \times (\mathbb{1}_{(PSNP=1)} \times \mathbb{1}_{(t=2013)}) + (\mathbf{C}_d \times \mathbb{1}_{(t=2013)})'\delta + \\
\mathbf{X}'_{dt}\theta + & \mathbf{H}'_i\zeta + \eta_d + \gamma \times \mathbb{1}_{(t=2013)} + \epsilon_{1,idt}
\end{aligned}
\tag{2.1}
$$

where Y_{idt} is the outcome for individual i in district d in year t. $\mathbb{1}_{(PSNP)}$ is an indicator equal to one if the district is targeted by the PSNP. $\mathbf{C}_d$ and $\mathbf{X}_{dt}$ are vectors of time-invariant and time-varying district controls, respectively. The indicator $\mathbb{1}_{(t=2013)}$ is equal to one for the year 2013 (the first LFS round after the start of the programme), which accounts for any aggregate-level factors affecting all districts in that specific year. $\mathbf{H}_i$ is a vector of individual controls. η_d is a district-specific fixed effect capturing unobserved characteristics of districts that do not change over time. The unobserved idiosyncratic component is denoted by $\epsilon_{1,idt}$. The coefficient β quantifies the effect of the PSNP. I present estimates of β with and without the individual-level controls. I cluster standard errors at the district level, following Bertrand et al. (2004).

To interpret the estimates at the district-level, I adjust individual observations using sampling weights to account for potential under-representation of larger districts in the data, following Imbert and Papp (2015).[15]

2.3 Results

This section presents the estimated effect of the PSNP on the labour market outcomes of individuals living in districts that were targeted by the programme. I find that the programme shifted self-employed individuals from agricultural to non-agricultural

[15] These weights, provided by the Central Statistical Authority (CSA), reflect the inverse probability of being sampled. The weighting strategy ensures that the sum of all weights within a district-year is constant over time for each district and proportional to the sampling weight of the rural population within that district. Additionally, I present unweighted estimates for robustness, but the results remain unaffected by the weighting strategy.

occupations. However, I do not find a significant impact on wages or employment.

2.3.1 Labour supply (extensive margin) and sectoral occupation

Table 2.3 presents the main results on the effect of the PSNP on employment and occupational categories. The coefficients are presented as percentage changes in the fraction of workers in each category.[16] I find no significant impact of the PSNP on labour market participation in targeted districts between 2005 and 2013. Adding individual-level controls reduces standard errors but does not change the overall results (Panel B). Finding no effect on the extensive margin of labour supply is unlikely to be spurious given the size of the program. The 90% confidence interval for the PSNP's effect on employment rate ranges from -4.3 to 3.2 percentage points, indicating little to no impact.

In the last four columns of Table 2.3, the results show changes in the composition of the labour force in PSNP-targeted districts. Specifically, there is a considerable increase of around 5 percentage points (p-value: 0.01) in the share of workers engaged in non-agricultural self-employment, from a baseline of 13 percentage points in the control group. Moreover, I also find a 0.29 percentage points increase in public sector labourers in targeted districts, statistically significant at the 10% level, consistent with a potential increase in PSNP participants engaged in public works.

To better understand the sectoral shifts, Appendix Table 2.8 reports the estimates for men and women, separately. This analysis shows that the increase in the share of workers engaging in non-agricultural activities is driven by women. All estimated coefficients for men are not statistically significant, except for a potential increase in unemployment rates, which is only significant at the 10% level (Panel B).

[16] To improve the readability of the tables, the indicator variables are multiplied by 100.

Table 2.3: Effects on employment participation and sectoral composition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. No individual controls							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	-0.575	0.978	-0.403	-6.359**	5.471**	0.018	0.292*
	(2.276)	(0.659)	(2.061)	(2.617)	(2.149)	(0.433)	(0.167)
Mean Dep. Var.	83.18	1.7	15.12	84.25	11.54	1.33	0.49
Observations	105,323	105,323	105,323	86,779	86,779	86,779	86,779
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	No	No	No	No	No	No
Panel B. Individual controls added							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	-0.16	0.936	-0.776	-5.826**	5.286**	-0.008	0.310*
	(2.277)	(0.655)	(2.066)	(2.427)	(2.122)	(0.434)	(0.168)
Mean Dep. Var.	83.18	1.7	15.12	84.25	11.54	1.33	0.49
Observations	105,323	105,323	105,323	86,779	86,779	86,779	86,779
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each cell reports an estimate of β for different dependent variables; standard errors in parenthesis are clustered at the district level. Each column has a different dependent variable. In Panel A, each model includes district fixed effects and district controls. In Panel B, each model includes district fixed effects, district controls and individual controls. The sample consists of individuals aged 17-65, pooling data from the 2005 and 2013 LFS rounds. Columns (4)-(7) restrict the sample only to those that are currently employed. Individual observations are weighted by sampling weights that are proportional to district population. All models are estimated using ordinary least squares. The means of district-level and individual-level controls are shown in Table 2.1. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

2.3.2 Demographic structure and labour supply (intensive margin)

Table 2.4 shows that the impact of the PSNP on the demographic composition and the intensive margin of labour supply is minimal and not statistically significant. The programme's effects on household in-migration rates and household size are small in magnitude and never close to statistical significance (Columns 1-3).[17] This finding implies that changes in the targeted districts' demographics are not likely to have confounded any labour market effects of the program.

When focusing on the currently employed individuals (Columns 4-7), I find no significant effect on different measures of the intensive margin of labour supply. I find no change in underemployment, changes in working hours during the lean agricultural season (at the time the survey took place), or engagement in more than one form of employment. The estimated effects are small in magnitude (below 10% of the untreated district mean) and not statistically significant. Overall, these findings do not provide evidence that the programme crowded out alternative forms of employment at the district-level.

2.3.3 Effects on the wage of private sector labourers

In Table 2.5, I focus on the wages of private sector labourers as they could theoretically be most influenced by the programme's general equilibrium effects. The coefficient in column 1 indicates a 31% reduction in wages in PSNP districts compared to control districts, but this effect is not statistically significant at the 10% level.[18]

It is challenging to determine if the observed private labourers' wages are represen-

[17] Household size is the integer number of household members. The second and third column report indicator variables (multiplied by 100) on whether the household has had at least one in-migrant in the last five or ten years.

[18] I use Kennedy (1981)'s transformation and interpret the estimated percentage effect on continuous variable measured in logs when a district switches from control to treatment as $100 \times [\exp(\hat{\beta} - 0.5 \times \hat{V}(\hat{\beta})) - 1]$, assuming normality of the errors.

Table 2.4: Effects on demographic composition and intensive margin of labour supply and unemployment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. No individual controls							
dependent variable:	Household Size	In-migrant (last 5 years)	In-migrant (last 10 years)	Underemployment	Has more than one activity	Hours worked in main activity	Hours worked in all activities
	0.015	-0.030	0.927	-4.004	-0.618	-0.328	-0.633
	(0.125)	(0.872)	(1.276)	(3.380)	(2.921)	(0.965)	(0.952)
Mean Dep. Var.	5.232	3.771	6.518	37.04	27.05	30.98	39.79
Observations	105,323	105,323	105,323	86,779	86,779	86,779	86,779
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	No	No	No	No	No	No
Panel B. Individual controls added							
dependent variable:	Household Size	In-migrant (last 5 years)	In-migrant (last 10 years)	Underemployment	Has more than one activity	Hours worked in main activity	Hours worked in all activities
	0.012	-0.348	0.639	-3.881	-0.574	-0.242	-0.560
	(0.095)	(0.800)	(1.194)	(3.375)	(2.908)	(0.962)	(0.941)
Mean Dep. Var.	5.232	3.771	6.518	37.04	27.05	30.98	39.79
Observations	105,323	105,323	105,323	159,902	159,902	159,902	116,321
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each cell reports an estimate of β; standard errors in parenthesis are clustered at the district level. Each column reports an estimate for a different dependent variable. Household size indicates the number of individuals normally residing in an household. In-migrant is an indicator variable equal to one if the individual has migrated into the district in the last 5 years (Column 2), or the last 10 years (Column 3). Columns (4)-(7) are conditional on being employed: the dependent variable in column (5) is an indicator variable equal to one if the individual has reported willingness to work more hours. The dependent variable in column (6) is a dependent variable equal to one if the individual has engaged in more than productive activity in the last seven days. The dependent variable in column (6) and (7) are in levels. In Panel A, each model includes district fixed effects and district controls. In Panel B, each model includes district fixed effects, district controls, and individual controls. The sample consists of individuals aged 17-65, pooling data from the 2005 and 2013 LFS rounds, sampled in 453 districts in each round. Individual observations are weighted by sampling weights that are proportional to district population. All models are estimated using ordinary least squares. The means of district-level and individual-level controls is shown in Table 2.1.* denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

tative of rural wages in Ethiopia, as wage employment may be more prevalent than official statistics suggest (Rizzo, 2011). The wage data is only observed for 1% of the sample, making it indicative rather than representative of rural markets in Ethiopia, as I do not observe wages for those self-employed. Due to potential selection bias and a small sample size, these results are illustrative, and no causal relationship is claimed. With these caveats in mind, my analysis finds no significant changes in private sector labourers' wages. To partly address the potential selection bias within the labourers' sub-sample, I show estimates of the programme's effect on various outcomes in columns 2 to 8 for this sub-sample. The sub-sample of labourers does not seem to be significantly affected by the programme across different measures of employment, although the effect sizes differ from the rest of the sample.

2.3.4 Discussion

There are two take-aways from my analysis. First, I find that a considerable proportion of workers, especially women, tend to transition from agriculture to non-agricultural self-employment activities as a result of the programme. Second, there is no significant effect of the programme on the local labour supply, considering both intensive and extensive margin.

I speculate that the observed shift towards non-agricultural self-employment in PSNP districts could be attributed to improved market access facilitated by the community assets built by the public works component of the programme, such as rural roads. This explanation aligns with previous research highlighting the importance of rural roads on labour market participation in rural areas (Dercon et al., 2009; Asher and Novosad, 2020). Based on this interpretation, the effect of the programme does not seem to distort the labour market or crowd out other employment opportunities but instead stem from the positive externalities generated by the assets created through

Table 2.5: Effects on private sector wage labourers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
dependent variable:	(log) Real monthly wage	Household Size	In-migrant (last 5 years)	In-migrant (last 10 years)	Underemployment	Has more than one activity	Hours worked in main activity	Hours worked in all activities
	-0.289	0.571	-19.526	-3.294	-14.809	-19.797	-1.173	-4.148
	(0.416)	(0.744)	(12.030)	(15.879)	(18.944)	(17.516)	(7.244)	(6.453)
Mean Dep. Var.	5.447	5.390	19.29	25.04	41.40	29.10	39.79	42.97
Observations	932	932	932	932	932	932	932	932
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each cell reports an estimate of β; standard errors in parenthesis are clustered at the district level. Each column reports an estimate for a different dependent variable. (log) Real monthly wage is computed is deflated to 2011 real prices using CSA regional deflators. Household size indicates the number of individuals residing in an household. In-migrant is an indicator variable equal to one if the individual has migrated into the district in the last 5 years (Column 3), or the last 10 years (Column 4). Columns The dependent variable in column (5) is an indicator variable equal to one if the individual has reported willingness to work more hours. The dependent variable in column (6) is a dependent variable equal to one if the individual has engaged in more than productive activity in the last seven days. The dependent variable in column (7) and (8) are in levels. The sample is restricted to private sector labourers aged 17-65, pooling data from the 2005 and 2013 LFS rounds, sampled in 453 districts in each round. There are only 81 districts where private sector labourer's are observed in both rounds. Individual observations are weighted by sampling weights that are proportional to district population. The means of district-level and individual-level controls is shown in Table 2.1.* denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

the public works component.

There are at least two possible explanations for the lack of labour supply response. First, the programme's design aims to support food-insecure households to become self-sufficient rather than creating new job opportunities. The capped days of employment for participants ensure time for other potential productive activities without replacing existing livelihood sources. Second, the labour supply in rural areas may be almost perfectly inelastic, in line with the experimental findings of Goldberg (2016) in Malawi. This interpretation is in stark contrast to the anecdotal assumption that this elasticity can be assumed to be infinite (Lewis, 1954). An inelastic labour supply could explain the muted response in terms of work's extensive margin, which would occur even if the PSNP increased the reservation wages among some rural workers.[19] My findings suggest that effects of social programmes into the labour market, like those found in NREGA, are less likely to be observed in the Ethiopian rural context where wage markets are thinner.

2.3.5 Unpacking within-district heterogeneity

Finally, I investigate whether the district-level effects of the programme are driven by changes within-targeted communities (*kebeles*) or whether these effect are due to changes in untargeted *kebeles* within targeted districts. To do so, I complement my results with a descriptive analysis of three waves of the Ethiopian Socio-Economic Surveys (ESS) collected in 2011, 2013, and 2015, combined with my district-level covariates. This additional panel of households allows me to observe variation in PSNP beneficiaries within targeted districts, at both the community-level and individual-

[19] In addition to this conceptual discussion, Appendix 2.5 sketches a more formal theoretical framework borrowed from Imbert and Papp (2015) and minimally adapted.

level.[20] I can still control for the targeting rule variables, to account for the selection at the *woreda*-level, but I cannot estimate my difference-in-differences specification as no wave of this survey was collected before the programme started. I therefore estimate two descriptive specifications that mimic my difference-in-differences analysis:

$$Y_{irwkt} = b \times \mathbb{1}_{\{PSNP=1\},w} + \mathbf{X}'_w\theta + \eta_r + \gamma_t + \epsilon_{1,irwkt} \tag{2.2}$$

where Y_{irwkt} is the outcome of interest for individual (or household) i, in the region r, in *woreda* w, in survey round t. $\mathbb{1}_{\{PSNP=1\},w}$ is an indicator equal to one if the *woreda* is targeted by the PSNP, $\mathbf{X}_w$ is a vector of time-invariant controls accounting for the geographic targeting rule of the programme; η_r and γ_t are region and survey-round fixed-effects. Finally, $\epsilon_{1,idt}$ is the unobserved idiosyncratic component, while b remains the main coefficient of interest.

Second, to study how the programme effects on the labour market decisions in untargeted communities inside targeted *woredas*, I extend the specification in Equation 2.2 to include an indicator for whether the *kebele* participated in the PSNP according to the community-level respondent.

$$\begin{aligned} Y_{irwkt} =& b_1 \times \mathbb{1}_{\{PSNP=1\},w} + b_2 \times \mathbb{1}_{\{PSNP=1\},k} + b_3 \times \mathbb{1}_{\{PSNP=1\},i} \\ &+ \mathbf{X}'_w\theta + \eta_r + \gamma_t + \epsilon_{2,irwkt} \end{aligned} \tag{2.3}$$

In Appendix Tables 2.9-2.11, Panel A reports estimates of b from estimating Equation 2.2 across different outcomes using the ESS. Each column presents this partial correlation for different outcome. Whereas Panel B reports estimates of b_1, b_2, b_3 from estimating Equation 2.3. Panel B unpacks whether any differences between targeted

[20] Appendix Figure 2.5 shows districts by their exposure to the PSNP and on whether they were sampled in the Ethiopian Socio-Economic Survey. Appendix 2.7 provides more details on the within-district distribution of PSNP targeting according to the survey data.

and untargeted *woredas* are stronger in targeted or untargeted *kebeles*. Moreover, the difference between b_1 and b_2 is of interest, because it indicates the relative difference of outcomes among untargeted households within targeted woredas, but across different *kebeles*. This difference can be interpreted as a descriptive non-causal spillover effect. The interpretation is not causal, since *kebeles* within targeted districts were selected based on the list of food insecure households compiled by the local administrators. The p-value testing the equality of these two coefficients is reported below Panel B.

The descriptive results are consistent with the main analysis. There is no significant differences in the extensive labor supply or the share of workers engaged in different occupations among non-beneficiaries in targeted and untargeted districts. The sign of the effects is consistent with the main analysis, but the magnitude is smaller.[21] However, there are some differences in the intensive labor supply. There are increases in the hours spent working in self-employment outside of agricultural for individuals in untargeted communities in PSNP districts. While non-beneficiaries in PSNP communities report spending more hours working in farming compared to non-beneficiaries in untargeted communities within the same district.[22] Non-beneficiaries in targeted communities demand more days of unpaid labour and report lower wages compared to both untargeted communities within the same district and individuals in untargeted districts (Column 2 and 3, Appendix Table 2.11). Individuals in untargeted communities in targeted districts report higher wages compared to those in untargeted districts, though this difference is not statistically significant. The descriptive analysis supports the notion that sectoral shifts may not be a direct effect of the programme but may be

[21] In particular, there is a 1.5 percentage point higher proportion of workers in non-agricultural self-employment in untargeted communities in PSNP-districts compared to non-PSNP ones (Column 3, Appendix Table 2.9).

[22] The difference is not statistically significant at the 10% level (p-value: 0.138), but it is non-negligible in magnitude, corresponding to 128 fewer hours (in a whole year) spent farming. Appendix Table 2.10 Panel B shows a 12% negative difference in the hours spent working in farming in untargeted communities in PSNP-districts compared to non-PSNP ones and a 15% positive difference in the hours spent working in agriculture among non-beneficaries in targeted communities.

driven indirectly from untargeted communities in PSNP districts. Overall, the findings align with the patterns identified earlier in the main analysis.

2.3.6 Robustness checks

To validate the results in Table 2.3 (in particular that the PSNP increased non-agricultural self-employment by about 5 percentage points) I employ four strategies: First, I run a placebo test replacing the indicator variable $\mathbf{1}(t = 2013)$ with $\mathbf{1}(t = 2005)$ and using the data from the 1999 and 2005 LFS rounds. Second, I include population density from the 2007 census (interacted with a dummy for the year 2013) as an additional control, although this variable might be considered a "bad control" since it is measured post-implementation. Third, I study heterogeneity of the main effects by whether districts experienced a pre-programme shocks. Finally, I present estimates of the main effects without district-level controls, without weights, and on the unbalanced panel of districts. None of my conclusions are affected by these robustness checks.

Placebo test

The first robustness check (Appendix Table 2.12) examines changes in employment and occupational categories before the PSNP started using data from 1999 and 2005 LFS rounds. The results show no significant changes in labour supply or self-employment activities. However, I do find an increase of 0.9 percentage points in public sector labourers in targeted districts, which is large (relative to the overall sample mean of 0.7%) and statistically significant at the 5% level. This effect may be consistent with the start of the first implementation phase of the PSNP public works by 2005.[23] Nonetheless, I regard 2005 as a baseline pre-programme year because, as the World

[23] A higher share of government employees is also plausibly due to the political and institutional factors related to the historical disbursements of aid in those districts, where sufficient administrative capacity had to be in place to monitor the transfers during times of emergency.

Bank (2010b, pp.1) states, the first phase of the programme (between 2005 and 2006) 'focused on testing and strengthening institutional arrangements and delivery systems', and facilitated the transition from the previous emergency system. Since 2007, the programme was seen to consolidate the changes and operate at a much larger scale. Hence, it is unlikely that within the first few months of the programme there would have been enough participants to strongly attenuate any market-level impacts of the programme by 2013. However, to be precise, my estimates should be seen as the additional effect of the programme relative to its initial adjustment phase. This placebo test is based on a sub-sample of the 391 districts due the challanges matching across the first two LFS rounds, further described in Appendix 2.7.

Adding population density as a control

The reduced availability of land, due to increased population growth, has been identified as one of the factors contributing to the reduction of productive assets in rural Ethiopia (World Bank, 2010a). To account for this dynamic, Appendix Table 2.13 presents the results adding population density as a control, a variable taken from the 2007 census. Before adding population density as a control, I drop observations from the Somali region in Panel A. I remove these observations because the 2007 census did not cover this region. Hence, the estimates of β, after including population density as a control (Panel B), should be compared to Panel A. After removing observations from the Somali region, the coefficients in Panel A are similar to the main results in Table 2.3. However, after including population density as a control in Panel B, the magnitude and significance of the coefficients decrease, particularly in columns 4 and 5. This suggests that the baseline controls were not fully accounting for the influence of population density. While the results remain consistent with the main analysis, this robustness check indicates that the magnitude of the effects may be about 1 percentage point smaller when accounting for changes in population dynamics.

Testing for heterogenous effects due to pre-programme shocks

As the third robustness check, I investigate whether pre-programme shocks affected the labour market outcomes differently in PSNP districts. I add to my previous specification a district-specific variable, W_d, likely correlated with pre-2005 shocks. I use two measures of pre-programme shocks, W_d: standardized cumulative rainfall for the 2002 *Belg* season and a dummy variable indicating continued relief assistance in 2005.[24]

$$
\begin{aligned}
Y_{idt} =& \beta \times (\mathbb{1}_{(PSNP=1)} \times \mathbb{1}_{(t=2013)}) + (\mathbf{C}_d \times \mathbb{1}_{(t=2013)})'\delta + \mathbf{X}'_{dt}\theta + \\
& \mathbf{H}'_i\zeta + \eta_d + \gamma \times \mathbb{1}_{(t=2013)} + \\
& \beta_2 \times (\mathbb{1}_{(PSNP=1)} \times \mathbb{1}_{(t=2013)} \times W_d) + \gamma_2 \times (\mathbb{1}_{(t=2013)} \times W_d) + \epsilon_{3,idt}
\end{aligned} \tag{2.4}
$$

Appendix Table 2.14 presents the estimates of β and β_2 from estimating Equation 2.4. The results indicate that including the interaction terms does not significantly alter the estimates relative to the main results. The estimates of β increase slightly in magnitude relative to the main results. This pattern suggests that pre-programme shocks could have attenuated the effects of the programme on the labour market outcomes considered, rather than bias them upwards. These estimates do not imply a failure of the parallel trends assumptions because of pre-programme shocks.

Removing weights and district-level controls

In Appendix Table 2.15, the results remain similar to the main estimates even when not using weights or when expanding the sample to include all individual observations in

[24] The first measure is based on the observation that districts affected by the widespread 2003 drought generally received limited rainfall during the 2002 *Belg* season (Gill, 2010). The second measure accounts for the possibility that PSNP-targeted districts that required emergency assistance in 2005 may have been more susceptible to experiencing negative shocks before the programme's implementation.

the 601 districts sampled in either the 2005 or 2013 LFS rounds. Although removing weights decreases standard errors, adding more districts does not change the main estimates but improves the precision of the control variables.[25] In Appendix Table 2.16, the effect of the PSNP on the main outcomes remains significant even without additional controls in the basic difference-in-differences model, demonstrating the overall robustness of the results presented in Table 2.3.

2.4 Conclusion

This paper examines the impact of the Productive Safety Net Programme (PSNP) on rural labour markets in Ethiopia using a difference-in-differences approach. The results indicate that the programme did not significantly affect the extensive and intensive labour supply in targeted districts. However, the PSNP led to a higher share of self-employed individuals engaging in non-agricultural activities in these districts. The PSNP primarily serves its main objective of ensuring food security for its beneficiaries rather than increasing overall employment. The programme does not appear to crowd out private sector activities. The results are consistent with the programme's productive assets having improved market access, leading to shifts in non-agricultural self-employment activities. The analysis of equilibrium wages remains limited due to the nature of the Ethiopian rural labor market data. Future research could explore how the rural-urban wage equilibrium may have been affected by both the rural and urban PSNP (which was launched in 2015).

[25] Solon et al. (2015) note that weighting may harm precision if the intra-group (district) correlation makes up a large proportion of the variance of the error term.

Appendix

This appendix has four sections. In Section 2.5 I report a conceptual framework that motivates the analysis of public works on equilibrium wages. Section 2.6 provides additional institutional details about the programme. Section 2.7 provides additional details about the dataset construction and sources of the covariates. Finally, Section 2.8 reports additional analytical checks described in main text.

2.5 Theoretical appendix

The exposition here follows Imbert and Papp (2015). This model illustrates theoretically how changes in public works can affect the labor market equilibrium.

2.5.1 A model of household labour supply and demand

Time is static. Households are indexed by i. D_i denotes household labour demand. Households operate a production function:

$$F_i(D_i) = A_i G(D_i) \tag{2.5}$$

where $A_i \in [\underline{A}, \overline{A}]$ are exogenous productive factors owned by the household. $G'(.) > 0$ and $G''(.) < 0$, i.e. the production function exhibits decreasing marginal returns to scale.

Households choose consumption (c_i), labour supply (L_i^s) and demand (D_i) to solve:

$$\max_{c_i, L_i^s, D_i} u(c_i, T - L_i^s) \quad \text{subject to}$$

$$\begin{aligned} c_i &= y_i + \tilde{W}_i L_i^s \\ y_i &= \pi_i \\ &= A_i G(D_i) - \tilde{W}_i D_i \end{aligned} \tag{2.6}$$

where y_i is non-labour (non-wage) income, π_i is profits from home production, $\tilde{W}_i$ is the shadow wage, which is the price of labour for the household that could be lower than the market wage W. Deriving first order conditions for Equation 2.6, given separability of consumption and production decisions, households will set the marginal product of labour equal to the shadow wage:

$$A_iG'(D^*) = \tilde{W}_i \tag{2.7}$$

2.5.2 Equilibrium with competitive labour markets

Suppose that labour markets are competitive, such that: $\tilde{W}_i = W$, the shadow wage that measures the opportunity cost of time is equal to the market wage for all households. If so, then $A_iG'(D^*) = W$.

If A_i is low, then $G'(D^*)$ will be high and because of $G''(.) < 0$ then D^* will be low for low-productivity households. In particular, low productivity households will be net-sellers of labour $D^* < L_i^{*s}$. Conversely, if A_i is high, then $G'(D^*)$ will be low and because of $G''(.) < 0$ then D^* will be high for high-productivity households. In particular, high productivity households will be net-buyers of labour $D^* > L_i^{*s}$.

2.5.3 Equilibrium with frictions

Suppose that due to labour market frictions (e.g. job search costs), there is a wedge $p \in [0,1]$ between the returns to one unit of labour for workers (pW) and its costs for employers (W). High-productivity households that are net-buyers of labour will then price the marginal value of labour according to $A_iG'(D^*) = W$. Low-productivity households that are net-sellers of labour will then price the marginal value of labour according to $A_iG'(D^*) = pW$. Households with intermediate productivity levels do not participate in the market and set $A_iG'(D^*) \in [pW, W]$. Denote with $\phi(W)$ the value of the productivity factor A_i such that labour supply is equal to labour demand for household i.

2.5.4 The effect of public works on labour market equilibrium

Suppose the government starts hiring labour at a wage W_g. Total labour hired in public works is $L^g = \int_i L_i^g di$ Then the households' non-labour income earned outside of labour markets (i.e. in own-farm agriculture or through public works) is:

$$y_i = \pi_i + (W_g - \tilde{W}_i)L_i^g \tag{2.8}$$

Define the labour market clearing condition that sets the total labour supply by net-sellers of labour equal to the total labour demanded by net-buyers of labour:

$$\underbrace{p\int_{\underline{A}}^{\phi(pW)} [L_i^s(pW) - D_i(pW) - L_i^g]dA_i}_{\text{low-productivity suppliers}} = \underbrace{\int_{\phi(W)}^{\overline{A}} [D_i(W) - L_i^s(W) + L_i^g]dA_i}_{\text{high-productivity buyers}} \tag{2.9}$$

To understand the equilibrium effects of public works, we would want to totally differentiate Equation (2.9) with respect to L_g. Note that we implicitly define the market wage W to be a function of L_g.

After applying Leibniz rule to differentiate an integral and several steps to simplify the algebra, we can define the total effect on market wage of public works as follows:

$$\frac{dW}{dL^g} = \frac{E_1 - E_2}{-E_3 + E_4} \tag{2.10}$$

where:

$$E_1 = p\int_{\underline{A}}^{\phi(pW)} \frac{dL_i^g}{dL^g}dA_i + \int_{\phi(W)}^{\overline{A}} \frac{dL_i^g}{dL^g}dA_i > 0 \tag{2.11}$$

E_1 is the crowding out of public employment from other sources of employment. The other sources of employment are wage labour, for the least productive households, and self-employment, for the more productive households.

$$E_2 = p \int_{\underline{A}}^{\phi(pW)} \frac{dL_i^s}{dy_i}(W_g - pW)\frac{dL_i^g}{dL^g} dA_i + \int_{\phi(W)}^{\overline{A}} \frac{dL_i^s}{dy_i}(W_g - W)\frac{dL_i^g}{dL^g} dA_i \tag{2.12}$$

E_2 is the effect on aggregate labour supply through non-labour income. This effect can be interpreted as the change in the total labour supply occurring from individuals shifting out of the wage market since they are getting an income directly through public works. $E_2 < 0$ if:

(i). $\frac{dL_i^s}{dy_i} < 0$ because of an income effect.

(ii). $(W_g - W) > 0$ by assumption of the programme, but this assumption is not valid in the PSNP case, where $(W_g - pW) \leq 0$.

Hence, the numerator of Equation (2.10) is generally positive, so long as (i). and (ii). are true.

If $E_1 > 0$ and $E_2 \geq 0$, then $E_1 - E_2$ can be ambiguous. In particular, if the income effect is larger than the crowding-out effect, i.e. $E_2 > E_1$, public works may reduce employment. Otherwise, if the income effect is small, then $E_1 > E_2$ will still make the numerator of Equation (2.10) positive, but smaller relative to the case where $E_2 < 0$. The latter scenario may occur if the public works wage is set to be below or equal to the market wage.

$$E_3 = p^2 \int_{\underline{A}}^{\phi(pW)} D'(pW) dA_i + \int_{\phi(W)}^{\overline{A}} D'(W) dA_i \tag{2.13}$$

E_3 is the effect on aggregate labour demand, which will generally be negative, based on the slope of the demand curve.

$$E_4 = p^2 \int_{\underline{A}}^{\phi(pW)} [\frac{dL_i^s}{dW}|_u + \frac{dL_i^s}{dy_i}(L_i^s - D_i - L_i^g)]dA_i + \int_{\phi(W)}^{\overline{A}} [\frac{dL_i^s}{dW}|_u + \frac{dL_i^s}{dy_i}(L_i^s - D_i - L_i^g)]dA_i \tag{2.14}$$

E_4 is the effect on aggregate labour supply via the wage equilibrium changes. If leisure is not a luxury good (which you consume a higher share of as you get richer), then $E_4 > 0$, which makes the denominator of Equation (2.10) also positive.

The effect on the wage of an increase in public works will be larger if $-E_3$ is small (i.e. aggregate demand is inelastic to the wage), or if E_4 is small because the labour supply is inelastic to the wage.

2.6 Programme details

2.6.1 Weather shocks and safety nets

Despite being one of Africa's fastest growing economies, Ethiopia's poverty rate remains high. While poverty reduction is one of the main objectives of the Ethiopian government, the number of individuals consuming less than US$1.25 per day (in purchasing power parity terms) was estimated to be 29.6% in 2010/11 (GFDRE, 2013). Food security is an unavoidable policy concern that Ethiopia has to address in pursuing poverty reduction.

To counter seasonal food shortages, Ethiopia has been receiving relief food aid from abroad, with amounts varying from year to year over the last 30 years. Until the establishment of the PSNP in 2005, the government resorted to annual appeals to the international community in order to secure assistance.[26] The emergency response system in place prior to 2005 had saved many lives, but was seen as not having protected the livelihoods of those affected by shocks (Kehler, 2004).

Following the 2003 drought, the GFDRE and a consortium of Development Partners[27] developed a Food Security Programme that aimed to overhaul the relief aid system, turning it into a more reliable safety net. The programme developers anticipated that the new system would allow both recipients and donors to plan support ahead of emergencies, rather than organising relief responses on a nearly annual ad-hoc basis. In particular, they argued that the provision of transfers over multiple years would

[26] The annual appeal system was considered unreliable, because food deliveries were often untimely and irregular, and unsustainable, because of instability in the global food marketing regime and uncertainties regarding donor pledges following the appeals (Rahmato, 2013). It could have taken up to three months after the outbreak of a food crisis for relief to reach those in need.

[27] The Development Partners comprise multilateral agencies such as the World Bank, the World Food Program, the European Union, and bilateral partners, such as USAID, the UK Department for International Development, Irish Aid, the Canadian International Development Agency and the Swedish International Development Agency.

allow recipients to curb the depletion of their own assets in times of need.[28] The PSNP was allocated the lion's share of the Food Security Programme's budget, and is the flagship component of this new strategy to counter food insecurity.[29] Drawing from existing studies and reports, I next provide an overview of how the programme is designed.

2.6.2 Overview of the PSNP

The PSNP aims to alleviate the incidence of food insecurity and avoid asset depletion among historically vulnerable rural communities. It primarily seeks to achieve this through timely and appropriate food and/or cash transfers, and the creation of productive community assets that can contribute to environmental rehabilitation, increase household productivity, and improve access to infrastructure and services (GFDRE, 2006).

The programme is managed by the GFDRE, but remains mostly donor-funded.[30] It has grown significantly in terms of budget requirements as the number of targeted beneficiaries has expanded. The fourth and latest phase of the programme, running from 2015 until 2019, has a budget requirement of US$3.6 billion, towards which the GFDRE has committed US$500 million, with the remainder financed by its Development Partners (World Bank, 2014).

After the first year, which was intended to test the administrative and logistic capacity to deal with the deployment of such a large programme, the number of districts went

[28] Short-selling of livestock in bearish market conditions is an example of a short-term coping mechanisms taken by households during food shortages. However, this practice may only contribute to less than a third of income smoothing after a drought (Fafchamps et al., 1998). Another short-term coping strategy is the deforestation of hill-sides for the production of charcoal. The PSNP seems to have had a modest positive impact on forest stock (Andersson et al., 2011), reducing environmental degradation of the agroecological conditions.

[29] The other components of the Food Security Programme were complementary to the PSNP, and were implemented in some, but not all, of the districts where the PSNP operated.

[30] The World Food Programme covers implementation in the Somali region.

up to 262. However, the increase in the number of districts was mostly due to large districts splitting, shortly after the 2005 elections. These administrative splits were partly justified on the grounds that large *woredas* were harder to administer and lacked sufficient governance.[31] Hence, the actual number of targeted districts, relative to the 2005 administrative boundaries, had not actually increased by 2006.

2.6.3 PSNP beneficiaries

The demographic characteristics of beneficiaries are relevant in choosing the appropriate labour market to focus on, and potential control variables for the analysis. The main beneficiaries of the PSNP transfers are chronically food insecure households, which the Programme Implementation Manual (PIM) defines as 'households that have been unable to meet their food needs for a period of three months or more in the last three years' (GFDRE, 2006, pp.4). In addition to chronically insecure households, the programme aims to provide transfers to households that are temporarily unable to meet their minimum food consumption requirements due to a negative shock, and households that have no means of support, such as remittances.

Eligible beneficiaries, who are able-bodied and above 16 years of age, receive transfers in return for participation in public works. In 2009, transfers conditional on public works participation comprised 84% of the total transfer to beneficiaries (World Bank, 2010b). Other eligible households, who cannot supply labour (either temporarily or permanently), receive an unconditional transfer (referred to as Direct support). Direct support beneficiaries include, but are not limited to, orphans, pregnant and breast-feeding women, the elderly, people with disabilities, and female-headed households with young children (GFDRE, 2006).

[31] I refer to the conversation I had with World Bank Officials in January 2016.

2.6.4 Public works

The main feature of the PSNP operations is its public works component. The public works supported under the PSNP are small-scale, labour-intensive community projects designed to provide unskilled, temporary employment for eligible households with able-bodied members. For all sub-projects in a district, the ratio of total labour inputs to total costs should be at least 80% (GFDRE, 2010). Annually around 46,000 public works sub-projects are undertaken (World Bank, 2009). The nature of the projects vary depending on the local environmental conditions and community needs. Most projects involve soil and water conservation activities aimed at fostering the local watershed development. Other PSNP-funded projects involve the construction of local roads, schools or health posts. The potential productivity effects of the infrastructure generated by these projects is what motivated the first "P" in the programme's acronym. These productivity gains can plausibly be the factor driving changes in the local labour market. However, because of a lack of a spatial database for public works program activities, it has been hard to accurately evaluate their impact (Subbarao et al., 2013).

The timing of public works is key. Public works run for 6 months each year, usually from January to June, to coincide with the agricultural slack season. The project's timing aims not to interfere with agricultural labour needs.[32] Participants usually work for eight hours a day for around 5 days/month. The actual days of individual employment vary depending on the household circumstances, as able-bodied members are expected to fulfil the workfare requirements (up to a maximum of 15 days/month) other household members that also receive transfers, but who do not participate in public works. The individual cap of 15 days/month was implemented for two reasons: budgetary constraints; and to enable participants to have sufficient time to engage in

[32] One may worry that the public works were not operating at the time in which the survey used in the analysis were collected. Luckily, the surveys were collected in March and June. I further elaborate on this point when discussing the potential limitations of my dataset.

other productive activities outside of the programme. As such, the programme was designed in a way that would not distort the intensive and extensive margin of the labour supply of participants.

In 2009, the World Bank estimated that the PSNP provided 190 million days of public works employment to 1.27 million households (World Bank, 2009). An additional 242,000 households were estimated to be direct support beneficiaries. The average household employed in public works received 129 days of employment in 2009, with some variation in this average across regions (Berhane et al., 2011). Administrative data on individual participation to the PSNP has been hard to find, even for the authors involved in the official impact evaluation of the programme (*Ibid* pp.131). As such, aside from the estimates of the independent evaluation and the official statistics, I am unable to observe directly whether individuals have taken up participation in the programme.[33]

2.6.5 Cash and food transfers

PSNP beneficiaries are remunerated with a daily payment in either cash or food, depending on their location. Overall, 60% of transfers are provided in cash, with factors such as local market conditions, beneficiaries' preferences and logistical constraints influencing which of the two is used.

The cash wage was meant to enable households to purchase the equivalent food transfers from the local market.[34] By design, this level is below the usual market wage

[33] This is a limitation of my study, if one worries about the potential institutional malfunctioning that could hinder the implementation of the programme. However, the high degree of scrutiny from the Development Partners, along with the fact that the evaluations of the programme were independent of the government, should provide some reassurance that the programme was operating.

[34] Food transfers are in general 3kg of cereals per day worked. In 2008, the rate was first increased from 6 to 8 birr/day to take into account the soar in food prices, with subsequent raises following roughly every 2 years. Until 2011, a uniform wage rate was employed across all recipient *woredas*, but, in 2012, it was decided to allow districts to change the wage rate so as to take into account the geographic heterogeneity in food availability and prices. In 2015, US$1 was exchanged for approximately 20 Ethiopian Birr (ETB).

for unskilled labourers (Subbarao et al., 2013). Currently, the wage rate is on average ETB23/day of work across all regions receiving cash transfers. In 2009, the estimated value of (annual) wages earned per average household recipient was US$137 (World Bank, 2009).

The parity of cash and food transfers has eroded over the years, with food becoming more expensive and cash transfers not adjusting fast enough. This disparity was particularly accentuated during the food price spike in 2008-2009, but the share of cash transfers never went below 50%. Economic theory suggests that if the public works wage is set above the market wage, then private labour supply may be crowded-out by public employment, raising the equilibrium wage for workers in the private sectors (Ravallion, 1991). The erosion in purchasing power of the wages offered by the programme, coupled with the fact that rates were intentionally set below market wages, could potentially reduce any aggregate effect of the programme occurring through changes in the demand for labour.[35]

[35] The reasoning is analogous to the introduction of a minimum wage above the market wage in a competitive market, which results in a higher equilibrium wage and lower employment.

2.7 Data appendix

2.7.1 Constructing a panel of districts

While the CSA made a big effort to cover both rural and urban areas in all regions of the country, its objective was not to cover all districts. There are only a few zones (and the districts within them) that are systematically omitted from the sampling frame. Appendix Figure 2.3 shows how the 2005 and 2013 round differ in their coverage of districts and what that means for the size of my balanced panel of districts.

I have to drop observations from the Gambella region and most of the Afar region, which make up about 0.7% and 1.5% of the total rural population of all districts sampled, respectively.[36] This is because rural districts in these regions were not included in the sampling frame in 2005. Aside from these cases, the sampling method was similar across survey rounds. Hence, the reason why a given district is not sampled in a round is (presumably) due to the realisation of the random draw of districts from the same population that were chosen to be sampled, except for those zones that were ex-ante excluded from the sampling frame. I do not expect there to be a bias in my estimates due to sample selection because of the survey design.

To merge the datasets, I follow this procedure: First, I construct a district identifier for the 2013 round of the LFS, which I match with the 2007 census. To create unique district identifiers across districts, I concatenate three numbers: an integer for the region, an integer defining the zone within a particular region, and an integer for the district within a particular zone. The CSA, which also carries out the census, did not change its maps since the 2007 census, so district identifiers are consistent between the 2013 LFS round and the 2007 census. This is how I obtain a list of district names in the 2013 LFS round, which was missing and is crucial for what follows next.

[36] Population estimates are calculated from the 2007 census.

Second, I digitalize the 2005 LFS district geographic identifiers, which were only available as a scanned file. As noted in the identification section, many new districts were formed following the 2005 election, by splitting large districts into two or more new ones. About 200 new districts were formed between 2005 and 2006. There are only a few instances in which two (pre-2006) districts were divided to jointly form a new district; I treat these few cases as if the new district was formed from a part of either of the two old ones. My challenge consisted in finding out which districts had split, and then assigning to each old district an identifier that was consistent with the 2013 round. I used the district names to identify which districts had split, using the information from two sources: recent administrative maps of Ethiopia[37], and the map plotting years of assistance, which was originally drawn using pre-2007 boundaries (before I converted it to post-2007 boundaries). Google searches were also used to confirm the validity of the district splits I identified.

After identifying which districts had split, I could have grouped the district boundaries in the 2013 round to reflect the old borders, aggregating back the new districts into their old borders. However, this procedure would have not taken into account the fact that the PSNP operates only within certain villages in each district, and not all newly formed districts that were originally contained in a geographically targeted district were targeted by the programme after 2006. As noted in the background section, the district officials were supposed to roll out the PSNP in the most needy villages based on the reports of the community food-security task force, which had drawn a list of food insecure households. Priority was given to villages with the highest number of food insecure households. There was no official cut-off that determined roll-out at the village-level. As such, the newly formed districts were not necessarily targeted by the programme following the boundary changes. Matching the old district to only one of the newly formed *woredas* would have incorrectly assigned treatment to certain

[37] Available at `http://tinyurl.com/ocha-map13`, accessed on 09/04/2016.

districts, which were not in fact recipients of the PSNP. Thus, I follow the approach suggested by Imbert and Papp (2015).

Using the 2005 LFS round, for I duplicate observations in districts that split into x copies, where x is the number of newly formed districts (usually two or three). Then, I assign a 2013 district identifier to each individual in a given copy of the x newly created districts. Finally, to adjust the sample for these artificial copies, I divide the survey weights by x for the observations that were duplicated x times. I apply the same procedure to the matched observations in the 1999 LFS round, which I use for my placebo test.

Issues combining the 1999 LFS round

Between 1999 and 2005, certain zones changed boundaries, and so did the integers that identify them. Unfortunately, the 1999 LFS round did not have district names like the 2013 round. To match this round with the 2005 round, I have to assume that the district numeric identifiers have remained constant across the two rounds. For the most part, it is unlikely that numeric identifiers changed between the two rounds for two reasons: First, the majority of districts splits in the last two decades occurred after the 2005 elections. Further, the CSA relies on census maps to assign geographical identifiers for most of its surveys, and there was no census collected between 1994 and 2007. However, in 2000, rather than districts splitting, some zones were divided.[38] I lack the information to unambiguously match the unique district identifiers across time and rounds in the zones that changed boundaries between the 1999 round and the 2005 round. Hence, for the placebo test, I have to drop the unmatched districts from the analysis, which makes up 10% of the observations collected in 1999. This restricts my balanced panel of districts for the placebo test to 391 *woredas*.

[38] Zones are the intermediate administrative unit between regions and districts, usually containing 5-10 *woredas*.

2.7.2 Sources of other covariates

Geographic targeting data

The geographic assignment of the PSNP mostly comes from the only two publicly available lists published in the Programme Implementation Manuals (GFDRE, 2006, 2010). I also compared the list of districts names with the maps contained in the World Bank (2010b) results report, by plotting the GFDRE's lists onto administrative shapefiles. With this procedure, I ensure that I match the geographic targeting of 290 districts by the end of 2009. The World Bank acts as the coordinator for all donor partners involved in the programme, which is why I rely on the information they publish.

Historical frequency of food aid

Districts were targeted based on their historical receipt of food aid prior to 2005. I collected data on the frequency of historical relief assistance at the district-level (between 1994 and 2005) from the National Disaster Risk Management Committee[39] of the GFDRE. I only observe an indicator for whether a district received aid assistance in a particular year, and not the quantity of aid received by a district in each of these years. I personally collect this data in a trip to Addis Ababa in January 2016. This information is shown in Appendix Figure 2.4. Its inclusion should capture some of the unobserved characteristics that are shared by targeted districts, such as the level of food insecurity, which, if omitted, could bias the estimated effect of the programme.

[39] Formerly known as Disaster Prevention and Preparedness Committee (DPPC). I am grateful to Lemlem Abraha and Negussie Kefeni for sharing their time in assisting me during such a demanding period.

Weather controls

Weather shocks could be part of the unobserved time-varying component, and may be more frequent in PSNP *woredas*, which is why I control for climatic variables in my main specifications using gridded data sources. Gridded data, which interpolates readings from weather stations with a statistical model, are frequently used by economists.[40] However, one of the difficulties of employing these data sources in low- and middle-income countries, particularly for rainfall, is that the stations tend to be highly dispersed, increasing the potential for measurement error. For this reason, I use data from the Global Precipitation Climatology Centre (GPCC) dataset as its station coverage has been found to be better than any other publicly available source of monthly rainfall (Becker et al., 2013). The GPCC dataset is maintained by the World Meteorological Organization and contains monthly estimates of total precipitation (mm) for the global land surface at $0.5° \times 0.5°$ resolution for all years between 1900 and 2014. For temperature, I employ the most recent version (V4.01) of the well-known Willmott and Matsuura (2015) series hosted by the University of Delaware, providing monthly temperatures at the same spatial resolution, for the period of interest. These data have been used in several other studies, such as Adhvaryu et al. (2019) and Theisen (2012), and were chosen because of their geographic scope and long time scale.[41] Since the gridded climatological data does not necessarily match the administrative district boundaries, a precipitation/temperature value is assigned to each *woreda* based on the values of the raster cells covering that *woreda*. If one single cell covers the *woreda* in question, then the *woreda* takes on the value of that cell. When two or more cells cover a single *woreda*, a weighted mean is calculated, where the weights are equal to the

[40] See Dell et al. (2014) for a review of the recent economic literature using weather data.

[41] I use data between 1979 and 2014 to construct a sample mean and standard deviation with which I calculate standardized values of cumulative rainfall and average temperature, for each year and each cropping season.

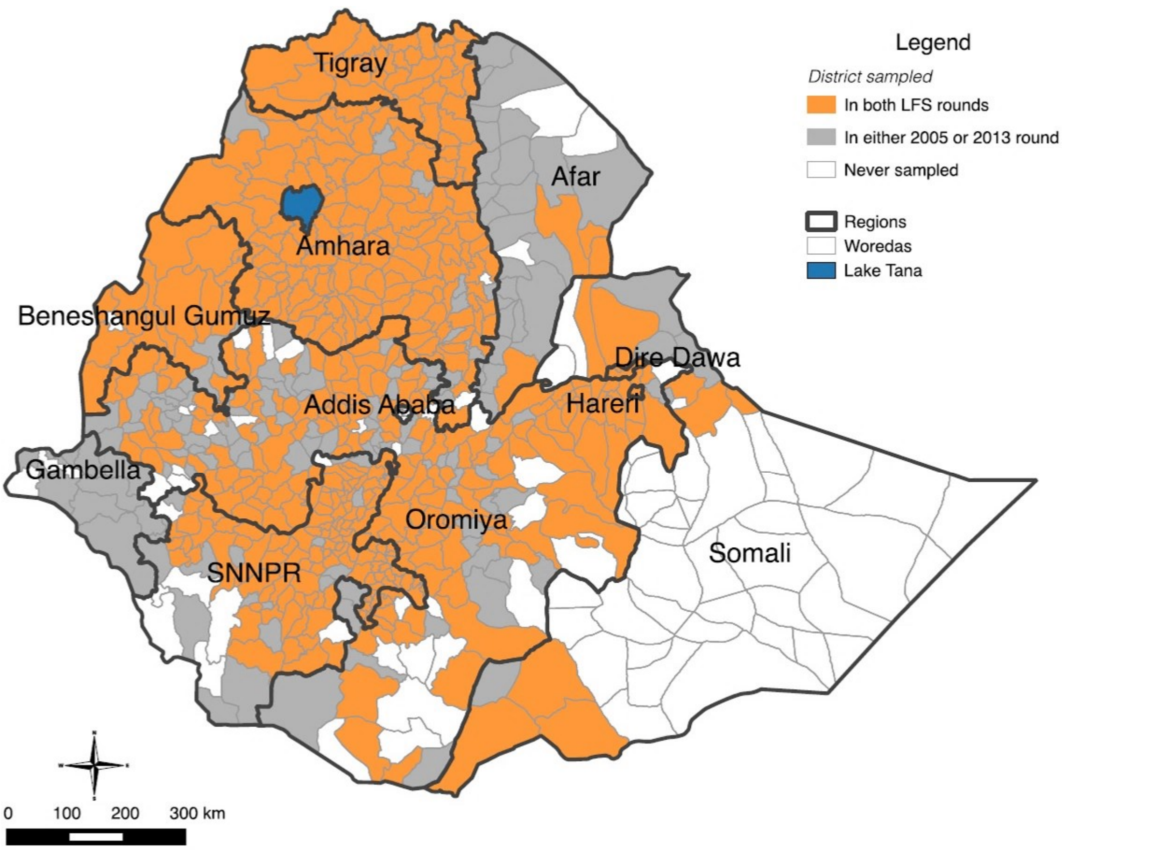


Figure 2.3: District balance in the Labour Force Survey

fraction of the polygon covered by each cell.[42]

Other controls, which I do not include in the main regressions (but that are shown in Appendix Table 2.7) come from the village-level 2007 census of Ethiopia, also carried out by the CSA. These variables could constitute a bad control, as they may have been affected by the PSNP between 2005 and 2007. Hence, I only include additional census variables controls as a robustness check, to explore whether my results could be explained by changes in the population dynamics.

[42] Temperature and rainfall data used are freely available at http://tinyurl.com/udel2014 and http://tinyurl.com/gpcc2014, respectively, accessed on 20/04/2016.

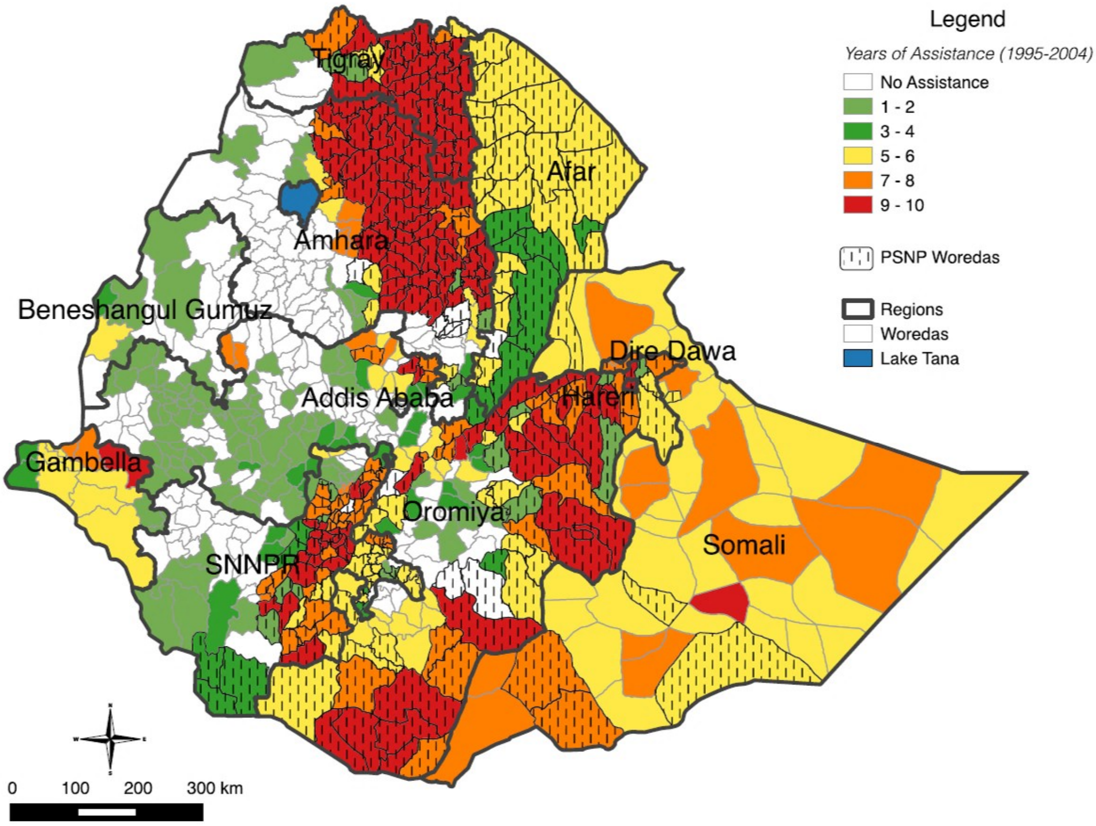


Figure 2.4: Cumulative years of aid receipts and PSNP targeting[a]

[a] *Notes:* PSNP assignment of 290 *woredas*, as of the end of PSNP Phase II (2007-2009). Years of assistance collected by the author from the National Disaster Risk Management Committee.

2.7.3 PSNP targeting in the Ethiopian Socio-Economic Survey

Table 2.6: PSNP targeting in the Ethiopian Socio-Economic Survey

	Mean	SD	Min	p25	p50	p75	Max
1 if PSNP-*woreda*	0.51	0.50	0.00	0.00	1.00	1.00	1.00
Years of assistance received between 1994-2005	4.94	3.78	0.00	1.00	6.00	9.00	10.00
1 if *woreda* received assistance in 2003	0.64	0.48	0.00	0.00	1.00	1.00	1.00
1 if *woreda* received assistance in 2004	0.56	0.50	0.00	0.00	1.00	1.00	1.00
1 if *woreda* received assistance in 2005	0.39	0.49	0.00	0.00	0.00	1.00	1.00
1 if PSNP-*kebele*	0.42	0.49	0.00	0.00	0.00	1.00	1.00
Number of kebeles	277						
Number of woredas	228						
Conditional on the woreda being targeted by the PSNP:							
Years of assistance received between 1994-2005	7.72	2.30	0.00	7.00	9.00	9.00	10.00
1 if *woreda* received assistance in 2003	0.94	0.23	0.00	1.00	1.00	1.00	1.00
1 if *woreda* received assistance in 2004	0.88	0.33	0.00	1.00	1.00	1.00	1.00
1 if *woreda* received assistance in 2005	0.60	0.49	0.00	0.00	1.00	1.00	1.00
1 if PSNP-*kebele*	0.77	0.42	0.00	1.00	1.00	1.00	1.00
Number of kebeles	140						
Number of woredas	112						

The first row of Appendix Table 2.6 shows that about 51% of the *kebeles* in this dataset are located in a *woreda* that was targeted by the PSNP. On average, these *kebeles* were in *woredas* that received about five years of aid assistance in the ten years prior to the start of the PSNP. The next three rows shows that there is some heterogeneity in the distribution of whether a *kebele* was in *woreda* that received aid assistance in the three years prior to the start of the PSNP. In the lower panel of the table, the same variables are conditioned on the *woreda* having been targeted by the PSNP. Importantly, the last

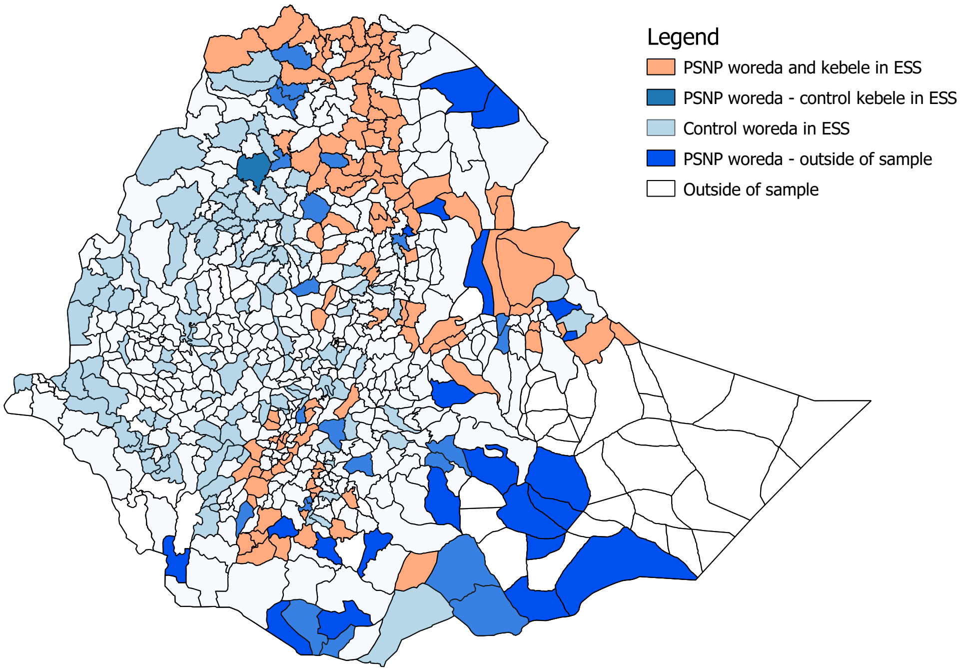


Figure 2.5: PSNP targeting in the Ethiopian Socio-Economic Survey.[a]

[a] Woredas with a salmon shading are those where I also observe that sampled *kebeles* were exposed to the PSNP, whereas the light blue *woredas* were not targeted by the PSNP and were sampled.

row shows that 77% of the *kebeles* inside a *woreda* targeted by the PSNP the dataset had also received the programme, which leaves 23% of *kebeles* as a comparison group to described within-districts differences in the variables of interest.

Appendix Figure 2.5 shows *woredas* by their exposure to the PSNP and on whether they were sampled in the Ethiopian Socio-Economic Survey. Woredas with a salmon shading are those where I also observe that sampled *kebeles* were exposed to the PSNP, whereas the light blue *woredas* were not targeted by the PSNP and were sampled.

2.8 Appendix tables

Table 2.7: Summary statistics on additional district covariates

Additional district-level controls and descriptive statistics	PSNP (1)	Control (2)	p-value (3)	Source (4)	Time-Varying? (5)
Fraction Orthodox	0.45	0.62	0.000	1999 LFS	No
Fraction Muslim	0.39	0.27	0.006	1999 LFS	No
Fraction Protestants	0.13	0.09	0.138	1999 LFS	No
Fraction in Other Religions	0.03	0.03	0.907	1999 LFS	No
Fraction Amhara	0.36	0.36	0.973	1999 LFS	No
Fraction Tigryina	0.15	0.01	0.000	1999 LFS	No
Fraction Somali	0.06	0.01	0.023	1999 LFS	No
Fraction Afari	0.00	0.00	0.358	1999 LFS	No
Fraction Oromo	0.28	0.43	0.002	1999 LFS	No
Fraction of other ethnicity	0.15	0.18	0.398	1999 LFS	No
Fraction of households with a death last year	0.06	0.05	0.001	2007 Census	No
Fraction of households with electrcity	0.03	0.02	0.425	2007 Census	No
Fraction of households with a private toilet	0.21	0.19	0.303	2007 Census	No
Fraction of households with a private kitchen	0.42	0.46	0.005	2007 Census	No
Population density (per sq. km)	250	167	0.000	2007 Census	No
Area (sq. km)	1097.34	1099.37	0.982	2007 Census	No
1979-2014 average cumulative Belg season rainfall (mm)	194.43	175.01	0.016	GPCC	No
1979-2014 average cumulative Meher season rainfall (mm)	581.59	816.37	0.000	GPCC	No
1979-2014 average Meher season temperature (°C)	19.44	17.85	0.000	UDel_AirT	No
1979-2014 average Belg season temperature (°C)	20.19	19.79	0.183	UDel_AirT	No
District Observations	215	238			
Individual Observations	31574	26805			

Notes: This table presents means of the district-level controls used in the additional regression models for different samples. Column 1 includes controls for districts that were targeted by the PSNP. Column 2 includes controls for districts that were not targeted by the PSNP (which form the control group). Column 3 presents the p-values of the student's t-test of equality of means. Standard errors for the student's t-test of equality of means are computed assuming correlation of individual observations within each district in a given year. The additional LFS controls are computed using the 1999 Labour Force Survey, with sampling weights adjusted for boundary changes. The sample is restricted to individuals of ages between 17-65, using information from the usual activity reported. Ethnicity and religion questions were not asked in the 2005 and 2013 round. Census controls are calculated aggregating the village-level 2007 census data. Cumulative rainfall is the 1979-2014 mean cumulative rainfall during the rain seasons for the *Meher* harvest (June-October) and *Belg* harvest season (February-May). Temperature is calculated as the 1979-2014 monthly averages for the respective pre-harvest rainy season.

Table 2.8: Effects on employment participation and sectoral composition by sex

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Employment and occupation effects on women							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	0.465	1.046	-1.446	-8.729**	8.697**	-0.221	0.065
	(3.738)	(1.006)	(3.512)	(3.973)	(3.872)	(0.470)	(0.200)
Mean Dep. Var.	75.46	2.144	22.37	79.63	17.80	0.584	0.416
Observations	54,770	54,770	54,770	40,792	40,792	40,792	40,792
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel B. Employment and occupation effects on men							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	-1.665	0.746*	0.902	-2.618	2.433	0.227	0.381
	(1.236)	(0.450)	(1.014)	(2.033)	(1.648)	(0.589)	(0.244)
Mean Dep. Var.	82.38	1.727	15.88	83.70	11.81	1.234	0.706
Observations	50,553	50,553	50,553	45,976	45,976	45,976	45,976
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each cell reports an estimate of beta for different dependent variables; standard errors in parenthesis are clustered at the district level. Each column has a different dependent variable.
In Panel A, the sample is restricted to women. In Panel B, the sample is restricted to men. The sample consists of individuals aged 17-65, pooling data from the 2005 and 2013 LFS rounds, in the 453 districts sampled in both rounds. Columns (4)-(7) restrict the sample only to those that are currently employed. Individual observations are weighted by sampling weights that are proportional to district population. All models are estimated using ordinary least squares. The means of district-level and individual-level controls are shown in Table 2.1. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table 2.9: Within-district analysis of labour supply (extensive margin)

	=1 if employed	=1 if self-employed farmer	=1 if non-farming self-employed	=1 if employee	=1 if temporary worker
Panel A: Differences across woredas					
1 if PSNP-*woreda*	-0.010	-0.024	0.014	0.004	0.011
	(0.037)	(0.038)	(0.023)	(0.010)	(0.009)
Panel B: Difference across kebeles and woredas					
1 if PSNP-*woreda*	-0.028	-0.042	0.015	0.008	0.018
	(0.039)	(0.038)	(0.025)	(0.012)	(0.013)
1 if PSNP-*kebele*	0.028	0.027	0.001	-0.007	-0.014
	(0.026)	(0.027)	(0.017)	(0.014)	(0.014)
1 if household participated in PSNP	-0.076***	-0.075***	-0.010	-0.015*	-0.011
	(0.026)	(0.026)	(0.015)	(0.009)	(0.008)
p-value : *kebele* vs. *woreda*	.274	.186	.698	.505	.201
Unit of obs.	Individual	Individual	Individual	Individual	Individual
# Clusters	228	228	228	228	228
# Obs.	37980	37980	37980	37980	37980
Dep. Var. Mean	.58	.49	.13	.04	.03
Dep. Var. St. Dev.	.49	.5	.33	.19	.17

Notes: Linear probability estimates of the difference in areas targeted by the Productive Safety Net Programme (PSNP). Each panel presents a separate regression model. Outcome variables are listed on top. The unit of observation is the individual. *Woredas* are districts and *kebeles* are wards within them. Pooled 2011,2013, and 2015 rounds of the Ethiopian Socio-Economic Surveys. All models control for survey round indicators, region indicators, the number of years of aid assistance received by the *woreda* prior to 2005, three indicators for whether the *woreda* received aid in 2004. Standard errors are in parentheses and are clustered at the *woreda*-level. Significance levels: *10%, **5%, and ***1%. "p-value : kebele vs. woreda" reports the p-value for a test of equality between the coefficients in the first and second row of panel B. The bottom panel displays the outcome mean, standard deviation, and total number of observations and clusters.

Table 2.10: Within-district analysis of labour supply (intensive margin)

Panel A: Differences across woredas	Hours worked as self-employed farmer	Hours worked as non-farming self-employed	Hours worked as temporary worker	Hours worked as employee
1 if PSNP-*woreda*	-12.449	17.432	1.651	23.816
	(71.997)	(33.394)	(5.529)	(22.837)
Panel B: Difference across kebeles and woredas				
1 if PSNP-*woreda*	-56.720	22.756	-1.458	28.546
	(72.313)	(34.497)	(5.916)	(29.226)
1 if PSNP-*kebele*	72.211**	-11.657	5.474	-7.646
	(36.584)	(17.720)	(4.264)	(12.755)
1 if household participated in PSNP	-87.386**	5.705	-9.104*	-8.154
	(40.325)	(14.009)	(5.526)	(7.500)
p-value : *kebele* vs. *woreda*	.138	.423	.404	.382
Unit of obs.	Individual	Individual	Individual	Individual
# Clusters	228	228	228	228
# Obs.	37736	37736	37736	37736
Dep. Var. Mean	455.18	120.42	23.64	34.11
Dep. Var. St. Dev.	675.22	428.2	180.42	262.63

Notes: Ordinary least squares estimates of the difference in areas targeted by the Productive Safety Net Programme (PSNP) in hours worked across different activities (annualised from a weekly recall). Each panel presents a separate regression model. Outcome variables are listed on top. The unit of observation is the individual. *Woredas* are districts and *kebeles* are wards within them. Pooled 2011,2013, and 2015 rounds of the Ethiopian Socio-Economic Surveys. All models control for survey round indicators, region indicators, the number of years of aid assistance received by the *woreda* prior to 2005, three indicators for whether the *woreda* received aid in 2004. Standard errors are in parentheses and are clustered at the *woreda*-level. Significance levels: *10%, **5%, and ***1%. "p-value : kebele vs. *kebele woreda*" reports the p-value for a test of equality between the coefficients in the first and second row of panel B. The bottom panel displays the outcome mean, standard deviation, and total number of observations and clusters.

Table 2.11: Within-district analysis of labour demand

	Days of hired labour post-harvest	Days of unpaid labour post-harvest	Daily wages for hired labourers post-harvest	Days of hired labour planting	Days of unpaid labour planting	Daily wages for hired labourers planting
Panel A: Differences across woredas						
1 if PSNP-*woreda*	-1.293	3.963	18.900	-20.896	-1.357	-58.234
	(5.061)	(5.060)	(65.268)	(17.486)	(2.881)	(36.112)
Panel B: Difference across kebeles and woredas						
1 if PSNP-*woreda*	-0.147	-2.371	56.651	-10.696	-3.349	-28.904
	(5.088)	(4.241)	(81.092)	(14.330)	(3.207)	(39.495)
1 if PSNP-*kebele*	-1.766	10.482**	-94.967*	-20.568	4.492	-64.398*
	(2.121)	(4.882)	(50.336)	(15.115)	(3.371)	(36.444)
1 if household participated in PSNP	-2.329*	-5.944*	-49.469**	-12.860	0.398	-15.915
	(1.347)	(3.253)	(20.412)	(11.966)	(2.984)	(52.772)
p-value : *kebele* vs. *woreda*	.773	.048	.214	.576	.168	.568
Unit of obs.	Household	Household	Household	Household	Household	Household
# Clusters	225	225	182	228	228	194
# Obs.	7903	7903	1943	8859	8859	1716
Dep. Var. Mean	11.51	14.64	122.41	27.26	11.65	116.93
Dep. Var. St. Dev.	133.53	93.58	400.12	450.71	68.12	291.6

Notes: Ordinary least squares estimates of the difference in areas targeted by the Productive Safety Net Programme (PSNP) in days of labour (hired or unpaid) and wages paid, before and after harvest. Each panel presents a separate regression model. Each panel presents a separate regression model. Outcome variables are listed on top. The unit of observation is the household. *Woredas* are districts and *kebeles* are wards within them. Pooled 2011,2013, and 2015 rounds of the Ethiopian Socio-Economic Surveys. All models control for survey round indicators, region indicators, the number of years of aid assistance received by the *woreda* prior to 2005, three indicators for whether the *woreda* received aid in 2004. Standard errors are in parentheses and are clustered at the *woreda*-level. Significance levels: *10%, **5%, and ***1%. "*p*-value : kebele vs. woreda" reports the *p*-value for a test of equality between the coefficients in the first and second row of panel B. The bottom panel displays the outcome mean, standard deviation, and total number of observations and clusters.

Table 2.12: Placebo test on employment participation and sectoral composition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. No individual controls							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	2.987	-1.168	-2.124	3.250	-3.542	-0.356	0.957**
	(2.486)	(1.046)	(2.050)	(2.848)	(2.477)	(0.380)	(0.451)
Mean Dep. Var. (%)	73.73	4.150	22.12	81.16	14.18	1.765	0.702
Observations	159,902	159,902	159,902	116,321	116,321	116,321	116,321
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	No	No	No	No	No	No
Panel B. Individual controls added							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	3.291	-1.168	-2.124	3.250	-3.542	-0.356	0.957**
	(2.486)	(1.046)	(2.050)	(2.848)	(2.477)	(0.380)	(0.451)
Mean Dep. Var. (%)	73.73	4.150	22.12	81.16	14.18	1.765	0.702
Observations	159,902	159,902	159,902	116,321	116,321	116,321	116,321
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each cell reports an estimate of β; standard errors in parenthesis are clustered at the district level. Each column reports an estimate for a different dependent variable.

In Panel A, each model includes district fixed effects and district controls. In Panel B, each model includes district fixed effects, district controls, and individual controls. Column (4)-(7) are conditional on being employed. The sample consists of individuals aged 17-65, pooling data from the 1999 and 2005 LFS rounds, sampled in 391 districts in each round. Individual observations are weighted by sampling weights that are proportional to district population. All models are estimated using ordinary least squares. The means of district-level and individual-level controls are shown in Table 2.1. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table 2.13: Effects on employment participation and sectoral composition, controlling for population density

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Somali region excluded							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	-1.996	0.904	1.119	-6.362**	5.770**	-0.029	0.272
	(2.204)	(0.685)	(1.947)	(2.660)	(2.345)	(0.473)	(0.186)
Mean Dep. Var.	83.29	1.703	15	84.26	11.50	1.340	0.500
Observations	100,731	100,731	100,731	83,319	83,319	83,319	83,319
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel B. Population density included as control							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	-1.867	0.794	1.104	-5.201*	4.011*	0.199	0.378*
	(2.297)	(0.707)	(2.026)	(2.687)	(2.364)	(0.475)	(0.212)
Mean Dep. Var.	83.17	1.701	15.12	84.25	11.54	1.328	0.494
Observations	100,731	100,731	100,731	83,319	83,319	83,319	83,319
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each cell reports an estimate of β for different dependent variables; standard errors in parenthesis are clustered at the district level. Each column has a different dependent variable. In Panel A, the sample excludes districts sampled in the Somali Region from either the 2005 or the 2013 LFS round. This region was not sampled in the 2007 census. In Panel B, district population density (000′ people/sq. km) estimated from the 2007 census, and interacted with a dummy variable equal to one if the year is 2013, is added as a control. The sample consists of individuals aged 17-65, pooling data from the 2005 and 2013 LFS rounds. Columns (4)-(7) restrict the sample only to those that are currently employed. Individual observations are weighted by sampling weights that are constant within a district across time. All models are estimated using ordinary least squares. The means of district-level and individual-level controls are shown in Table 2.1. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table 2.14: Effects on employment participation and sectoral composition, controlling for pre-PSNP shocks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Interaction with belg rainfall in 2002							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
Estimate of β:	-0.989	0.504	0.495	-7.078**	5.912**	0.160	0.662**
	(2.708)	(0.807)	(2.409)	(3.152)	(2.840)	(0.497)	(0.309)
Coef. On Interaction term:	-1.650	-1.004	2.617	-3.579	1.426	0.567	1.024*
	(3.457)	(0.972)	(3.083)	(4.326)	(4.007)	(0.671)	(0.587)
Mean Dep. Var. (%)	83.17	1.701	15.12	84.25	11.54	1.328	0.494
Observations	105,323	105,323	105,323	86,768	86,768	86,768	86,768
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel B. Interaction with emergency assistance received in 2005							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
Estimate of β:	0.162	0.511	-0.627	-6.317**	6.168**	-0.002	0.217
	(2.317)	(0.743)	(2.041)	(2.999)	(2.622)	(0.587)	(0.257)
Coef. On Interaction term:	0.140	0.826	-1.032	3.040	-3.500	0.108	-0.047
	(3.661)	(1.115)	(3.649)	(4.499)	(3.958)	(0.619)	(0.355)
Mean Dep. Var. (%)	83.17	1.701	15.12	84.25	11.54	1.328	0.494
Observations	105,323	105,323	105,323	86,768	86,768	86,768	86,768
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The first row in each panel reports an estimate of β for different dependent variables; standard errors in parenthesis are clustered at the district level. Each column has a different dependent variable. The second row in each panel reports the estimated coefficient of an interaction term with the standardized measure of rainfall for the 2002 *Belg* rainy season (Panel A), and a dummy variable equal to one if the district has received emergency assistance in 2005 (Panel B). The sample consists of individuals aged 17-65, pooling data from the 2005 and 2013 LFS rounds, in the 453 districts sampled in both rounds. Columns (4)-(7) restrict the sample only to those that are currently employed. Individual observations are weighted by sampling weights that are proportional to district population. All models are estimated using ordinary least squares. The means of district-level and individual-level controls are shown in Table 2.1. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table 2.15: Effects on employment participation and sectoral composition, using the unbalanced sample and without weights

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Individual observations from unbalanced panel of districts							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	-0.356	0.891	-0.510	-5.673**	5.605**	-0.064	0.262
	(2.227)	(0.640)	(2.030)	(2.447)	(2.169)	(0.429)	(0.166)
Mean Dep. Var.	83.02	1.699	15.27	84.36	11.46	1.331	0.478
Observations	111,674	111,674	111,674	91,676	91,676	91,676	91,676
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel B. Individual observations are unweighted							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	1.571	0.424	-1.991	-5.413**	5.390***	-0.358	0.716**
	(2.650)	(0.506)	(2.591)	(2.246)	(2.020)	(0.333)	(0.349)
Mean Dep. Var.	82.38	1.727	15.88	83.70	11.81	1.234	0.706
Observations	105,323	105,323	105,323	86,768	86,768	86,768	86,768
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each cell reports an estimate of beta for different dependent variables; standard errors in parenthesis are clustered at the district level. Each column has a different dependent variable.

In Panel A, the sample is restricted to women. In Panel B, the sample is restricted to men. The sample consists of individuals aged 17-65, pooling data from the 2005 and 2013 LFS rounds, in the 453 districts sampled in both rounds. Columns (4)-(7) restrict the sample only to those that are currently employed. Individual observations are weighted by sampling weights that are proportional to district population. All models are estimated using ordinary least squares. The means of district-level and individual-level controls are shown in Table 2.1. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table 2.16: Effects on employment participation and sectoral composition, without controls or district fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. DID estimates: No controls and no district fixed effects							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	-2.261	0.978	-0.403	-2.704	3.819**	-0.491	-0.754**
	(1.525)	(0.358)	(1.437)	(2.079)	(1.711)	(0.390)	(0.368)
Mean Dep. Var.	83.18	1.7	15.12	84.25	11.54	1.33	0.49
Observations	105,323	105,323	105,323	86,779	86,779	86,779	86,779
District Fixed Effects	No	No	No	No	No	No	No
District Controls	No	No	No	No	No	No	No
Individual Controls	No	No	No	No	No	No	No
Panel B. DID estimates with district fixed effects and no controls							
dependent variable:	Employed	Unemployed	Inactive	Self-employed in agriculture	Self-employed out of agriculture	Private Labourer	Public Labourer
	-2.349	0.323	2.026	-3.247	4.162**	-0.390	-0.772**
	(1.514)	(0.354)	(1.430)	(2.030)	(1.682)	(0.394)	(0.379)
Mean Dep. Var.	83.18	1.7	15.12	84.25	11.54	1.33	0.49
Observations	105,323	105,323	105,323	86,779	86,779	86,779	86,779
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	No	No	No	No	No	No	No
Individual Controls	No	No	No	No	No	No	No

Notes: Each cell reports an estimate of beta for different dependent variables; standard errors in parenthesis are clustered at the district level. Each column has a different dependent variable.
In Panel A, each model does *not* include any district fixed effects or district controls. In Panel B, each model includes only district fixed effects. The sample consists of individuals aged 17-65, pooling data from the 2005 and 2013 LFS rounds. Columns (4)-(7) restrict the sample only to those that are currently employed. Individual observations are weighted by sampling weights that are proportional to district population. All models are estimated using ordinary least squares. The means of district-level and individual-level controls are shown in Table 2.1. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Chapter 3

Medium-term Effects of a Role Model Intervention in Rural Ethiopia

Acknowledgements: This chapter is co-authored with Tanguy Bernard (Bordeaux School of Economics, University of Bordeaux & CNRS); Stefan Dercon (University of Oxford); Kate Orkin (University of Oxford); Alemayehu Seyoum 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.

3.1 An experiment to raise aspirations

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).

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

[2] Recent studies include Haushofer et al. (2020); McKelway (2021); Angelucci and Bennett (2021); Ashraf et al. (2022); Barker et al. (2022); Ghosal et al. (2022). See Kremer et al. (2019) for a review.

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 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 et al. (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 et al., 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 et al., 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 et al., 2021; Cecchi et al., 2022; McKenzie et al., 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 psychological characteristics on decision-making in poor settings beyond aspirations, including self-regulation, self-efficacy, grit, and preferences (Heller et al., 2016; Blattman et al., 2017; Ashraf et al., 2022; Campos et al., 2017; Alan et al., 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 et al., 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

[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).

countries (Greene et al., 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 et al., 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 3.2 provides a simple theoretical framework to model aspirations, combining elements from existing theoretical models of aspirations. Section 3.3 describes the context of the study. Section 3.4 discusses our intervention, design, and gives a brief description of our main estimation strategy and tests for experimental integrity. Section 3.5 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 3.6 reports a series of tests for the presence of spillovers. Section 3.7 concludes.

3.2 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.

3.2.1 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 et al., 1999; Dalton et al., 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 $z(c_t - a_t) \leq 0$

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

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 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:

$$u_{c_t} + z_{c_t} = \beta.(1 + f_{k_t} - \delta).(u_{c_{t+1}} + z_{c_{t+1}}) \tag{3.1}$$

$$u_{l_t} = \beta.f_{e_t}.(u_{c_{t+1}} + z_{c_{t+1}}) \tag{3.2}$$

Equation 3.1 governs choices between consumption today versus saving and investing for future consumption; Equation 3.2 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).

3.2.2 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 3.1 and 3.2. 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 3.2, 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 3.1 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 3.1, 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 3.1 and reduce $u_{c_{t+1}}$ until equality across both Equation 3.1 and 3.2 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 et al., 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 3.2 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 et al., 1986; Madrian and Shea, 2001), goals (Heath et al., 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 et al. (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.

3.3 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.

3.3.1 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 3.2 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 3.9 details the construction of the variables used in our analysis.

3.3.2 Characteristics of the sample

Table 3.1 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 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 3.9).

In comparison, there are fewer differences in other beliefs and preferences by wealth terciles (Appendix Table 3.9). 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 et al., 2016; Kahneman et al., 1986).

[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?*".

Table 3.1: 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 (baseline full sample)						
Value of durable assets excluding tools (USD) per ad. equiv. PPP	18.72	0.00	5.49	51.25	0.00	1119
Value of tools (USD) per ad. equiv. PPP	24.32	16.28	17.63	38.18	0.00	1111
Total savings (USD) per ad. equiv. PPP	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 (baseline full sample)						
Value of livestock (USD) per ad. equiv. PPP	411.49	281.09	348.13	590.91	0.00	1110
Value of cattle (USD) per ad. equiv. PPP	366.87	235.44	296.43	524.95	0.00	1118
Value of sheep or goat (USD) per ad. equiv. PPP	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 (baseline full sample)						
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 (baseline full sample)						
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 (baseline full sample)						
Value of house (USD) per ad. equiv. 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 3.9. 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 3.9.1.

3.4 Experimental design and estimation strategy

3.4.1 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 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

[6] The documentaries, with English subtitles, and one of four placebo segments are available at https://www.youtube.com/channel/UCqfoNjCzt8YPjTRWQaMQfAg. Appendix Section 3.8 summarises two documentaries and one placebo segment.

in Section 3.5.2. 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

[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.

alone, we also invited another group of households, in the same villages, to a "placebo" 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.

3.4.2 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 3.3. 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 3.6, 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 3.6 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 3.5 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 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 3.13). There are no differences in compliance rates across treatment and placebo groups.

3.4.3 Empirical strategy

Our main specification is:

$$y_i = \alpha + \delta T_i + \rho P_i + X'_{i1}\pi + \varepsilon_i \tag{3.3}$$

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

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

and highest school grade completed are for the head of the household.

We will use Equation 3.3 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 3.10.3 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 3.10 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 et al. (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 outcomes reported in our main exhibits. Following Bessone et al. (2021) and Kling et al. (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

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

spillovers in our experiment in Section 3.5.4 and 3.6, though we had not pre-specified these summary indices in our analysis plan.

3.4.4 Balance and attrition

Appendix Table 3.10 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).

3.5 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 3.3: 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.

3.5.1 Effects on economic outcomes five years after the screening

Effort, investment and productive assets

In Table 3.2, 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

[10] Results remain broadly robust to alternative specifications presented in Appendix Table 3.18. 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.

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 3.18.

We combine all outcomes from the first three panels of Table 3.2 into a single agricultural investment index. Treated households significantly increased this index of investment by 0.14 and 0.18 standard deviations relative to the placebo and control

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

group, five years after exposure to the role models in the videos. Overall, Table 3.2 gives support to our theoretical predictions.

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 3.2 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 3.3 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

[12] As noted in Section 3.4.4, all outcomes in Table 3.3 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.

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 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 3.3 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.

Table 3.2: 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 3.9. 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 3.7.

Table 3.3: 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 3.9. 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 3.7.

Consumption, durable goods and well-being

Table 3.4 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 3.4). 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 3.2.2, 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 findings are consistent with the substi-

Table 3.4: 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 3.9. 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 3.7.

tution 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 3.4). 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 3.4). 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 3.4 (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.

3.5.2 Where do these results come from?

Early impact

Our video intervention began to change household behaviour soon after the screenings. In Table 3.5, 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 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 3.2, 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 3.25) suggestive of cash savings that were later invested in relatively lumpy assets such as livestock as reported in Table 3.2. 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 3.3. As in Table 3.3 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.

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 3.1 displays effects on our indices of aspirations and expectations, across survey rounds after our intervention. Appendix Table 3.20 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

Table 3.5: 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 3.9. 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 3.9.1. 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 3.3, 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.

education (Appendix Table 3.20). 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 3.21).

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

3.9.4 details the construction of the measures used to test these alternative mechanisms.

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 3.6). After the intervention, there is no change in the share of patient, impatient, very impatient respondents, or present-biased respondents — categorised as in Ashraf et al. (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]

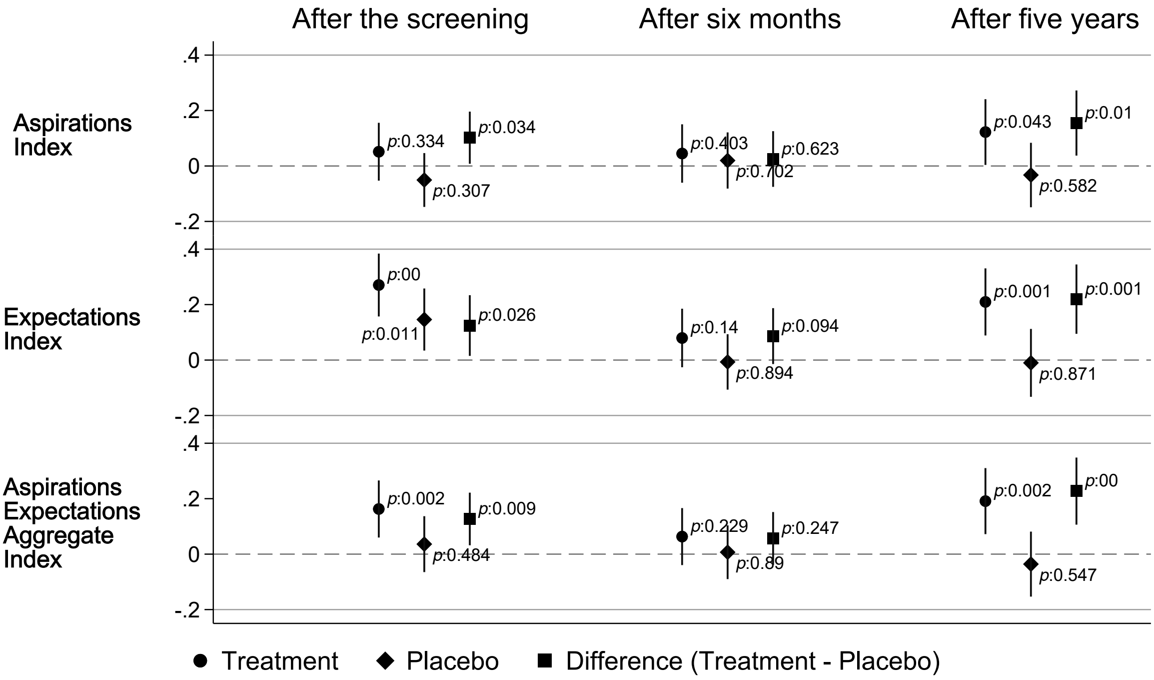
While risk does not enter explicitly the theoretical framework, it could be trivially extended. For example, if future returns are risky, then higher risk aversion would induce less effort and investment in the future. The increased salience of a plausible future through the intervention may have reduced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

Risk preferences have not shifted either (top panel, Table 3.6). 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.

[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 et al. (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.



[a] *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 3.20 reports results across all three indices and their individual components.

Figure 3.1: Treatment effects on the aspirations and expectations indices. [a]

Table 3.6: 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 3.9. 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 et al., 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 3.9. 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.

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 3.6). 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 3.6). 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 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 3.6, 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.

3.5.3 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 3.5 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 3.1 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 et al. (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 3.7).

3.5.4 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 3.4).[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

[15] In an exploratory analysis, we also test whether effects vary by the terciles of durable assets used to categorise our sample in Table 3.1. 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 3.22).

Table 3.7: 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 3.9. 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 3.2, with daily minutes in leisure being recoded to be negative. The educational investment index includes all outcomes reported in Table 3.3. The welfare index includes all outcomes reported in Table 3.4, 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 et al. (2007). As the omnibus index is for the whole household, we use the household head's aspirations and expectations aggregate index.

(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]

3.6 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

[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.

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 3.11 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:

$$y_{i3} = \alpha + \delta^s T_i + \rho^s P_i + \varphi C_i + X_{i2}'\pi^s + \varepsilon_i \tag{3.4}$$

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 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

[17] See Appendix Figure 3.5 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 3.14.

Table 3.8: 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 3.2, with daily minutes in leisure being recoded to be negative. The educational investment index includes all outcomes reported in Table 3.3. The welfare index includes all outcomes reported in Table 3.4, 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 et al. (2007). As the omnibus index is for the whole household, we use the household head's aspirations and expectations aggregate index.

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 3.8). 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 3.5. Our within-village comparisons of treated household relative to placebo and within-village controls remain very close to those reported throughout Section 3.5 (Column 4 and 5).

In our second empirical specification to identify spillovers, we expand the model in Equation 3.4 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 3.23). 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"

[19] Appendix Section 3.10.3 provides the equation of the estimated specification.

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 include two additional terms to Equation 3.4: 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 3.5 graphically illustrates the spatial variation in treatment intensity in our sample.[21] We find no evidence of spillovers across villages (Column 4, Appendix Table 3.24). 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. At least in our setting, there is either no evidence of spillovers within communities or such effects, if present, take longer than five years to occur or a larger sample of villages to detect them.

[20] In Appendix Section 3.10.3 we provide details of how we empirically selected the length of the radius to be 1km, following the methodology of Egger et al. (2022).

[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.

3.7 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.

Appendix

Section 3.8 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 3.9 provides details on additional data sources as well as procedures to construct all variables used throughout the paper. Section 3.10 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 3.11 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.

3.8 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.

3.9 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.

3.9.1 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 rebased 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).

3.9.2 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 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]

[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.

[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.

3.9.3 Agricultural investment

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 twelve months: feed, veterinary supplies, and hired labour. Spending on purchase of livestock is the sum of spending on all animals in the last twelve months.

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.

[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.

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 twelve 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.

3.9.4 Beliefs, preferences, and information

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 et al. (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:

(i). Would they prefer the gift of 100 ETB today or could instead choose to receive a gift of 125 ETB in one month?

(ii). If they answer 100 ETB to question 1, they are asked if they prefer 100 ETB today or 150 ETB in one month.

(iii). 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]

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 et al. (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 3.16 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 et al. (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

[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.

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.

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 3.15. 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 3.15. 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

[26] The distribution of individuals across categories is similar to results from the same measure in the Ethiopian Rural Household Survey (Hill et al., 2013).

[27] For example, Harrison et al. (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.

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).

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

[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 et al., 2005; Reimanis and Posen, 1980; Van Haaften and van de Vijver, 1999)

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 et al. (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".

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 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.

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.

3.9.5 Consumption, food security, housing, and well-being

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). 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

[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.

Agriculture's food insecurity questionnaire (Bickel et al., 2000; Andrews et al., 2000) adapted for Ethiopia (Hadley et al., 2008).

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.

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.

3.10 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.

3.10.1 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.

3.10.2 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.

(i). **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 3.6.

- **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

[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.

available upon request).

(ii). **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 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 3.19). We did pre-specifiy that this would be a secondary analysis that we had planned to carry out using the household-member-level dataset,

[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.

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.

(iii). **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 3.25).

  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.

(iv). **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 3.26).

- **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.

(v). **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 3.9.4.
- **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 3.9.4.

- **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 3.9.4.

3.10.3 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 3.6.

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 3.5 of the main text focus on treated villages only. In Section 3.6, 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:

$$y_{i3} = \alpha + \varphi_1 TI_v + \varphi_2 PI_v + \delta_1 T_i * TI_v + \delta_2 T_i * PI_v + \rho_1 P_i * TI_v + \rho_1 P_i * PI_v + + X'_{i2}\pi^{si} + \varepsilon_i \tag{3.5}$$

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 3.4 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 $r \in \{0, 1\}$; $\ldots$; 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:

$$\begin{aligned} y_{i3} =& \alpha + \varphi_1^r C_i + \delta_1^r T_i + \rho_1^r P_i + \\ & \sum_{d=1}^{D} \left(\delta_1^d \mathrm{V}_{i\neg v}^d + \delta_2^d \mathrm{TH}_{i\neg i}^d \right) + X_{i2}' \pi_{1r} + \varepsilon_i \end{aligned} \tag{3.6}$$

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:

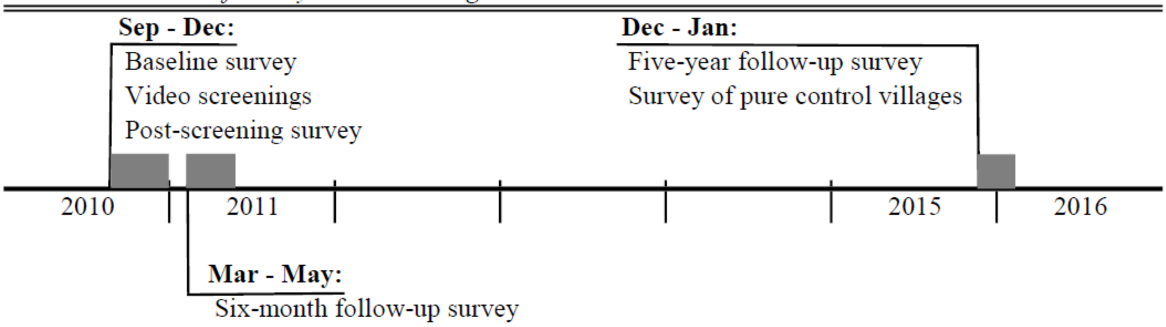
(a) δ_1^r gives us the (direct) intent-to-treat effect of the aspirational videos on individuals/households in treated villages.

(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.

3.11 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 3.2: Timeline of the study [a]

[a] *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.

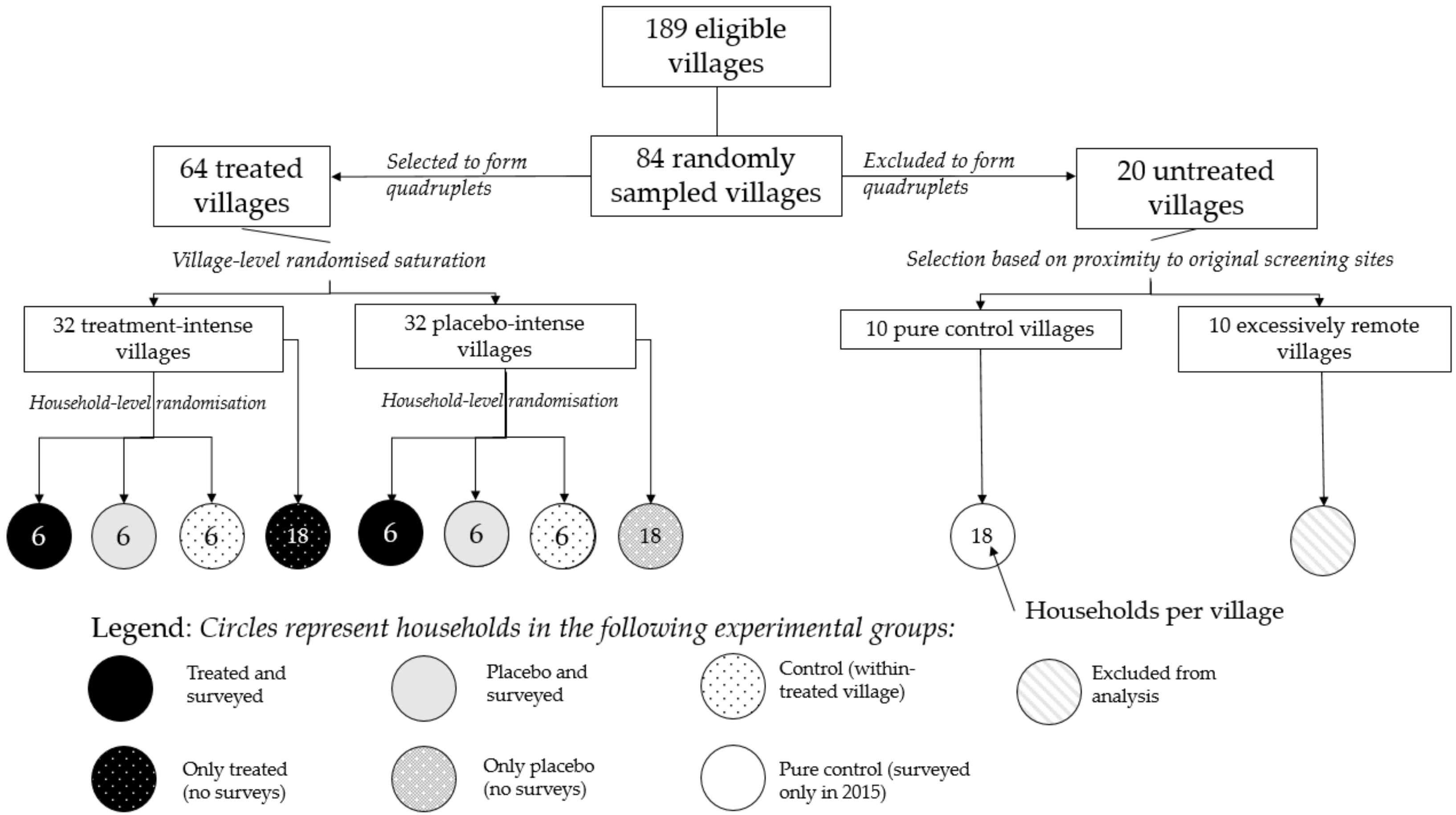
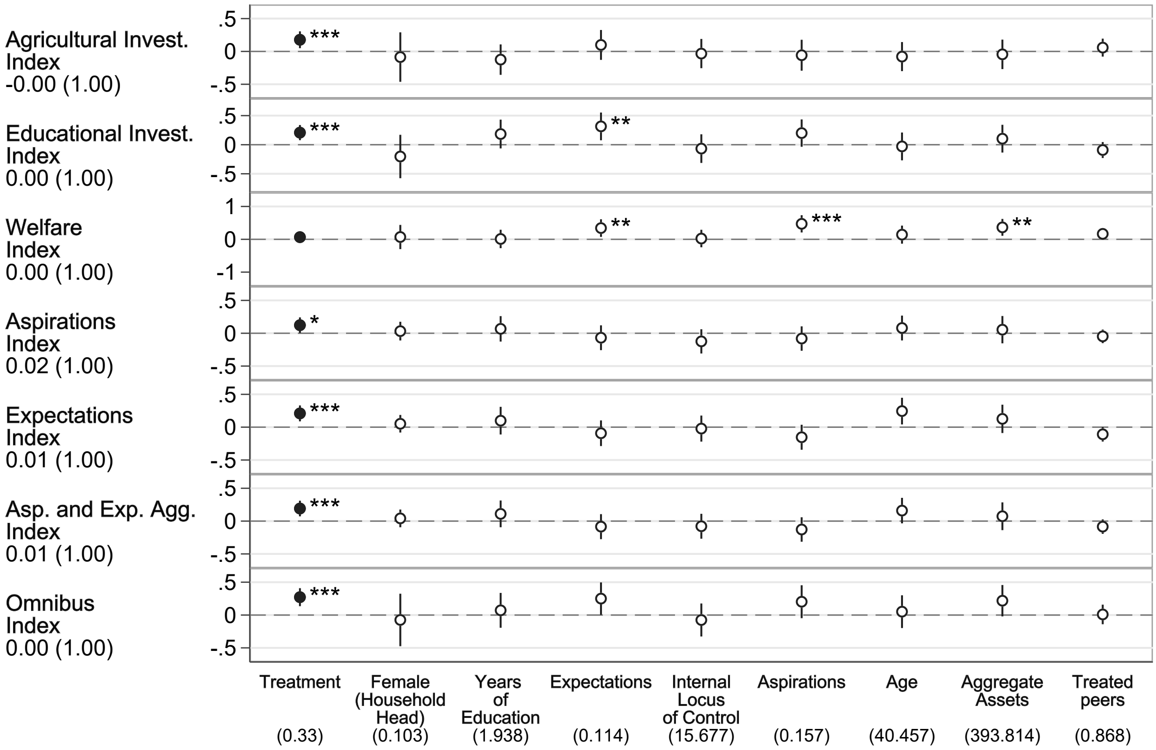


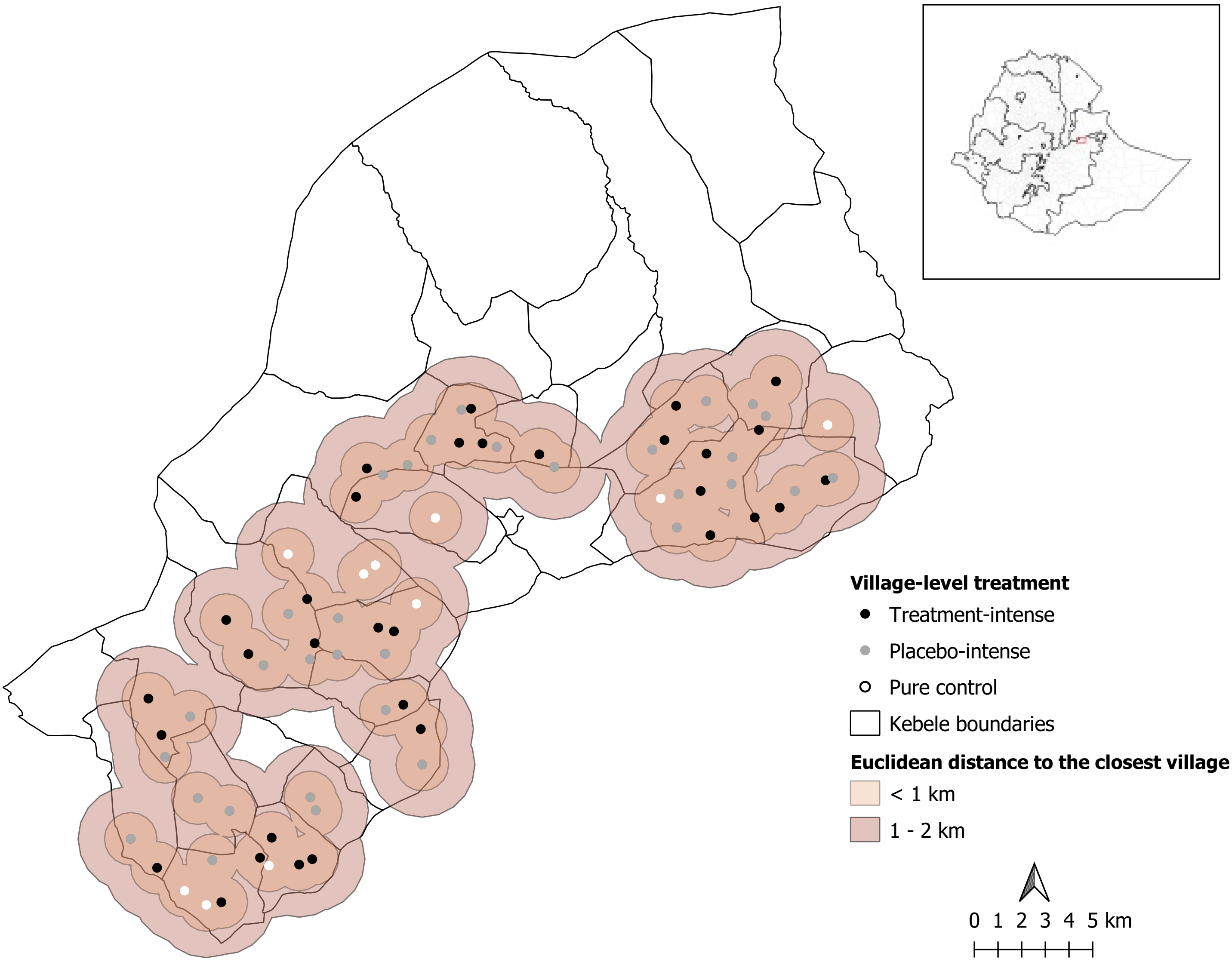
Figure 3.3: Study design [a]

[a] *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.



[a] *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 3.9.4. 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 et al., 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 3.7, are inverse-covariance-weighted averages standardised relative to the within-village control group, following Anderson (2008).

Figure 3.4: Heterogeneous treatment effects on summary indices [a]



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.

Figure 3.5: Villages in the study

Table 3.9: 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
Risk and time preferences						
Risk aversion: most to least risk averse (1 to 5)	3.20	3.30	3.10	3.22	0.35	2010
% 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

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 et al., 2006). All other variables are defined in detail in Appendix 3.9.

Table 3.10: 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 et al. (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 3.9.1.

Table 3.11: 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 3.9. 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 3.12: 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 3.13: 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 3.14: 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 3.15: 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 et al. (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 3.16: Tabulation of responses to hypothetical time preference questions

			Indifferent between ETB 100 ETB in one and month and X in two months			
			Patient X <125	Somewhat impatient 125 <X <150	Most impatient 150 <X	Total
Indifferent between 100 ETB now and X in one month	Patient	X <125	**602** **28.9%**	**45** **2.2%**	**38** **1.8%**	685 32.8%
	Somewhat impatient	125 <X <150	**61** **2.9%**	**91** **4.4%**	**116** **5.6%**	268 12.8%
	Most impatient	150 <X	**53** **2.5%**	**37** **1.8%**	**1,044** **50%**	1134 54.3%
	Total		716 34.3%	173 8.3%	1,198 57.4%	2087 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 3.9.4.

Table 3.17: 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 3.9. 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 3.9.1. 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 3.18: Robustness test for household-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.
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 3.9. 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 3.9.1. The unit of observation is the household. The number of observations varies slightly across rows because some respondents do not answer all questions.

Table 3.19: 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 3.5 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 3.9. 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 3.9.1. The unit of observation is the household. The number of observations varies slightly across rows because some respondents do not answer all questions.

Table 3.20: 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 3.9. 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 3.9.1. 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 3.21: 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 3.9. 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 3.9.1. 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 3.22: Heterogeneous treatment effects on summary indices after five years by terciles of durable assets

After five years	(1) Medium	(2) High	(3) Treat.#Low	(4) Treat.#Med.	(5) Treat.#High	(6) Control Mean (SD) Total obs.
Agricultural investment index	0.12	0.28***	0.17	0.20*	0.11	-0.00
	(0.08)	(0.09)	(0.11)	(0.10)	(0.11)	(1.00)
	[0.24]	[0.00]***	[0.40]	[0.06]*	[0.34]	1061
Educational investment index	0.09	0.22**	0.09	0.31**	0.22**	0.00
	(0.08)	(0.09)	(0.10)	(0.12)	(0.11)	(1.00)
	[0.34]	[0.02]**	[0.57]	[0.02]**	[0.16]	1061
Welfare index	0.24**	0.51***	-0.01	-0.01	0.17	0.00
	(0.10)	(0.10)	(0.12)	(0.12)	(0.13)	(1.00)
	[0.04]**	[0.00]***	[0.92]	[0.91]	[0.34]	1063
Aspiration index	0.07	0.12	-0.01	0.26***	0.08	0.02
	(0.08)	(0.08)	(0.10)	(0.10)	(0.09)	(1.00)
	[0.36]	[0.13]	[0.92]	[0.02]**	[0.36]	1901
Expectations index	0.13*	0.37***	0.15*	0.27***	0.12	0.01
	(0.08)	(0.08)	(0.09)	(0.10)	(0.10)	(1.00)
	[0.22]	[0.00]***	[0.40]	[0.02]**	[0.34]	1900
Asp. and exp. aggregate index	0.11	0.27***	0.08	0.32***	0.11	0.01
	(0.08)	(0.08)	(0.10)	(0.09)	(0.10)	(1.00)
	[0.24]	[0.00]***	[0.57]	[0.01]***	[0.34]	1901
Omnibus index	0.22***	0.47***	0.11	0.25**	0.27**	0.00
	(0.08)	(0.09)	(0.10)	(0.11)	(0.12)	(1.00)
	[0.04]**	[0.00]***	[0.57]	[0.03]**	[0.15]	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 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 are inverse-covariance-weighted averages, standardized relative to the within-village control group, following Anderson (2008), as described in Table 3.7.

Table 3.23: 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 3.2, with daily minutes in leisure being recoded to be negative. The educational investment index includes all outcomes reported in Table 3.3. The welfare index includes all outcomes reported in Table 3.4, 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 et al. (2007). As the omnibus index is for the whole household, we use the household head's aspirations and expectations aggregate index.

Table 3.24: 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 3.6 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 3.2, with daily minutes in leisure being recoded to be negative. The educational investment index includes all outcomes reported in Table 3.3. The welfare index includes all outcomes reported in Table 3.4, 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 et al. (2007). As the omnibus index is for the whole household, we use the household head's aspirations and expectations aggregate index.

Table 3.25: 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 3.9. 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 3.26: 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 3.9. 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.

Concluding remarks and future work

In conclusion, this thesis adds to our understanding of the effects that different constraints and policies have on rural economic development, focusing on studies based in Guinea-Bissau and Ethiopia. In the first chapter, I show that informational asymmetries between traders and producers can be reduced by a scale-able digital intervention in order to increase producers' incomes. In the second chapter, I provide evidence that a large social protection programme indirectly affected non-beneficiaries, through changing the occupational activities in rural labour markets towards non-agricultural self-employment. In the third chapter, I find that altering individuals' goals for their future can be an effective tool to increase investments over the medium-term, especially in children's human capital and in additional inputs for agricultural enterprises.

A common policy thread runs through the first and third chapters, emphasizing the significance of digital interventions, such as video-based or mobile-based platforms, as tools to disseminate information. These cost-effective methods can have meaningful economic changes and enable rural households to make more informed decisions to improve their livelihoods.

Furthermore, this thesis emphasizes the importance of market interactions and spillover effects when designing and assessing rural development programmes. All studies in the thesis were designed to estimate spillover effects. The first two chapters find evidence that indirect effects of the policies studied are considerable, and the result of complex market interactions between beneficiaries and non-beneficiaries in rural markets. The final chapter does not find evidence of spillover effects, possibly because there are fewer or smaller market interactions as a result of the intervention.

I am expecting my future research to build on this thesis in three ways. First, in an-ongoing follow-up study, I am testing complementary and novel ways of delivering market information among a similar sample of cashew-nut producers in Guinea-Bissau. This follow-up study aims to test the effectiveness of sending information through focal points selected by the villages, rather than to a randomly selected group of producers. This selection criteria has the advantage of being able to deliver information using cheaper internet-based apps, such as WhatsApp, and increasing interaction with the users. Second, I plan to further explore the general equilibrium effects of this type of information interventions, by designing an experiment where the share of users targeted by the service at a local-level (across different villages) is explicitly randomised. Finally, I am in the process of obtaining the latest round of the Ethiopian Labour Force Survey to study whether my existing findings have changed over time, particularly as structural transformation in Ethiopia sped up before the pandemic in 2020.

In sum, this thesis calls for continued attention to context-specific solutions and a comprehensive understanding of the complex interactions within rural economies to effectively address the challenges of poverty and promote lasting prosperity.

Bibliography

Abdulai, A., C. B. Barrett, and J. Hoddinott (2005): "Does Food Aid Really Have Disincentive effects? New Evidence from Sub-Saharan Africa," *World Development*, 33, 1689–1704.

Abebe, G., S. Franklin, C. Imbert, and C. Mejia-Mantilla (2021): "Urban Public Works in Spatial Equilibrium: Experimental Evidence from Ethiopia," *CEPR Discussion Paper No. DP16691*.

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

Adhvaryu, A., J. Fenske, and A. Nyshadham (2019): "Early Life Circumstance and Adult Mental Health," *Journal of Political Economy*, 127, 1516–1549.

Aggarwal, S., E. Francis, and J. Robinson (2018): "Grain Today, Gain tomorrow: Evidence from a Storage Experiment with Savings Clubs in Kenya," *Journal of Development Economics*, 134, 1–15.

Aker, J. C. (2010): "Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger," *American Economic Journal: Applied Economics*, 2, 46–59.

Aker, J. C., I. Ghosh, and J. Burrell (2016): "The Promise (and Pitfalls) of ICT for Agriculture Initiatives," *Agricultural Economics (United Kingdom)*, 47, 35–48.

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

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

ALBUQUERQUE, R. A., B. ARAUJO, L. BRANDAO-MARQUES, G. MOSSE, P. VLETTER, AND H. ZAVALE (2022): "Market Timing, Farmer Expectations, and Liquidity Constraints," *SSRN Electronic Journal*, 1–48.

ALDERMAN, H. AND R. YEMTSOV (2012): "Productive Role of Safety Nets," *Social Protection and Labor Discussion Paper*, 1–93.

ALLEN, T. (2014): "Information Frictions in Trade," *Econometrica*, 82, 2041–2083.

ANDERSEN, S., G. W. HARRISON, M. I. LAU, AND E. E. RUTSTROM (2008): "Eliciting Risk and Time Preferences," *Econometrica*, 76, 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, 1481–1495.

ANDERSSON, C., A. MEKONNEN, AND J. STAGE (2011): "Impacts of the Productive Safety Net Program in Ethiopia on Livestock and Tree Holdings of Rural Households," *Journal of Development Economics*, 94, 119–126.

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*.

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*, ed. 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.

Asher, S. and P. Novosad (2020): "Rural Roads and Local Economic Development," *American Economic Review*, 110, 797–823.

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, 635–672.

Bagga, A., M. Holmlund, N. Khan, S. Mani, E. Mvukiyehe, and P. Premand (2023): "Do Public Works Programs Have Sustained Impacts?: A Review of Experimental Studies from LMICs," *World Bank Policy Research Working Paper 10471*, 1–40.

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, 844–860.

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

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

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

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, 187–206.

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

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

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, 824–59.

Bari ®, Faisal and Malik ®, Kashif and Meki ®, Muhammad and Quinn ®, Simon (2021): "Asset-based Microfinance for Microenterprises: Evidence from Pakistan," *Centre for the Study of African Economies Working Paper 2021-03*, 1–46.

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, 527–45.

Basu, A. K. (2013): "Impact of Rural Employment Guarantee Schemes Impact of Rural Employment Guarantee Schemes on Seasonal Labor Markets : Optimum Compensation and Workers ' Welfare," *The Journal of Economic Inequality*, 1–38.

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

Beaman, L., A. BenYishay, J. Magruder, and A. M. Mobarak (2021): "Can Network Theory-Based Targeting Increase Technology Adoption?" *American Economic Review*, 111, 1918–43.

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

Becker, A., P. Finger, A. Meyer-Christoffer, B. Rudolf, K. Schamm, U. Schneider, and M. Ziese (2013): "A Description of the Global Land-Surface Precipitation Data Products of the Global Precipitation Climatology Centre with Sample Applications Including Centennial (Trend) Analysis from 1901 to Present," *Earth System Science Data*, 5, 71–99.

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

Behrman, J. R. (1999): "Labor Markets in Developing Countries," in *Handbook of Labor Economics*, vol. 3, 2859–2939.

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "Inference on Treatment Effects after Selection among High-Dimensional Controls," *Review of Economic Studies*, 81, 608–650.

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

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

Berg, E., S. Bhattacharyya, D. Rajasekhar, and R. Manjula (2018): "Can Public Works Increase Equilibrium Wages? Evidence from India's National Rural Employment Guarantee," *World Development*, 103, 239–254.

Berhane, G., D. O. Gilligan, J. Hoddinott, N. Kumar, and A. S. Taffesse (2014): "Can Social Protection Work in Africa? The Impact of Ethiopia's Productive Safety Net Programme," *Economic Development and Cultural Change*, 63, 1–26.

Berhane, G., J. Hoddinott, N. Kumar, A. Seyoum, M. T. Diressie, Y. Yohannes, R. Sabates-wheeler, M. Handino, J. Lind, M. Tefera, and F. Sima (2011): "Evaluation of Ethiopia's Food Security Program: Documenting Progress in the Implementation of the Productive Safety Nets Programme and the Household Asset Building Programme," Tech. rep., International Food Policy Research Institute, Washington, D.C.

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, 189–224.

Bertrand, M., E. Duflo, and S. Mullainathan (2004): "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119, 249–275.

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

Besley, T. and S. Coate (1992): "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs," *American Economic Review*, 82, 249–261.

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, 1887–1941.

Bettinger, E., N. Cunha, G. Lichand, and R. Madeira (2021): "When the Effects of Informational Interventions Are Driven by Salience — Evidence from School Parents in Brazil," *Mimeo*, 1–128.

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, 395–407.

Bisin, A. and T. Verdier (2001): "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory*, 97, 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, 1165–1206.

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

Borusyak, K. and P. Hull (2021): "Non-Random Exposure to Exogenous Shocks: Theory and Applications," Working Paper 27845, National Bureau of Economic Research.

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

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, annurev–economics–080218–030333.

Broussar, N. H. and T. G. Tekleselassie (2012): "Youth Unemployment, Ethiopia Country Study," *International Growth Centre (IGC) Working Paper*, 1–39.

Bruhn, M. and D. McKenzie (2009): "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 1, 200–232.

Bryan, G., S. Chowdhury, and A. M. Mobarak (2014): "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh," *Econometrica*, 82, 1671–1748.

Burke, M., L. Falcao Bergquist, and E. Miguel (2019): "Sell Low and Buy High: Arbitrage and Local Price Effects in Kenyan Markets," *Quarterly Journal of Economics*, 134, 785–842.

Camacho, A. and E. Conover (2019): "The Impact of Receiving SMS Price and Climate Information on Small Scale Farmers in Colombia," *World Development*, 1–211.

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

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

Cardell, L. and H. Michelson (2023): "Price Risk and Small Farmer Maize Storage in Sub-Saharan Africa: New Insights into a Long-Standing Puzzle," *American Journal of Agricultural Economics*, 105, 737–759.

Casaburi, L., R. Glennerster, and T. Suri (2013): "Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone," *SSRN Electronic Journal*.

Casaburi, L. and T. Reed (2022): "Using Individual-Level Randomized Treatment to Learn about Market Structure," *American Economic Journal: Applied Economics*, 14, 58–90.

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.

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, 152–188.

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

Cole, S. A. and A. N. Fernando (2021): "'Mobile'izing Agricultural Advice Technology Adoption Diffusion and Sustainability," *The Economic Journal*, 131, 192–219.

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

Conley, T. G. and C. R. Udry (2010): "Learning About a New Technology: Pineapple in Ghana," *American Economic Review*, 100, 35–69.

Cont, W. and G. Porto (2014): "Measuring the Impact of a Change in the Price of Cashew Received by Exporters on Farmgate Prices and Poverty in Guinea-Bissau," *World Policy Research Working Paper 7036*, 1–47.

Courtois, P. and J. Subervie (2015): "Farmer Bargaining Power and Market Information Services," *American Journal of Agricultural Economics*, 97, 953–977.

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

de Chaisemartin, C. and X. D'Haultfoeuille (2022): "Two-way Fixed Effects and Difference-in-Differences with Heterogeneous Treatment Effects: A Survey," Tech. rep., National Bureau of Economic Research.

de Chaisemartin, C. and J. Ramirez-Cuellar (2023): "At What Level Should One Cluster Standard Errors in Paired and Small-Strata Experiments?" *American Economic Journal: Applied Economics*, Forthcoming, 1–65.

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

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

Deaton, A. and G. Laroque (1992): "On the Behaviour of Commodity Prices," *The Review of Economic Studies*, 59, 1–23.

Dell, M., B. F. Jones, B. a. Olken, and M. Gates (2014): "What Do We Learn from the Weather? The New Climate-Economy Literature," *Journal of Economic Literature*, 52, 740–798.

Dercon, S., D. O. Gilligan, J. Hoddinott, and T. Woldehanna (2009): "The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages," *American Journal of Agricultural Economics*, 91, 1007–1021.

Dercon, S. and P. Krishnan (1996): "Income Portfolios in Rural Ethiopia and Tanzania: Choices and Constraints," *Journal of Development Studies*, 32, 850–875.

Dinkelman, T. (2011): "The Effects of Rural Electrification on Employment: New Evidence from South Africa," *American Economic Review*, 101, 3078–3108.

Diop, B. Z. (2023): "Upgrade or Migrate: The Consequences of Input Subsidies on Household Labor Allocation," *University of Oxford Working Paper*, 1–94.

Duflo, E., R. Glennerster, and M. Kremer (2008): "Using Randomization in Development Economics Research: A Toolkit," in *Handbook of Development Economics*, ed. by T. Schultz and J. Strauss, North Holland, chap. 61, 3895–3962.

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

Eble, A. and M. Escueta (2022): "Demand, Supply, and Learning in a Very Low-income Context," *EdWorkingPaper No . 21-473*, 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, 2603–2643.

FABREGAS, R., M. KREMER, AND F. SCHILBACH (2019): "Realizing the Potential of Digital Development: The Case of Agricultural Advice," *Science*, 366, eaay3038.

FAFCHAMPS, M. AND R. V. HILL (2008): "Price Transmission and Trader Entry in Domestic Commodity Markets," *Economic Development and Cultural Change*, 56, 729–766.

FAFCHAMPS, M. AND B. MINTEN (2012): "Impact of SMS-Based Agricultural Information on Indian Farmers," *The World Bank Economic Review*, 26, 383–414.

FAFCHAMPS, M., C. UDRY, AND K. CZUKAS (1998): "Drought and Saving in West Africa: Are Livestock a Buffer Stock?" *Journal of Development Economics*, 55, 273–305.

FALCAO BERGQUIST, L., C. MCINTOSH, , AND M. STARTZ (2021): "Search Costs, Intermediation, and Trade: Experimental Evidence from Ugandan Agricultural Markets," *Mimeo*, 1–58.

FALK, A., A. BECKER, T. DOHMEN, B. ENKE, D. HUFFMAN, AND U. SUNDE (2018): "Global Evidence on Economic Preferences," *Quarterly Journal of Economics*, 133, 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.

FETZER, T. (2020): "Can Workfare Programs Moderate Conflict? Evidence from India," *Journal of the European Economic Association*, 18, 3337–3375.

FIELD, E. (2007): "Entitled To Work: Urban Property Rights and Labor Supply in Peru," *Quarterly Journal of Economics*, 122, 1561–1602.

Filipski, M. J., J. E. Taylor, G. A. Abegaz, T. Ferede, A. S. Taffesse, and X. Diao (2016): "General Equilibrium Impact Assessment of the Productive Safety Net Program in Ethiopia," *3ie Final Report*.

Filmer, D. and L. H. Pritchett (2001): "Estimating Wealth Effects Without Expenditure Data—or Tears: an Application to Educational Enrollments in States of India," *Demography*, 38, 115–132.

Foster, A. D. and M. R. Rosenzweig (1995): "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103, 1176–1209.

Franklin, S. (2014): "Youth Unemployment and Job Search in Urban Ethiopia," *World Bank Background paper*, 1–54.

——— (2020): "Enabled to Work: The Impact of Government Housing on Slum Dwellers in South Africa," *Journal of Urban Economics*, 118, 103265.

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

Gazeaud, J. and V. Stephane (2023): "Productive Workfare? Evidence from Ethiopia's Productive Safety Net Program," *American Journal of Agricultural Economics*, 105, 265–290.

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

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

GFDRE (2006): "Productive Safety Net Programme: Programme Implementation Manual, Phase II," Tech. rep., Government of the Federal Democratic Republic of Ethiopia. Ministry of Agriculture and Rural Development (MoARD), Addis Ababa.

——— (2010): "Productive Safety Net Programme: Programme Implementation Manual, Phase III," Tech. Rep. May, Government of the Federal Democratic Republic of Ethiopia. Ministry of Agriculture and Rural Development (MoARD), Addis Ababa.

——— (2013): "Poverty and Growth in Ethiopia (1995/96-2010/11)," Tech. rep., Government of the Federal Democratic Republic of Ethiopia. Ministry of Finance and Economic Development (MoFED), Addis Ababa.

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, 431–448.

Gill, P. (2010): *Famine and Foreigners: Ethiopia since Live Aid*, Oxford University Press.

Gilligan, D., J. Hoddinott, and A. S. Taffesse (2009): "The Impact of Ethiopia's Productive Safety Net Programme and Its Linkages," *Journal of Development Studies*, 45, 1684–1706.

Gilligan, D. O., J. Hoddinott, N. Kumar, and A. S. Taffesse (2011): "Impact of Social Protection on Food Security and Coping Mechanisms: Evidence from Ethiopia's Productive Safety Nets Program," *Journal of vestibular research : equilibrium & orientation*, 21, 297–8.

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

Goldberg, J. (2016): "Kwacha Gonna Do? Experimental Evidence about Labor Supply in Rural Malawi," *American Economic Journal: Applied Economics*, 8, 129–149.

Goyal, A. (2010): "Information, Direct Access to Farmers, and Rural Market Performance in Central India," *American Economic Journal: Applied Economics*, 2, 22–45.

Green, M. and T. Brock (2000): "The Role of Transportation in the Persuasiveness of Public Narratives," *Journal of Personality and Social Psychology*, 79, 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, 393–398.

Grosh, M., C. Del Ninno, E. Tesliuc, and A. Ouerghi (2008): *For Protection & Promotion: The Design and Implementation of Effective Safety Nets*, The World Bank, Washington, D.C.

Gupta, A., J. Ponticelli, and A. Tesei (2021): "Access to Information, Technology Adoption and Productivity: Large-scale Evidence from Agriculture in India," *Social Science Research Network Electronic Journal*, 1–46.

Hadley, C., D. Lindstrom, F. Tessema, and T. Belachew (2008): "Gender Bias in the Food Insecurity Experience of Ethiopian Adolescents," *Social Science & Medicine*, 66, 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, 80–104.

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.

Headey, D., F. B. Nisrane, I. Worku, M. Dereje, and A. S. Taffesse (2012): "Urban Wage Behavior and Food Price Inflation: The Case of Ethiopia," *ESSP II Working Paper 41*, 1–32.

Heath, C., R. P. Larrick, and G. Wu (1999): "Goals as Reference Points," *Cognitive psychology*, 38, 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, 1–54.

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, 385–398.

Hoddinott, J., G. Berhane, D. O. Gilligan, N. Kumar, and A. S. Taffesse (2012): "The Impact of Ethiopia's Productive Safety Net Programme and Related Transfers on Agricultural Productivity," *Journal of African Economies*, 21, 761–786.

Hoddinott, J., D. O. Gilligan, and A. S. Taffesse (2011): "The Impact of Ethiopia's Productive Safey Net Program on Schooling and Child Labor," in *Social Protection for Africa's Children*, March, 71–95.

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

Hussmanns, R. (2007): "Measurement of Employment, Unemployment and Underemployment: Current International Standards and Issues in their Application," *Bulletin of Labour Statistics*, 1985, 1–23.

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

Imbert, C. and J. Papp (2015): "Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee," *American Economic Journal: Applied Economics*, 7, 233–263.

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. (2007): "The Digital Provide: Information (Technology), Market Performance, and Welfare in theSouth Indian Fisheries Sector," *Quarterly Journal of Economics*, 122, 879–924.

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

JOHN, A. AND K. ORKIN (2021): "Can Simple Psychological Interventions Increase Preventive Health Investment?" *Journal of the European Economic Association*, jvab052.

KADJO, D., J. RICKER-GILBERT, T. ABDOULAYE, G. SHIVELY, AND M. N. BACO (2018): "Storage Losses, Liquidity Constraints, and Maize Storage Decisions in Benin," *Agricultural Economics (United Kingdom)*, 49, 435–454.

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, 83–106.

KEHLER, A. (2004): "When Will Ethiopia Stop Asking for Food Aid?" *Humanitarian Exchange, Overseas Development Institute*.

Kennedy, P. E. (1981): "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations (The Interpretation of Dummy Variables in Semilogarithmic Equations)," *American Economic Review*, 71, 801.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007): "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75, 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, 1133–1165.

Kremer, M., G. Rao, and F. Schilbach (2019): "Behavioral Development Economics," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 2, 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, 791–827.

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

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, 920–929.

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

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

Lewis, W. A. (1954): "Economic Development with Unlimited Supplies of Labour," *The Manchester School*, 22, 139–191.

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, 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, 709–753.

——— (2019): "Hope as Aspirations, Agency, and Pathways: Poverty Dynamics and Microfinance in Oaxaca, Mexico," in *Economics of Poverty Traps*, ed. by C. B. Barrett, M. R. Carter, and J.-P. Chavas, Chicago, IL: University of Chicago Press, chap. 4, 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*, ed. by C. B. Barrett, M. Carter, J.-P. Chavas, and M. R. Carter, Chicago: University of Chicago Press, 325–356.

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, 251–261.

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, 1149–1187.

MAGRUDER, J. R. (2018): "An Assessment of Experimental Evidence on Agricultural Technology Adoption in Developing Countries," *Annual Review of Resource Economics*, 10, 299–316.

MANI, A. AND E. RILEY (2021): "Social Networks, Role Models, Peer Effects, and Aspirations," in *Social Mobility in Developing Countries: Concepts, Methods, and Determinants*, ed. by V. Iversen, A. Krishna, and K. Sen, Oxford, United Kingdom: Oxford University Press, chap. 17, 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, 19–50.

MCCORD, A. (2013): "Public Works and Resilient Food Systems," Tech. Rep. March, Overseas Development Institute (ODI), London.

MCCORD, A. AND R. SLATER (2013): "Learning from the PSNP: The Influence of Ethiopia's Social Protection Experience in Sub-Saharan Africa and Beyon," in *Food Security, Safety Nets and Social Protection in Ethiopia*, ed. by D. Rahmato, A. Pankhurst, and J.-G. van Uffelen, Addis Ababa: Forum for Social Studies, Ethiopia, 41–65.

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, 159–217.

Mitra, S., D. Mookherjee, M. Torero, and S. Visaria (2018): "Asymmetric Information and Middleman Margins: An Experiment with Indian Potato Farmers," *The Review of Economics and Statistics*, 100, 1–13.

Mobarak, A. M. and M. Rosenzweig (2014): "Risk, Insurance and Wages in General Equilibrium," *National Bureau of Economic Research Working Paper*, 19811, 1–47.

Moneke, N. (2020): "Can Big Push Infrastructure Unlock Development? Evidence from Ethiopia," *University of Oxford Working Paper*, 1–99.

Mukherjee, S. W., L. Falcao Bergquist, M. Burke, and E. Miguel (2021): "Unlocking the Benefits of Credit through Saving," *Working Paper 29542, National Bureau of Economic Research Working Paper Series*, 1–58.

Muralidharan, K., P. Niehaus, and S. Sukhtankar (2023): "General Equilibrium Effects of (Improving) Public Employment Programs," *Econometrica*, 91, 1261–1295.

Nakasone, E. (2013): "The Role of Price Information in Agricultural Markets: Experimental Evidence from Rural Peru," *Agricultural & Applied Economics Association Conference Paper*, 1–54.

Nakasone, E., M. Torero, and B. Minten (2014): "The Power of Information: The ICT Revolution in Agricultural Development," *Annual Review of Resource Economics*, 6, 533–550.

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

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

Osborne, T. (2004): "Market News in Commodity Price Theory: Application to the Ethiopian Grain Market," *Review of Economic Studies*, 71, 133–164.

Parker, C., K. Ramdas, and N. Savva (2016): "Is IT enough? Evidence from a Natural Experiment in India's Agriculture Markets," *Management Science*, 62, 2481–2503.

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

Quisumbing, A. and Y. Yohannes (2005): "How Fair is Workfare? Gender, Public Works, and Employment in Rural Ethiopia," *World Bank Policy Research Working Papers*, 1–67.

Rahmato, D. (2013): "Food Security and Safety Nets: Assessment and Challenges," in *Food Security, Safety Nets and Social Protection in Ethiopia*, ed. by D. Rahmato, A. Pankhurst, and J.-G. van Uffelen, Addis Ababa: Forum for Social Studies, Ethiopia, 113–136.

Ravallion, M. (1991): "Market Responses to Anti-Hunger Policies: Effects on Wages, Prices, and Employment," *World Institute for Development Economic Research (WIDER) of the United Nations University Working Paper 28.*, 1–58.

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

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

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.

Rizzo, M. (2011): "Rural Wage Employment in Rwanda and Ethiopia: A Review of the Current Policy Neglect and a Framework to Begin Addressing It," *International Labour Office Working Paper*, 1–40.

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

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

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.

Santangelo, G. (2019): "Firms and Farms: The Local Effects of Farm Income on Firms' Demand," *Mimeo*.

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, 163–183.

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

Soldani, E., N. Hildebrandt, Y. Nyarko, and G. Romagnoli (2023): "Price Information, Inter-village Networks, and "Bargaining Spillovers": Experimental Evidence from Ghana," *Journal of Development Economics*, 103100.

Solon, G., S. J. Haider, and J. M. Wooldridge (2015): "What Are We Weighting For?" *Journal of Human Resources*, 50, 301–316.

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, 255–270.

Subbarao, K., C. Rodríguez-Alas, C. del Ninno, and C. Andrews (2013): *Public Works as a Safety Net: Design, Evidence, and Implementation*, Washington, DC: World Bank.

Svensson, J. and D. Yanagizawa (2009): "Getting Prices Right: The Impact of the Market Information Service in Uganda," *Journal of the European Economic Association*, 7, 435–445.

Theisen, O. M. (2012): "Climate Clashes? Weather Variability, Land Pressure, and Organized Violence in Kenya, 1989-2004," *Journal of Peace Research*, 49, 81–96.

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.

Willmott, C. J. and K. Matsuura (2015): "Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1900-2014)," *http://www.esrl.noaa.gov/psd/data/gridded/data.UDel_AirT_Precip.html*.

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, 1–38.

World Bank (2009): "Project Appraisal Report No: 48633-ET," Tech. rep., Washington, D.C.

——— (2010a): "Designing and Implementing A Rural Safety Net in a Low Income Setting: Lessons Learned from Ethiopia's Productive Safety Net Program 2005-2009." Tech. rep., Washington, D.C.

——— (2010b): "Implementation Completion and Results Report (Report No: ICR00001676)," Tech. rep., Washington, D.C.

——— (2014): "Project Appraisal Document: Productive Safety Nets Project 4. Report No: PAD1022," Tech. rep., Washington, D.C.

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.

Zimmermann, L. (2020): "Why Guarantee Employment? Evidence from a Large Indian Public-works Program," *GLO Discussion Paper*, 504, 1–64.