Nudges, Integrated Pest Management, and Livelihoods Experimental Evidence from Rural Ethiopia

Solomon Balew Deffersha

2026

Propositions

1. Customizing nudges to fit local contexts enhances their effectiveness.
   (This thesis)

2. Subsidizing Push-Pull Technology (PPT) does not ensure its initial and long-term adoption.
   (This thesis)

3. Innovation that emphasizes sustainability takes precedence over novelty.

4. The better theoretical models are those aiming for precision rather than generality.

5. Information has a greater impact on social cohesion than shared values and trust do.

6. Failures undermine resilience rather than they strengthen it.

Propositions belonging to the thesis, entitled:

Nudges, Integrated Pest Management, and Livelihoods Experimental Evidence from Rural Ethiopia

Solomon Balew Deffersha
Wageningen, 29 January 2026

Nudges, Integrated Pest Management, and Livelihoods
Experimental Evidence from Rural Ethiopia

Solomon Balew Deffersha

Thesis committee

Promotor
Prof. Dr E.H. Bulte
Professor of Development Economics
Wageningen University & Research

Co-promotor
DrM.Kassie
Head of Integrated Data and Analytics Platform International Center of Insect Physiology and Ecology (icipe), Nairobi, Kenya

Other members
Dr F.Alpizar, Wageningen University & Research
Prof. Dr R. Lensink, University of Groningen
Dr M.Jaleta, International Maize and Wheat Improvement Center (CIMMYT), Addis Abeba, Ethiopia
Dr K.Shikiku, International Livestock Research Institute (ILRI), Nairobi, Kenya

This research was conducted under the auspices of the Wageningen School of Social Sciences (WASS).

Nudges, Integrated Pest Management, and Livelihoods
Experimental Evidence from Rural Ethiopia

Solomon Balew Deffersha

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,Prof. Dr C. Kroeze,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Thursday 29January 2026
at 3:30 p.m. in the Omnia Auditorium.

Solomon Balew Deffersha
Nudges, Integrated Pest Management, and Livelihoods Experimental Evidence from Rural Ethiopia,
184 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2026)
With references, with summary in English
https://doi.org/10.18174/681005

To my late dad, and to my mum, wife, and brothers for their support in many aspects.

Table of contents

Chapter 1

Introduction

1.1. Problem statement

In Africa, insects play a complex dual role within ecosystems and human livelihoods. On one hand, they are essential as pollinators, predators, and parasites that support agriculture and biodiversity (FAO, 2013a). On the other hand, when their populations become unmanageable, insects transform into pests and vectors of disease, posing significant threats to food security, health, and economic stability across the continent. Notable examples include locusts, tsetse flies, termites, mosquitoes, and crop-damaging insects such as fall armyworm and stem borers. For instance, tsetse flies transmit sleeping sickness in humans and nagana in animals, severely affecting health and livestock productivity (Harrington & Vreysen, 2017). Termites cause extensive damage to wooden structures and crops, leading to substantial economic costs (Kellermann & van der Cingel, 2019). Crop pests like fall armyworm and stem borers can reduce yields by 25–80%, threatening the livelihoods of millions of smallholder farmers (FAO, 2020). Mosquitoes, as vectors of malaria, dengue, and yellow fever, are among the deadliest insects, contributing to widespread illness and fatalities (WHO, 2019).

Beyond direct damage, these pests impose heavy economic burdens through control efforts and productivity losses. Malaria alone accounts for over USD 12 billion annually in healthcare costs and lost productivity (WHO, 2021). Locust swarms can devastate crops and pastures, resulting in food shortages and economic losses exceeding USD 1 billion during severe invasions such as those experienced in 2020–2021 (CABI, 2020; FAO, 2021). In crop production alone, farmers generally spend between $20 and $50 per hectare annually on pesticides and pest control measures, this translates into total annual pest control costs potentially reaching into the billions of dollars.(FAO, 2018b). These examples highlight the critical importance of effective pest management strategies, which are essential not only for safeguarding health but also for economic stability.

However, reliance on chemical pesticides remains problematic. The frequent and often indiscriminate use of these chemicals has led to the emergence of insecticide-resistant pests and disease vectors, which reduces the efficacy of conventional control methods. This resistance necessitates the use of higher doses or more potent chemicals, further exacerbating environmental pollution and health risks. Additionally, chemical pesticides can harm non-target species, including beneficial insects, animals, and microorganisms, disrupting ecological balances. Concerns about contamination of soil, water sources, and food products also raise serious health issues for farmers, consumers, and communities.

In response, innovative and environmentally friendly strategies are gaining attention—such as house screening for malaria. House screening involves sealing entry points like eaves, doors, and windows to prevent mosquitoes and other disease vectors from entering indoor spaces. This method not only reduces dependency on chemical insecticides but also provides long-lasting protection against vector-borne diseases. Despite its promising potential, however, house screening remains underutilized in both policy frameworks and practical applications (Kirby et al., 2009; WHO, 2008).

Another innovative and environmentally-friendly strategy to manage insect-caused damages is push-pull technology (PPT), which protects crops against pests such as the stemborer and fall armyworm.. Introduced in 1997, PPT involves intercropping maize with Desmodium (Desmodium spp.), which naturally repels pests such as stem borers and Striga, while planting trap crops like Brachiaria around the main crop to lure and divert pests away from maize. By leveraging natural plant interactions, PPT significantly reduces dependence on chemical pesticides. In addition to pest control, this method provides several other benefits, including enhanced soil fertility through nitrogen fixation by Desmodium, weed suppression, and increased biodiversity within farming systems. These combined advantages lead to higher crop yields, improved livestock feed, greater climate resilience, and increased income opportunities for smallholder farmers. Furthermore, adopting PPT contributes to poverty alleviation by promoting environmentally sustainable and eco-friendly agricultural practices (Khan et al., 2014; Kassie et al., 2018; Midega et al., 2018).

While house screening and push-pull technology (PPT) present promising opportunities to enhance public health and boost agricultural productivity, several important knowledge gaps are preventing their widespread adoption. House screening, as a relatively new intervention, currently lacks clear evidence regarding its economic viability for malaria prevention. Additional research is needed to identify the most cost-effective screening methods and to develop strategies that ensure consistent and efficient implementation. Key factors such as the cost per malaria case averted and the motivational drivers that promote effective and sustained community participation must be better understood.

Similarly, the limited adoption of PPT highlights the need for effective dissemination strategies that overcome informational barriers, increase awareness of its benefits, and foster community engagement. Moreover, questions remain about the economic feasibility of PPT and whether farmers are willing to accept and adopt the technology. Addressing these knowledge gaps will enable stakeholders to make more informed and effective decisions about promoting, scaling, and integrating these interventions into broader health and agricultural programs.

1.2. Objectives and research questions

The main objective of this thesis is to evaluate how promotion and information interventions influence the adoption of push-pull technology (PPT) by smallholder farmers in Ethiopia. Additionally, it evaluates the effectiveness and tests how to encourage behavior that increases the effectiveness of investments in house screening against malaria.

A secondary objective is to test the impacts of interventions on malaria prevalence and agricultural productivity. Throughout this thesis, the aims are to promote effective methods for information dissemination concerning agricultural innovations, encourage wider and sustained adoption, enhance productivity, and employ strategies to amplify the impact of new technologies in Ethiopia. In particular, the research questions in the four core chapters are as follows:

Chapter 2:

- How do different types of incentives (material versus social) and their framing (gain-based framing versus loss-based framing) impact the effort of farmers in disseminating knowledge to their peers, and adoption of a new technology?
- Is there interaction between incentive type and framing in terms of incentivizing farmers to disseminate knowledge?

Chapter 3:

- How does a one-time subsidy (free provision) of a new technology, combined with additional incentives for experimentation, influence willingness to pay (WTP) for its ongoing use?

Chapter 4:

- Does implementation of Push Pull Technology (PPT) affect households' livestock productivity and livestock-based income, and what are the mechanisms linking PPT to livestock keeping?

Chapter 5:

- Does house screening reduces the prevalence of malaria and the number of malaria sick days for household members?

- Does the impact of a house screening intervention improve when screening is combined with an information intervetion that aims to promote behavioral change (a nudge)?

1.3. Background and context

In sub-Saharan Africa, the adoption of PPT remains low due to a combination of market failures, informational barriers, and lack of access to seeds, skills, and labor—particularly during planting seasons. Moreover, the need for multiple forage harvests to avoid competition with maize further complicates adoption efforts. Evidence indicates that social learning from peers is more effective in overcoming these informational barriers than extension efforts alone (Pramila & Patnam, 2013). Most studies on social networks and incentives have primarily focused on fertilizer and improved seed adoption (Bandiera & Rasul, 2006; Benyishay et al., 2018; Krishnan & Patnam, 2014).

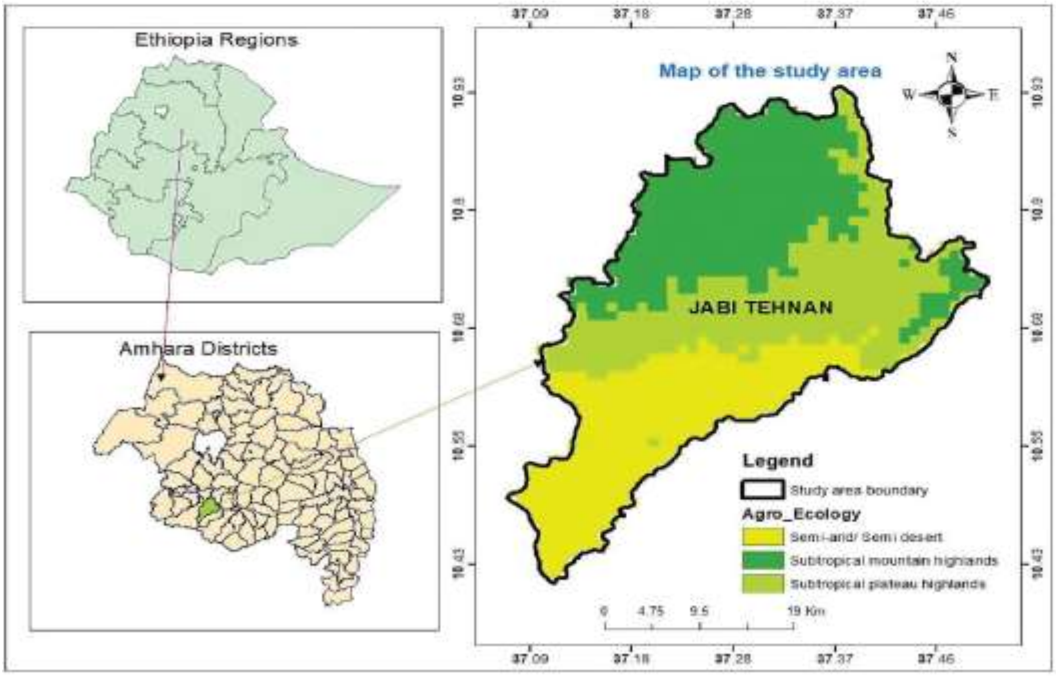
Building upon recent research (e.g., BenYishay & Mobarak, 2018; Shikiku & Bulte, 2019), I evaluated a novel extension approach in Ethiopia: a lead farmer assigned to five follower farmers, who are 'nudged' through social recognition or private incentives. The goal was to promote knowledge sharing and technology adoption by leveraging social capital and small group dynamics. This approach aimed to strengthen the connection between extension services and households via peer communication. While initial findings appeared promising, the model was also controversial. The government viewed it as a tool for mobilization (Leta et al., 2017), but critics argued it risked being exploited for political control, potentially limiting genuine technology diffusion. These concerns, however, had yet to be thoroughly tested.

In this region, malaria remains a major public health issue as well, causing significant illness and death. Vulnerable groups, such as children and pregnant women, face a higher risk of infection. Each year, hundreds of thousands of people die from malaria, with about 95% of these deaths occurring in Sub-Saharan Africa (WHO, 2021). The disease also leads to many cases of illness, resulting in outpatient visits, hospital stays, and economic losses due to decreased productivity. The reported incidence rate in Sub-Saharan Africa is roughly 300 to 400 cases per 1,000 people at risk each year, based on WHO reports and global malaria data, highlighting the heavy burden of the disease. Screening methods like rapid diagnostic tests (RDTs) and microscopy are crucial for early detection and treatment (WHO, 2020b). Although progress has been made in reducing malaria cases over the past decade, recent data show that the decline has slowed in some areas, partly because of insecticide resistance and difficulties in expanding screening efforts (WHO, 2022). Continued investment in screening, vector control, and effective treatment is essential to reduce the impact of malaria further across the continent.

The study was conducted in Jabi Tehnan district, located in northwest Ethiopia, an area comprising 38 villages and 114 sub-villages. The district faces significant development challenges, largely stemming from dependence on rain-fed agriculture, low productivity, and limited access to infrastructure and social services. Most residents rely on small-scale farming and livestock rearing, often constrained by environmental degradation, climate variability, and resource limitations—factors that contribute to persistent food insecurity and economic vulnerability. While there are opportunities to improve livelihoods through agricultural modernization, diversification, and community-based natural resource management, progress remains hindered by inadequate market access, low educational attainment, and limited off-farm income sources. Consequently, addressing these issues requires coordinated, integrated efforts across multiple sectors.

Both health and PPT technologies introduced during the study were new to the community and involved high initial costs, especially given limited access to credit in the region. As a result, the risks and expenses associated with early adoption served as barriers, leading to initial reliance on free or subsidized distributions. However, the long-term impact of such subsidies is complex; they can either encourage sustained adoption or discourage future participation (Dupas, 2014; Fisher et al., 2017). Therefore, understanding how to balance immediate support with sustainable adoption remains a key challenge.

Figure 1: Map of the Study area



Jabi Tehnan District is located in the West Gojjam Zone of the Amhara Regional State, Northwestern Ethiopia. The map illustrates the district's diverse agroecological zones, depicted as follows: yellow indicates semi-arid low land areas, green represents subtropical dry mountain highlands, and light green denotes subtropical plateau highlands.

1.4. Push-Pull Technology (PPT)

PPT (Push-Pull Technology) is a pioneering method for pest management, designed to protect cereal crops from damaging pests such as stem borers and striga weeds—problems that pose significant threats to smallholder farmers throughout sub-Saharan Africa. This innovative approach was developed through a collaborative effort involving the International Centre of Insect Physiology and Ecology (icipe) and other practitioners, this innovative system harnesses the natural interactions between plants and pests. Its essence lies in "pushing" pests away from the main crop using repellent plants such as Desmodium, while simultaneously "pulling" them toward trap crops like Brachiaria or Napier grass, which lure and confine pests for easier management (Khan et al., 2010). This harmonious blend of biological strategies offers a sustainable alternative to chemical pesticides, fostering ecological balance and resilience.

In practical terms, farmers cultivate maize alongside Desmodium, which emits fragrant odors that deter pests like stem borers and suppress striga weeds. Surrounding the main crop, trap crops are planted to attract pests away from the vital maize plants. These pests then gather in specific trap areas, where they can be managed through simple methods such as manual removal or biological controls. Beyond its pest-fighting prowess, PPT enriches the soil, provides fodder for livestock, and cuts costs by reducing reliance on expensive chemicals—an all-encompassing approach that champions environmentally friendly agriculture (Khan et al., 2010).

Ultimately, the core goal of PPT is to boost crop yields, enhance food security, and promote sustainable farming among vulnerable smallholder farmers. Its success stories span across Kenya, Ethiopia, Uganda, and beyond, where it offers an affordable, eco-conscious alternative to chemical solutions. By fostering biodiversity, improving ecological health, and strengthening climate resilience, PPT exemplifies a forward-thinking model of integrated farming—one that nurtures both the land and the livelihoods of those who depend on it (Khan et al., 2010; Vurro et al., 2010).

1.5. House screening (HS)

HS is an effective and sustainable malaria prevention strategy that involves installing physical barriers, such as fine mesh screens on windows and doors, to prevent mosquitoes—primarily Anopheles species—from entering homes. This intervention reduces human-mosquito contact, thereby lowering the risk of malaria transmission. Evidence shows that house screening can significantly decrease indoor mosquito populations and malaria cases, especially when combined with other control measures like insecticide-treated nets (ITNs) and indoor residual spraying (IRS) (Lindsay et al., 2002; Kelly et al., 2010). Made from durable, chemical-free materials, screens provide ongoing protection while enhancing indoor comfort by minimizing insect nuisances. When properly implemented and maintained, house screening offers a cost-effective, long-term solution that complements existing malaria control efforts (Killeen et al., 2014).

Behavioral change can play a crucial role in maximizing the benefits of house screening for disease prevention. Simple practices such as consistently closing windows and doors before mosquitoes become active, and reducing outdoor activities during peak mosquito hours, can significantly enhance the protective effects of screening. When residents adopt these behaviors, the likelihood of mosquitoes entering the home diminishes, further reducing the risk of vector-borne diseases like malaria and dengue. Together with physical barriers, these behavioral modifications can create a comprehensive approach to mosquito control, fostering healthier living environments and amplifying the overall impact of disease prevention efforts.

1.6. Theory

In economics, an incentive is defined as something that encourages individuals or organizations to act in specific ways through rewards or penalties, thereby influencing their choices (Mankiw, 2014). These incentives serve to align personal interests with broader economic goals, guiding decision-making and resource allocation. To fully understand their impact, it is essential to distinguish between material incentives—such as cash or goods—and social or status-based motivations. While material incentives offer tangible, immediate benefits, social incentives appeal to desires for approval, recognition, and positive self-image. Consequently, social incentives foster intrinsic motivation and promote sustained engagement (Deci & Ryan, 1985; Bénabou & Tirole, 2003).

Another fundamental concept in behavioral economics is loss aversion, which states that individuals tend to feel the pain of losses more acutely than the pleasure of equivalent gains (Kahneman & Tversky, 1979). This bias significantly influences decision-making, as people are often more motivated to avoid losses than to pursue gains. Therefore, interventions can be more effective if they frame choices in ways that minimize perceived losses or emphasize potential gains, thereby increasing the likelihood of desired behaviors (Thaler & Sunstein, 2008).

Moving forward, the role of subsidies—particularly one-time subsidies—is also crucial. A one-time subsidy refers to a single-payment financial assistance provided to individuals, businesses, or organizations to promote specific behaviors, support projects, or alleviate financial burdens (Mankiw, 2014). For example, providing a subsidy for adopting new agricultural technologies can reduce initial risks, thereby encouraging experimentation and skill development. Moreover, as farmers gain experience, they are more likely to adopt and efficiently use the technology even after the subsidy ends, which fosters long-term behavioral change (Feder et al., 1985; Bandiera & Rasul, 2006).

However, it is important to recognize that subsidies can also shape perceptions by establishing reference points. When farmers receive substantial support, they may overestimate the benefits or underestimate the costs associated with new technologies. This can lead to reluctance to pay full prices later or increased dependence on aid—potentially hindering sustained adoption (Tversky & Kahneman, 1974; Cialdini, 2001). Empirical evidence presents mixed effects: some studies indicate that beneficiaries anchored to low prices resist paying full costs, while others suggest that understanding the benefits increases willingness to pay later (Dupas, 2014).

Furthermore, for experience-based technologies, learning plays a vital role in mitigating reference effects. Nonetheless, concerns about improper use or dependency on subsidies persist. The impact of higher prices on adoption varies across studies; some find that increased costs improve usage, while others observe no significant effect (Ashraf et al., 2010; Cohen & Dupas, 2010; Fisher et al., 2017). Therefore effective subsidy design should aim to promote learning and adoption while minimizing dependency and negative perception effects.

1.7. Methods

To investigate the research questions outlined in section 1.2, this thesis employs a combination of methodologies, including randomized controlled trials (field experiments), experimental auctions, economic modeling and econometric analysis. The following sections will provide a detailed overview of each method, discussing their strengths, and potential limitations.

1.7.1. Randomized Controlled Trials (RCTs)

A Randomized Controlled Trial (RCT) is a rigorous research method used to assess the effectiveness of an intervention by randomly assigning participants to either a treatment or control group. This randomization minimizes selection bias and ensures that observed differences are due to the intervention itself. Recognized as the gold standard in experimental research, RCTs provide strong causal evidence about the effects of specific variables, making them essential in fields such as medicine, public health, and social sciences (Hariton & Locascio, 2018). By establishing causal relationships, RCTs not only determine what works but also challenge assumptions and identify ineffective approaches, thereby guiding effective decision-making (Duflo, 2006).

A solid theoretical foundation is essential for designing effective RCTs, guiding researchers toward relevant questions and specific, testable predictions (Duflo, 2006). While RCTs have a long-standing history in agricultural and biomedical research, their application in economics is relatively recent. In recent years, the use of randomized field experiments has expanded significantly within development economics (Athey and Imbens, 2017). By leveraging randomization in real-world environments, researchers can control for confounding factors and determine causal effects on individuals or groups (Baldassarri and Abascal, 2017).

For this thesis I conducted two RCTs, one on HS and another on PPT, with randomization at the sub-village level. The trials involved a total of approximately 5,470 lead farmers (LFs), who are key adopters of new technologies, with a random sample of 754 LFs selected for the PPT study. These lead farmers were randomly assigned to four groups: a control group receiving no intervention(C), a group receiving training, seeds (brachiaria and desmodium), and guidance (T1), a group receiving the full package plus private incentives (T2), and a group receiving the full package plus social recognition (T3). This design facilitated comparison of private versus social incentives in influencing knowledge transfer and adoption rates. For treatments 2 and 3, incentives were split into “gain” and “loss” versions

For the HS intervention, a sample of 914 households was randomly selected. These households were then randomly assigned to four groups: one served as the control group and received bed nets, while the others received HS and bed nets (T1). Among the intervention groups, one received HS and bed nets along with health messages emphasizing gains (T2), and another received HS bed nets with health messages emphasizing losses (T3). This experimental design aimed to evaluate whether additional informational interventions affect the effectiveness of house screening technology, as well as to compare the impact of loss-framed versus gain-framed health messages on health behaviors and outcomes.

While RCTs are widely regarded as the gold standard for establishing causality, they also present notable limitations, including ethical considerations (Shadish, Cook, & Campbell, 2002; Meinert, 2012). In my studies, ethical concerns arose particularly in the HS intervention, where withholding treatment from the control group posed a dilemma. To address this, bed nets were distributed to control households to ensure they received some benefits.

Data collection took place in multiple waves. For Chapter 2, it took place in two waves from lead farmers and followers, with baseline data collected between June and August 2018, and follow-up data collected in November 2020. The data for Chapters 3, 4, and 5 were collected in three waves, exclusively from farmer leaders within the district. Similar to Chapter 2, the baseline data were obtained from June to August 2018. Subsequent follow-up data collection occurred in November 2020, marking the midline, and endline in February 2022.

1.7.2. Experimental auctions

Experimental auctions are a valuable tool in valuation research, enabling accurate assessment of how individuals value non-market goods—items not typically exchanged in traditional markets. They are useful for cost-benefit analyses of public projects, evaluating technological impacts, predicting market success, and understanding consumer and citizen behavior. By combining revealed preferences (choices in real settings) and stated preferences (expressed willingness to pay in hypothetical scenarios), experimental auctions offer more reliable, incentive-compatible valuation estimates, minimizing biases common in other methods (Lusk & Shogren, 2007).

Among various auction formats, the Becker-DeGroot-Marschak (BDM) auction is particularly prominent. It offers several key advantages, foremost among them its incentive compatibility, which means it encourages participants to reveal their true valuations since honest bidding is their optimal strategy (Becker, DeGroot, & Marschak, 1964). Its simplicity and transparency make it easy for participants to understand and execute, reducing errors and confusion. Additionally, the BDM effectively captures private valuations that are otherwise difficult to observe, making it highly versatile across various research contexts. Moreover, the mechanism minimizes strategic manipulation, ensuring that the data collected accurately reflects genuine preferences. Overall, these qualities enhance the reliability and validity of valuation measurements obtained through the BDM method.

In Chapter 3, I used the BDM auction to assess farmers' WTP for PPT. Farmers submit bids compared to a randomly selected strike price; if their bid exceeds the strike, they can purchase at that price, encouraging honest valuation.

Although the BDM mechanism is often criticized for certain limitations—such as the potential for misunderstandings leading to biased bids among novices, as well as its susceptibility to framing effects and hypothetical bias (Lusk & Norwood, 2016)—recent evidence from Burchardi et al. (2021) suggests that participants can largely comprehend the process. In their study, 94% of participants answered comprehension questions correctly, and 86% bid optimally. These findings indicate that, under certain conditions, participants are capable of effectively engaging with and accurately navigating the mechanism. To further reduce understanding bias, I provided extensive training to enumerators on the BDM procedure and rules, and ensured they trained participants prior to bidding each year.

1.7.3. Economic modeling

Economic modeling serves as a tool for representing, analyzing, and predicting the behavior of economic systems. It employs simplified mathematical equations or computational algorithms to create models that capture the key elements and interactions within an economy. These models are particularly useful for understanding the complex relationships among variables such as prices, income levels, and employment rates, which are often interdependent and dynamic. By providing a structured framework, economic models facilitate scenario analysis, allowing policymakers and researchers to explore the potential impacts of different policies, market conditions, or external shocks. They also aid in forecasting future economic trends and informing strategic decisions across various sectors.

Despite their usefulness, economic models are inherently based on assumptions that aim to simplify reality. These simplifications can sometimes overlook certain nuances or complexities of actual economies, which may lead to inaccuracies or oversights. Additionally, the reliability of model outputs heavily depends on the quality of the underlying data and the specific choices made in model design and parameters. Therefore, while models are powerful tools for gaining insights and guiding policy, they are not perfect representations of real-world economic systems.

Nevertheless, the importance of economic modeling remains unquestioned. It provides a systematic approach to understanding economic dynamics, testing hypotheses, and evaluating potential policy outcomes. The ongoing process of constructing, testing, and refining models helps economists better grasp how economies function and evolve. As Varian (1992) emphasizes, despite their limitations, economic models are indispensable for informing policies, addressing market failures, and advancing our overall knowledge of economic behavior and interactions.

In Chapter 3, I employ economic modeling to explore how treatment interventions impact farmers' demand for the new agricultural technology. The analysis estimates demand curves for each group, depicting the relationship between the share of farmers willing to pay and the price in ETB per hectare. Using survey data, the model assesses how each treatment influences this relationship, revealing differences in willingness to pay and demand elasticity across the various groups.

1.7.4. Econometric analysis

Econometrics is a specialized field within economics that leverages statistical and mathematical techniques to analyze economic data and evaluate economic theories. It focuses on constructing and applying quantitative models to uncover the relationships among various economic variables, estimate their impacts, and generate reliable forecasts. By integrating economic principles with empirical data, econometrics enables researchers and policymakers to assess policy effects, identify causal linkages, and interpret complex economic phenomena with enhanced accuracy and confidence (Greene, 2018).

A key component of econometrics is the use of techniques such as regression analysis, which helps in examining relationships between variables and testing specific hypotheses. Additionally, econometrics is widely used for forecasting future economic and financial trends based on historical data, providing valuable insights for decision-making.

In this thesis, I utilized field experiments alongside econometric analysis to rigorously test hypotheses. This combined approach ensures that the findings are grounded in empirical evidence and relevant to real-world economic behavior.

1.8. Outline

The rest of this thesis is structured as follows. In Chapter 2, I examined how incentives given to leader farmers, alongside behavioral nudges, influence their efforts to share information and promote the adoption of push-pull technology (PPT) among follower farmers. The primary goal of this intervention was to address challenges related to information flow and dissemination by motivating lead farmers to pass on knowledge about the new technology to their followers. I found that both private rewards and social prestige incentives significantly encouraged lead farmers to increase their efforts in spreading information, which facilitated broader knowledge sharing within the group. Moreover, social prestige incentives, compared to private rewards, proved more effective in motivating lead farmers to organize training events and to reach a larger portion of their group members.

Building on these findings, I also observed that both loss-framed and gain-framed incentives effectively motivated lead farmers to intensify their information dissemination efforts. Notably, loss-framed incentives—where the perceived damage of losing prestige outweighs the potential benefits—had a stronger impact than gain-framed incentives in encouraging lead farmers to organize training sessions for their followers. This suggests that framing incentives as potential losses to social standing can be a powerful motivator.

Furthermore, I explore the interaction between incentives and loss-framed messages in this chapter. Specifically, I document that loss-framed incentives enhanced the effectiveness of social prestige incentives, with the threat of shame serving as a potent motivator. Conversely, I find that loss-framed messages only modestly increased effort when paired with private rewards, possibly because the credibility of threats to reclaim material rewards if performance falls short is limited. Interestingly, although group leaders increased their efforts in disseminating information, this did not translate into greater experimentation or adoption among follower farmers. This indicates that barriers to adoption extend beyond mere information deficits, implying that simply providing access to information is insufficient. It is also important to note that the rewards depended on successful information diffusion by leaders, not on the followers' experimentation or adoption. Therefore, while leaders responded to incentives, they did not exhibit additional behavioral changes beyond their immediate efforts to disseminate information.

This chapter advances the existing literature in several key ways: (i) by introducing a novel social prestige incentive that relies solely on symbolic recognition rather than material rewards; (ii) by utilizing pre-existing semi-formal local institutions for knowledge dissemination instead of creating artificial institutions for the experiment; and (iii) by combining private or social prestige incentives with loss-framed incentives, and exploring their interaction through the perspectives of loss aversion and shame. Significantly, this study is among the first to apply loss-framed incentives in a real-world field experiment within a low-income country context, where leveraging loss aversion is notably more complex than in controlled laboratory settings. Finally, from a methodological standpoint, (iv) in addition to evaluating the impact of loss-framed incentives, I also assess their feasibility in environments characterized by imperfect contracting and long production cycles.

In Chapter 3, I examine the effects of providing farmers with a one-time full subsidy for PPT on their future willingness to pay (WTP). I find that a single free provision of the PPT package produced mixed impacts on subsequent WTP. For farmers who did not experiment with PPT, being assigned to receive the package for free appeared to generate anchoring effects that negatively influenced their WTP afterward. This may stem from the fact that these farmers lacked practical experience with the technology, which limited their understanding of its benefits. Instead, simply being entitled to receive the package for free during subsidy programs might have caused them to base their expectations on initial impressions. Such anchoring could have distorted their perceptions and reduced their motivation to invest in the technology later on.

Conversely, I also identified that farmers who did experiment with PPT and gained hands-on experience were more likely to benefit from the subsidy through enhanced learning and belief updating regarding its advantages. While the anchoring and learning effects operate in opposite directions, the latter generally dominates among farmers who actively engaged with the technology. As a result, the overall effect of a one-time full subsidy on future WTP tends to be positive for those farmers who adopted the technology and are likely to continue using it.

These insights underscore the importance of caution when promoting new products via full subsidies, particularly when additional costs accompany adoption. Without complementary interventions that encourage experimentation, temporary full subsidies could inadvertently suppress future demand. Regarding IPM, which provides private benefits to adopters and some public benefits through reduced pest spread, the case for sustained subsidies remains uncertain. Therefore, I believe that evaluating the long-term implications after the subsidy ends is crucial. More than three years post-intervention, I still observe sizable coefficients consistent with anchoring effects; however, these effects are no longer statistically significant at conventional levels, suggesting that the anchoring effect diminishes over time or is overshadowed by other sources of information.

This chapter also makes a novel contribution to the existing literature by introducing an experimental design capable of providing an unbiased estimate of the anchoring and learning effects associated with full subsidies for new technologies. The two-stage approach relies on fewer assumptions than previous studies aiming to distinguish and quantify these effects. For example, Dupas (2014) estimated the overall impact of subsidized bed nets for malaria prevention and then attempted to differentiate between anchoring and learning effects by imposing additional assumptions based on a simplified experience-good model. Similarly, Shukla et al. (2022) conducted an experiment involving hermetic storage bags, randomly assigning farmers to three treatment groups: free provision, a flat-rate price, and a random strike price in a BDM auction. They then compared willingness to pay (WTP) for additional bags in the second stage. However, because farmers faced different prices across treatment arms in the first stage, adoption rates varied, and it is possible that different types of farmers purchased the bags in each arm. Consequently, variations in WTP during the second stage could be driven not only by anchoring effects but also by selection effects. Interpreting differences in bids across arms as causal effects of the zero-price condition thus requires additional assumptions, complicating the analysis. The used design minimizes these assumptions but limits causal inference to a subset of respondents—namely, the "complying farmers"—who are induced to experiment with the new technology due to the extra incentive provided.

Beyond these methodological advancements, I emphasize a focus on a specific type of innovation—Integrated Pest Management (IPM) technologies—that are labor-intensive and require complementary inputs for successful implementation. This makes the IPM package comparable to many other agricultural innovations, such as improved seeds or conservation agriculture, but distinct from innovations like malaria bed nets, PICS bags, or solar lamps, which have been more extensively studied in the literature. As research in this area expands, future studies are expected to examine the robustness of anchoring and learning effects across different categories of innovations—particularly those with and without external inputs or benefits.

In Chapter 4, I examined the impacts of PPT intervention on farmers' livestock productivity and household income. I find that for farmers who experimented with the technology, the intervention significantly boosted both livestock productivity and household income. Specifically, it helps farmers increase herd sizes, milk production, income from feed and livestock sales. An important observation is that the positive effects of the intervention on productivity and income grow over time. Overall, the intervention for the adopting household raises income by about 2,356 ETB (approximately USD 43) at midline and by 3,323 ETB (approximately USD 60) at endline. This likely occurs because, as the PPT plots mature, forage yields from the push-pull system improve, enabling households to expand their herd sizes. Larger herds then translate into higher income from livestock. Additionally, as forage yields increase with plot maturity, households are able to produce surplus feed, which can be sold for additional income. These findings suggest that the benefits of the PPT intervention accumulate and become more pronounced as the technology matures.

This chapter advances the existing literature on the economic impacts of PPT by employing experimental data to establish causality, thereby surpassing the limitations of previous correlational studies. Unlike prior research, my approach provides clear evidence of a direct causal relationship between PPT adoption and improvements in farm household income and productivity. Beyond this methodological breakthrough, the chapter also differentiates between short-term and long-term effects, illuminating how the timing of PPT adoption influences farmers' returns. This nuanced analysis fills a critical gap in the literature and offers valuable insights into the evolving economic benefits that arise from sustained adoption of the technology over time.

In Chapter 5, I analyze the health and economic impacts of house screening(HS) as an intervention to combat malaria. The chapter also examines how strategic use of information and thoughtful design can influence household decisions to adopt house screening, thereby maximizing its benefits. My findings show that house screening is effective in reducing malaria transmission. Quantitatively, households that received screening experienced a 73% reduction in malaria cases during the first year—equivalent to 0.70 fewer episodes per household—highlighting the immediate health benefits. Over a longer period, two years after installation, the reduction remains significant at 33% or 0.32 fewer episodes per household, demonstrating the durability of the intervention's effects. At the household level, these reductions translate into tangible wellbeing improvements. During peak malaria season, households with screening report approximately 6.8 fewer sick days in the short term and about 4.4 fewer days in the long term. These decreases not only improve individual health but also enhance productivity and household income stability.

Building on these results, I explore how coupling house screening with targeted health messaging can further enhance outcomes. The analysis indicates that when combined, the effects are even more pronounced. For example, one year after the intervention, households that only received screening saw a 66% reduction in malaria episodes—about 0.63 fewer episodes. When paired with gain-framed messages, the reduction was nearly identical (−0.64 episodes), but loss-framed messages led to a larger decrease of 0.86 episodes. This suggests that emphasizing potential losses from malaria can motivate stronger protective behaviors.

In terms of sick days, combining house screening with health messaging produces notable improvements. Screening alone reduces sick days by about 5.5 days; with gain-framed messaging, this increases to 6.2 days, and with loss-framed messaging, the reduction reaches approximately 9.3 days. The stronger response to loss framing indicates that households respond more vigorously to messages emphasizing risks and potential losses, although differences between message types are not always statistically significant.

An important observation is that the effects of combined interventions tend to decline over time, reflecting diminishing returns as households may revert to previous behaviors or enthusiasm wanes. Nonetheless, the impact of loss-framed messaging persists even after two years, highlighting its potential for long-term influence.

From a policy perspective, these results are highly relevant. Our model suggests that integrating house screening with loss-framed health messaging can significantly improve social welfare by reducing malaria and its economic burdens. While house screening alone offers substantial benefits, pairing it with behavioral insights amplifies these effects, leading to better health outcomes and increased household resilience. Overall, the study underscores the importance of combining physical interventions with targeted messaging to maximize impact.

This chapter also makes several contributions to the literature. First, it examines how a technical intervention (house screening) interacts with information treatments (health messaging), comparing gain-and loss-framed messages. Despite their low cost, these messages can substantially enhance the effectiveness of the more expensive physical intervention. We find early evidence that message framing influences behavioral responses, aligning with research on how framing affects health decisions. Second, we provide valuable insights into the long-term effects of behavioral interventions—a dimension often missing in studies focused on short-term lab experiments. A single framing message at the start of the project had lasting effects across multiple malaria seasons, suggesting habits or learning processes among households. However, these effects gradually weaken over time, possibly due to factors like screen depreciation. Finally, this study is among the first to evaluate the economic impact of house screening at the household level, moving beyond traditional measures like mosquito density or malaria incidence to assess broader economic benefits.

In Chapter 6, I will summarize the key findings, lessons learned, and policy implications of the research. It will emphasize the effectiveness of social prestige and loss-framed incentives in motivating information sharing and driving behavioral changes to support house screening. The chapter will discuss how subsidy design affects sustained demand and highlights the importance of encouraging experimentation. It will present evidence of long-term economic benefits from push-pull systems and health improvements resulting from house screening combined with behavioral messaging. Key lessons include leveraging social and emotional motivators, fostering active engagement, and developing context-specific interventions. Additionally, the chapter will suggest avenues for future research on long-term sustainability, transferability, and the behavioral factors underlying these outcomes.

1.9 References

Ashraf, N., J. Berry, and J. Shapiro (2010). Can Higher Prices Stimulate Product Use? Evidence From a Field Experiment in Zambia, *American Economic Review*, 100 (5), 2383–2413.

Athey, S., and Imbens, G. W. (2017). The State of Applied Econometrics: Causality and Policy Evaluation. *Journal of Economic Perspectives*, 31(2),3– 32 https://doi.org/10.1257/jep.31.2.3

Baldassarri, D., and Abascal, M. (2017). Field Experiments Across the Social Sciences. *Annual Review of Sociology*, 43(1), 41–73. https://doi.org/10.1146/annurev-soc-073014-112445

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring Utility by a Single-Response Sequential Method. *Behavioral Science*, 9(3), 226–232.

Bénabou, R., & Tirole, J. (2003). Incentives and Prosocial Behavior. *American Economic Review*, 93(5), 165-168.

BenYishay, A., and Mobarak, A.M. (2014). Social learning and communication. Working Paper 20139, National Bureau of Economic Research (NBER) Working Paper Series, Cambridge, MA, USA.

BenYishay, A., and Mobarak, A.M., 2018. Social Learning and Incentives for Experimentation and Communication. *Review of Economic Studies*, 0: 1–34.

Burchardi, K. B., de Quidt, J., Gulesci, S., & Lerva, B. (2021). Testing willingness to pay elicitation mechanisms in the field: Evidence from Uganda. *Journal of Development Economics*, 154, 102774. https://doi.org/10.1016/j.jdeveco.2021.102774

CABI. (2020). *Desert Locust Bulletin*. Center for Agriculture and Bioscience International.

Cialdini, R. B. (2001). *Influence: Science and Practice.* Allyn & Bacon.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media.

Duflo, E. (2006). *Field Experiments in Development Economics*. https://www.semanticscholar.org/paper/Field-Experiments-in-Development- Economics-1-Duflo/7bf153f4e0c7480572a01ce00e8c494b4732f9ec

Dupas, P. (2014). Djort-run subsidies and long-run adoption of new health products: Evidence from a field experiment. *Econometrica*, 82(1), 197–228.

FAO. (2013). *Insect Pests of Crops in Africa*. FAO Plant Production and Protection Paper 230. Rome: FAO.

FAO. (2013a). *The Role of Insects in Biological Control*. FAO Plant Production and Protection Paper 218. Rome: FAO.

FAO. (2018a). *The State of Food Security and Nutrition in the World 2018*. Food and Agriculture Organization of the United Nations.

FAO. (2018b). *The State of Food and Agriculture 2018: Migration, Agriculture and Rural Development.* Food and Agriculture Organization of the United Nations.

FAO. (2020). *The Fall Armyworm: Impact and Management*. Available at: https://www.fao.org/in-action/fall-armyworm/impact/en/

FAO. (2021). *Locusts and their control in Africa*. Available at: https://www.fao.org/locusts/en/

Feder, G., Just, R. E., & Zilberman, D. (1985). Adoption of agricultural innovations: A review. *Economic Development and Cultural Change*, 33(2), 255-298.

Fisher, M., Smith, J., & Lee, R. (2017). Evaluating subsidy impacts on agricultural productivity. *Journal of Development Economics*, 125, 1-15.

Greene, W. H. (2018). *Econometric Analysis* (8th ed.). Pearson.

Hariton, E., & Locascio, J. J. (2018). Randomized controlled trials — the gold standard for effectiveness research: Study design: How to choose the appropriate design. *American Journal of Obstetrics and Gynecology*, 218(4), 379–389. https://doi.org/10.1016/j.ajog.2017.01.031

Harrington, R., & Vreysen, M. J. (2017). Tsetse and trypanosomiasis control: the role of sterile insect technique. *Journal of Insects and Society*, 64(2), 245–260. https://doi.org/10.1007/s13752-017-0267-1

Kassie, M., Stage, J., Diiro, G., Muriithi, B., Muricho, G., Ledermann, S.T., Pittchar, P., Midega, C., and Khan, Z. (2018). Push–Pull farming system in Kenya: Implications for economic and social welfare. *Land Use Policy*, 77, 186–198.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263-291.

Kellermann, J., & van der Cingel, N. (2019). Termites as pests and beneficial insects. *Insects*, 10(11), 400. https://doi.org/10.3390/insects10110400

Kelly, A. H., et al. (2010). Effectiveness of house screening and insecticide-treated nets for malaria control in The Gambia. *Malaria Journal*, 9, 262.

Khan, Z. R., van den Berg, J., & Kropff, M. J. (2010). *Push-pull technology: a new approach to integrated pest management*. African Journal of Agricultural Research, 5(16), 1888–1894.

Killeen, G. F., et al. (2014). A research agenda for the next decade of malaria control. *PLoS Medicine*, 11(4), e1001644.

Killeen, G.F., Govella, N.J., Mlacha, Y.P. and Chaki, P.P., (2019). Suppression of malaria vector densities and human infection prevalence associated with scale-up of mosquito-proofed housing in Dar es Salaam, Tanzania: re-analysis of an observational series of parasitological and entomological surveys. *Lancet Planet. Heal*., 3(3), e132–e143. Available at: https://www.sciencedirect.com/science/article/pii/S254251961930035X [Accessed April 26, 2019].

Kirby, M. J., Ameh, D., Bottomley, C., Green, C., Jawara, M., Milligan, P. J., Snell, P. C., Conway, D. J., & Lindsay, S. W. (2009). Effect of two different house screening interventions on exposure to malaria vectors and on anaemia in children in The Gambia: A randomized controlled trial. *PLoS Medicine*, 6(11), e1000165. https://doi.org/10.1371/journal.pmed.1000165

Krishnan, P., & Patnam, M. (2013). Neighbors and extension agents in Ethiopia: Who matters more for technology adoption? *American Journal of Agricultural Economics*, 95(4).

Leta, T., Mengistu, B., & Alemu, T. (2017). Facilitating social learning for improved crop management in Ethiopia. *Development Studies Review*, 12(3), 210-225.

Lindsay, S.W., M., E.P. and Charlwood, J.D. (2002). Reducing malaria by mosquito-proofing houses. *Trends Parasitol*., 18(11), pp.510–514.

Lusk, J. L., and Shogren, J. F. (2007). *Experimental Auctions: Methods and Applications in Economic and Marketing Research*. Cambridge University Press.

Lusk, J. L., & Norwood, F. B. (2016). *Commodity Choice and Commodity Valuation*. In *Experimental Auctions: Genetic Algorithms and Market-Based Mechanisms* (pp. 103-124).

Mankiw, N. G. (2014). *Principles of Economics* (7th ed.). Cengage Learning.

Midega, C.A.O., Pittchar, J., Pickett, J.A., Hailu, G., Khan, Z. (2018). A climate-adapted push-pull system effectively controls fall armyworm, *Spodoptera frugiperda* (J E Smith), in maize in East Africa. *Crop Protection*, 105, 10-15.

Midingoy, S-K. G., Kassie, M., Muriithi, B., Gracious, D., and Ekesi, S. (2018). Do farmers and the environment benefit from adopting integrated pest management practices? Evidence from Kenya. *Journal of Agricultural Economics*, doi: 10.1111/1477-9552.12306.

Pramila, R., & Patnam, M. (2013). Impact of agricultural interventions on rural livelihoods in India. *Journal of Rural Development Studies*, 29(4), 345-360.

Ranson, H., and Lissenden, N. (2016). Insecticide resistance in african anopheles mosquitoes: A worsening situation that needs urgent action to maintain malaria control. *Trends in Parasitology*, 32, 187–196. doi:10.1016/j.pt.2015.11.010

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin.

Shikiku, K. M., & Bulte, E. H. (2019. The Role of Social Networks in Agriculture: Evidence from Kenya. *World Development*, 117, 132-146.

Shukla, Pallavi, Hemant Pullabhotla, and Kathy Baylis. (2022). Trouble with Zero: The Limits of Subsidizing Technology Adoption. *Journal of Development Economics*, 158: 102920. https://doi.org/10.1016/j.jdeveco.2022.102920

Sternberg, E. D., & Thomas, M. B. (2017). Insights from agriculture for the management of insecticide resistance in disease vectors. Department of Entomology and Center for Infectious Disease Dynamics, The Pennsylvania State University, University Park, PA, USA.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Tusting, L.S, Thwing, J., Sinclair, D., Fillinger, U., Gimnig, J., Bonner, K.E, Bottomley, C., Lindsay, S.W. (2015) Mosquito larval source management for controlling malaria. *Cochrane Database Syst Rev*. ; 8: CD008923. doi:10.1002/14651858.CD008923.pub2.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Varian, H. R. (1992). *Microeconomic Analysis* (3rd ed.). W. W. Norton & Company.

Vurro, M., Rizzello, C. G., & Cattivelli, L. (2010). *Sustainable pest management in cereal crops: the push-pull approach*. Pest Management Science, 66(12), 1411-1413.

World Health Organization. (2008). WHO position statement on integrated vector management. Retrieved from http://apps.who.int/iris/bitstream/10665/69745/1/WHO_HTM_NTD_VEM_2008.2_eng.pdf

World Health Organization.(2019). *World Malaria Report 2019*. WHO. Available at: https://www.who.int/publications/i/item/9789241565721

World Health Organization.(2020a). Malaria Policy Advisory Committee (MPAC) meeting report.

World Health Organization. (2020b). *World Malaria Report 2020*

[https://www.who.int/publications/i/item/9789240015791]

(https://www.who.int/publications/i/item/9789240015791)

World Health Organization. (2021). *Malaria Fact Sheet.*

https://www.who.int/news-room/fact-sheets/detail/malaria

World Health Organization. (2022). *World Malaria Report 2022*.

[https://www.who.int/publications/i/item/9789240040496

](https://www.who.int/publications/i/item/9789240040496)

Chapter 2

Incentivizing and Nudging Farmers to Spread Information: Experimental Evidence from Ethiopia

Abstract: Information does not flow freely through social networks. We use an experiment to study knowledge diffusion about an innovation (integrated pest management, IPM) in farmer groups in Ethiopia. Group leaders are incentivized to share knowledge with members through the conditional provision of material- or social prestige rewards. We combine incentives with loss-framed messaging to leverage loss aversion. Incentives increase diffusion effort, and combining incentives with loss-framed messaging increases effort further. However, the treatments failed to induce follower farmers to experiment with IPM. We also document that re-claiming material rewards is difficult after a long delay, attenuating the effectiveness of the loss frame.

Keywords: agricultural extension, reference-dependent utility, social prestige vs material rewards, integrated pest management, clawback.

JEL codes: O13, O33, Q12,

Publication status: Balew, S., Bulte, E., Abro, Z., & Kassie, M. (2023). "Incentivizing and nudging farmers to spread information: Experimental evidence from Ethiopia." *American Journal of Agricultural Economics*, 105, 994-1010.

2.1. Introuction

While widespread adoption of modern inputs and practices by smallholders is a precondition for meeting the Sustainable Development Goals related to food security and poverty in Africa, the uptake of such productivity-enhancing inputs remains incomplete (Pamuk et al. 2014). One well-known reason for this is lack of information (e.g., Foster and Rosenzweig 1995, 2010).[1] If farmers are not informed about new technologies, do not know how to implement them, or are uncertain about their benefits, adoption will fail. Most governments invest in extension systems to promote the diffusion of information. A typical approach to extension involves targeting and training so-called lead (or model) farmers—often recruited from the subset of successful, entrepreneurial, and relatively well-educated farmers. After receiving some training, these lead farmers are supposed to share information with their co-villagers and peers.

This approach has produced disappointing outcomes in sub-Saharan Africa (De Janvry et al. 2016; Pamuk et al. 2014). The main reason is presumably that most extension systems are wildliy underfunded. Other possible reasons include the absence of clear incentives for public sector extension workers, resulting in the application of low levels of effort (Dar et al. 2021), or selection of the “wrong” subsample of farmers as lead farmers. While it is convenient for extension experts to work with advanced farmers, ordinary farmers are more likely to learn from peers similar to them rather than peers who set a shining example for the rest (Ben Yishay and Mobarak 2019). Outcomes may also be disappointing because the implicit assumption that knowledge spreads spontaneously among intended beneficiaries, perhaps as a side-effect of normal social interaction, is false. The current paper asks whether applying incentives and nudges (applying loss framed messages) promotes the diffusion of information about a specific agricultural innovation.

Several papers analyze social learning and study how information spreads through social networks (e.g., Bandiera and Rasul 2006; Conley and Udry 2010; Magnan et al. 2015; Beaman et al. 2019). New information does not always spread spontaneously, and typically the source

[1] There are several other well-known reasons for incomplete adoption of modern inputs. For example, farmers must incur costs when trying to access markets and may lack the liquidity to pay for modern inputs (e.g., Feder et al. 1985; Duflo et al. 2011). Heterogeneity in returns to technologies among farmers implies that the adoption of new technologies may not be in the interest of all farmers (Suri 2011).

should invest additional effort—especially when it is difficult or risky for farmers to mimic the behavior of successful adopters (Munshi 2004). Additional instruction or information may be needed to implement an innovation successfully. However, unless there are strategic complementarities in adoption, lead farmers may not be interested in teaching others as doing so involves an opportunity cost. Spreading information is a conscious choice that involves an effort cost, so economic reasoning based on incentives is likely relevant when explaining lead farmer behavior.

Shikiku et al. (2018) propose that lead farmers may be motivated by altruism, social prestige, or private gains when deciding to engage in information sharing. Hence, one solution is to promise a material incentive to lead farmers if they share information—compensating them for their effort cost. Recent research suggests that providing private incentives to farmers to spread information increases diffusion effort (Ben Yishay and Mobarak 2019).[2] Shikuku et al. (2018) demonstrate that social prestige incentives may work equally well. The social prestige incentive they use is a material gift to the local community rather than the lead farmer herself. Gifts like a weighting scale for the community may be a source of status for individuals who "earned" them.

This paper evaluates the effectiveness of different approaches to incentivizing lead farmers to spread information about a new agricultural technology. We probe the robustness of earlier findings concerning private rewards (Ben Yishay and Mobarak 2019) and social prestige (Shikuku et al. 2018) for a new technology (integrated pest management) in a novel context (rural Ethiopia). We extend the literature by (i) considering a new type of social prestige incentive with only symbolic value (rather than a material reward for the community), (ii) leveraging *existing* semi-formal local institutions for knowledge diffusion rather than working with artificial institutions created during the experiment, and (iii) combining the private or social prestige incentive with a loss framed incentive (see below), and probing their interaction by leveraging loss aversion (and maybe shame). We are among the first to apply loss-framed incentives in a real field experimental setting in a low-income country condition, where leveraging loss aversion

[2] The effectiveness of incentives for technology diffusion was also demonstrated in other domains. For example, Sseruyange and Bulte (2018) study incentives and the spread of financial knowledge, and Alem and Dugoua (2022) study incentives and diffusion of knowledge about the benefits of a solar lamp.

is much more complex than in controlled (lab) settings. This implies a final (methodological) contribution: (iv) in addition to evaluating the *impact* of loss-framed incentives, we also examine their *feasibility* in a context characterized by imperfect contracting and long production cycles.

To study the impact of incentives and loss-framed messages in the context of smallholder social learning, we implement a randomized controlled trial (RCT) in northwestern Ethiopia. We introduce a new integrated pest, weed, and soil management system that promises several benefits to adopters—protection from pests and invasive weeds and providing high-quality feed for livestock. In our study area, the conventional extension system uses a farmer training center (FTC) and the so-called 'one-to-five groups' approach to disseminating agricultural information (ATA and MoA, 2014). Groups of six smallholder farmers have one designated leader who receives training on the new technologies or practices and is expected to share lessons with his group members. We construct a sample frame consisting of 571 farmer groups from the lists of available groups in the villages and randomly allocate these groups to one of five experimental arms: (a) a control group, (b) two arms where group leaders receive a private incentive (a modern sickle) for diffusing information, and (c) two arms where leaders receive a social prestige incentive (a framed certificate of recognition). The sickle is an important farm tool for weeding and harvesting crops. The framed certificate signals the leader's successful participation in the experiment and only has symbolic value. Orthogonal to the incentive treatments, we introduce loss and gain frames so that, in total, we have four treatment groups.

In the gain frame, leaders are promised a reward in case of sufficient performance. In the loss frame, instead, leaders received their reward *up-front*, but they have to return it in case of insufficient performance. While gain and loss-framed incentives are perfectly isomorphic from a conventional economics perspective, the literature in behavioral economics emphasizes that their incentive effects are different if leaders are loss-averse—i.e. have a utility curve that is "kinked" at the endowment level. If returning a reward involves greater disutility than not winning a same-sized reward, leaders should supply greater effort with loss-framed incentives (Bulte et al. 2020).

From an intervenionist's perspective, it is an open question whether all types of rewards (incentives) can be "clawed-back" in practice, as intended in the loss frame regime. While up-front incentives can be easily re-claimed in lab-style settings, it is not obvious whether this is credible in the setting of a field experiment—especially if complete contracting is impossible and there is a long delay between the timing of the up-front payment and the actual measurement of performance. For example, workers can hide their reward or claim it is stolen or broken. Anticipating that they can hide their reward later, lead farmers may have little incentive to change their behavior now, and the effectiveness of the loss-framed incentive would be attenuated or even nullified.[3]

Our main results are as follows. We find that both private and social prestige incentives increase diffusion effort by lead farmers and promote knowledge diffusion to other farmers. These findings confirm insights from the existing literature. Extending the literature, we find evidence of interaction between incentives and the loss-framed message. Loss-framed incentives increase the effectiveness of social prestige incentives—losing prestige is worse than not obtaining it. This could reflect the power of shame. We also find that loss-frames induce little extra effort, on average, when combined with a private reward. As mentioned above and further discussed below, this could reflect a lack of credibility of our threat to re-claim the material reward in case of insufficient performance.

Interestingly, we also find that diffusion effort by group leaders does not translate in additional experimentation and adoption by follower farmers. This suggests farmers face other constraints besides a lack of information, or that access to information is a necessary but not a sufficient condition for adoption. It is also important to realize that earning the rewards was conditional on information diffusion (instead of contingent on experimentation by follower farmers, as in BenYishay and Mobarak 2019). In other words, the leaders responded to incentives, but accomplished nothing more. Further study is required, but potential adoption barriers in the context of our study region could include labor and skill constraints, and a preference for intercropping with food crops rather than forage crops.

This paper is organized as follows. Section 2 provides additional information about loss-framed nudging. In section 3 we sketch the experiment and provide testable predictions. In section 4 we

[3] Fryer et al. (2018) study clawback incentives for teachers in a real school setting, with a full academic year between up-front payment and performance assessment. However, teachers are employed by the school, and the experimenters wrote a contract specifying outcomes in case of non-performance. Instead, casual employment is common in low-income countries, and imperfect contracting is the norm. It is not clear that the clawback or loss-frame be implemented under such circumstances.

summarize our data and outline our identification strategy. Section 5 provides our regression results and examines how the combination of incentives and loss-framed messages affects diffusion effort and diffusion patterns. The discussion and conclusions ensue.

2.2. Theoretical background

The experiment involves combinations of incentives and loss-framed messages to promote knowledge diffusion about a particular agricultural innovation. The notion that farmers respond to incentives needs no additional explanation, and neither does the idea that the prospect of earning a private reward can be a powerful stimulus (e.g., Lazear 2000). The idea that people care about their social position and prestige, and that this affects their behavior, is also rather well-established and has been worked out in theory (e.g., Benabou and Tirole 2006) and demonstrated in empirical work (e.g., Ashraf et al. 2014; Carpenter and Myers 2010). However, the hypothesis that the provision of effort can be manipulated by framing rewards as a gain or a loss is more controversial. We will elaborate on that hypothesis in this section.

Focus on the case of material rewards first. An assumption underlying the effectiveness of loss-framed incentives is that a sizable fraction of the population displays reference-dependent preferences or is loss averse (e.g., Kahneman and Tversky 1979). Kőszegi and Rabin (2006) developed a model of reference-dependent utility where an individual's utility consists of two components: conventional consumption utility and so-called "gain-loss utility". Specifically, utility is assumed to depend on a consumer's k-dimensional consumption vector $\boldsymbol{c}$ and a reference vector $\boldsymbol{r}$, as follows:

$$u(\boldsymbol{c}|\boldsymbol{r}) = \sum_k m_k(c_k) + \sum_k \mu(m_k(c_k) - m_k(r_k)), \tag{1}$$

The first term on the right-hand side captures utility derived from consuming good k. The second term introduces reference dependence and captures gain-loss utility. Value function μ is defined as: $\mu(x) = \eta x$ for $x > 0$ and $\mu(x) = \eta\lambda x$ for $x < 0$, where parameter η is the idiosyncratic weight attached to gain-loss utility and λ is the consumer's coefficient of loss aversion. Parameter λ is assumed to be greater than unity ($\lambda > 1$), so that utility loss associated with outcomes c_k below reference value r_k are greater than utility gains from equal-sized realizations in excess of r_k.[4] Kahneman and Tversky (1979) speculate that reference points may be based on current *endowments* (agents want to keep what they have), *expectations* (agents want to obtain what they expect to receive), or *aspirations*. Recent experimental work examined the role of endowments and expectations in reference point formation (e.g., Abeler et al. 2011; Ericson and Fuster 2011; Banerji et al. 2014; Heffetz and List 2014).

If economic agents maximize (1), then manipulating reference points r_k will provoke a behavioral response. Specifically, if reference points shift in response to varying the timing of rewards, then the framing of incentive designs invites an effort response. The idea is that up-front payment of rewards is akin to providing that agent with an endowment which shifts reference point r_k. If recipients anchor on owning the reward, equation (1) says that losing the reward is worse than not earning it (as $\lambda > 1$). The theory of loss-framed incentives exploits this idea. It proposes that up-front bonus payments, which should be returned in case of non-compliance, extract greater effort from agents than the promise of an *ex-post* bonus of the same size (in case of compliance).[5] Because workers have to return their rewards in case of underperformance, loss-framed incentives are also known as "clawback" incentive regimes.

If information about individual behavior is public, then "image payoffs" may also enter. The idea of using social incentives to motivate desired behavior is to leverage public recognition. This

[4] The implications of reference-dependent preferences are studied in several domains, including technology adoption (Dupas 2014), housing demand (Simonsohn and Loewenstein 2006), and labor supply (Crawford and Meng 2011).

[5] Some evidence supports this line of reasoning. Several lab studies indicate that agents work harder in a loss-frame incentive regime. In a lab-style real effort experiment in Uganda, Bulte et al. (2020) found that workers are 30% more productive under a loss-frame than a gain-frame regime. Moreover, recognizing their own productivity response, many "sophisticated" workers self-select in a loss-framed incentive regime when given that option—using it as a commitment device for supplying high effort (see also Imas et al. 2017, and Brownback and Sadoff 2020). Hossain and List (2012) worked with a Chinese electronics company and found that the threat of clawing back up-front bonuses raised the productivity of teams of workers. Fryer et al. (2018) found that up-front payments raise the productivity of schoolteachers compared to a conditional bonus scheme. Not all evidence supports the effect of loss-framed incentives. Smaller and sometimes insignificant effects are found in experiments on dieting (List and Samek 2015), educational performance of school children (Levitt et al. 2016), recruitment (de Quidt 2018), and effort provision during simple online tasks (DellaVigna and Pope 2018).

means that information about behavior or performance should be shared with a relevant group of other people. In case a leader meets the norm, this may lead to pride. In contrast, falling short of the norm could cause feelings of shame. Tangney et al. (2007) emphasize that, from a psychological perspective, pride and shame are separate emotions of different valences. In other words, the welfare gain caused by an increase in status (pride) may be smaller than the welfare loss caused by an equal-sized decrease in status (shame). Indeed, Butera et al. (2022) study the welfare effects of positive and negative image payoffs (pride and shame, respectively), and find that the so-called *public recognition utility function* is non-linear. The public recognition utility function may be "kinked"— akin to the consumption-based utility function that includes gain-loss utility. If so, applying a loss frame when using social incentives would be more effective than applying a gain frame.

2.3. Experimental design and predictions

2.3.1. The new technology

The intervention we study is a new agronomic approach in the study area developed to reduce pest infestation and improve soil fertility and fodder production. Push-pull technology (PPT) is a novel cropping system designed for integrated pests (e.g., stemborer, fall armyworm), weed, and soil management in cereal-livestock production systems (Picket et al. 2014; Khan et al., 2018). It combines the target cereal crop with a pest repellent plant (push) and a trap plant (pull) to control pests. PPT is an intercropping system where maize or sorghum are grown with two other fodder crops: *Desmodium* and *Brachiaria*.[6] Both companion crops are perennial forage legumes, important to diversify income through enhancing livestock production and selling the fodder (Kassie et al. 2018). The push-pull system's high-quality forage supply is an additional incentive to farmers, given the shortage of grazing land in the study areas.

[6] *Desmodium* is a fodder legume with repellent chemicals, grown between rows of cereals in the field—that "push" stemborers away from the field. Moreover, the root exudates of *Desmodium* control the parasitic Striga weed, causing abortive germination, and the plant contributes to improved soil fertility by nitrogen fixation, natural mulching, and erosion control. *Brachiaria* is grown at the field's edges and attracts stemborers—"pulling" them from the nearby cereal crop (e.g., Picket et al. 2014; Khan et al. 2018). *Desmodium* and *Brachiaria* seeds can be produced locally to enhance seed availability at a lower price.

Stemborer and fall armyworm are the main pests affecting maize production in the study area, reducing maize yields by on average 12.3 % and 16.2%, respectively. PPT reduces crop damage by pests (Midega et al., 2018; Fetene et al., 2021). It also increases and diversifies income (Kassie et al., 2018), increases women's nutrition security (Kassie et al., 2020) and farmers' resilience to climate shocks (Gugissa et al., 2022), and improves soil health (Ndayisaba et al. 2020; 2021). Some evidence suggests PPT reduces labor demand once established (Diiro et al. 2021). However, there are barriers to adoption. First, PPT is a labor-intensive technology during the first season of land preparation, planting, weeding, and trimming and harvesting of forage crops to avoid competition with main crops. Second, the companion (forage) crops compete with the main (food) crops for space. Third, a well-functioning seed system for companion crops does not exist yet. Fourth, adopting the technology requires knowledge—farmers require information and training to apply it on their farms.[7] Any of these constraints may be binding, and we focused on relaxing the third and fourth one. We made seed available via extension workers which farmers could obtain at zero cost, and organized an extension intervention to promote the diffusion of information.

In this extension experiment we leverage an existing institution to promote the diffusion of knowledge about PPT. So-called "one-to-five groups" are a key component of the extension system in Ethiopia. Farmers form groups of six and then propose one group member as their "leader". This leader figure is targeted for training by extension officers promoting new technologies or practices and instructed to share new knowledge with other group members (also known as followers). Diffusion of information via one-to-five groups is therefore planned but not supervised or monitored. In our experiment, we try to increase the performance of existing one-to-five groups by incentivizing and loss-framed messaging group leaders to spend more time and effort spreading knowledge about PPT.

[7] Planting companion crops and separating *desmodium* from weeds during establishment are difficult tasks of implementing the technology. Until farmers have seen desmodium seedlings growing, they struggle to distinguish between weeds and the crop.

2.3.2. *The treatment arms*

We organize our experiment in Jabi Tehnan district (Ethiopia), in 38 kebeles (or villages) and 113 sub-kebeles. We organized a census of the one-to-five groups in all 113 sub-kebeles in the district, and randomly picked approximately five groups per sub-kebele for participation in the experiment. The total sample consists of 571 groups (and group leaders). All selected group leaders participated in a two-day theoretical and practical training session in their villages, offered by the International Centre of Insect Physiology and Ecology (*icipe*) and the Jabi Tehnan district agricultural office experts. The training took place in May 2018. Group leaders received training on PPT implementation techniques (correct spacing, row planting, timely weeding and harvesting of companion crops) and learned about its benefits. Group leaders were also offered seeds of the companion crops (*Desmodium* and *Brachiaria*). We made seeds available via village extension offices to stimulate experimentation, and both leaders and group members could access them at zero cost (in season 1). They were advised to use vegetative propagation after season 1, but could also purchase new seeds at cost-recovery prices (in season 2).

After participating in the training, group leaders were assigned to one out of five experimental arms and learned about the incentive regime they would be working in. Importantly, farmers heard about incentives *after* participating in the training, so their effort and learning are unaffected by the assignment to treatment arms. We assigned all group leaders from the same sub-kebele to the same experimental arm to avoid spillover effects.

One-third of the sub-kebeles were assigned to the control arm (38 sub-kebeles, 190 group leaders). These group leaders received the PPT training and were asked to share the information with their group members but received no incentives for knowledge diffusion. The remaining leaders were promised an incentive in case of sufficient knowledge diffusion within their own one-to-five group. Specifically, as a threshold for adequate leader performance we used the rule that leaders should successfully train at least half of their group members (so that these members could pass a simple knowledge test—see below). One-third of the sub-kebeles were assigned to a private reward incentive, and group leaders were promised a sickle if they trained enough group members (38 sub-kebeles, 190 group leaders). Finally, one-third of the sub-kebeles were assigned to the social prestige incentive. These leaders receive a “good performance” framed certificate if they trained at least half of their group members (38 sub-kebeles, 191 leaders).

After being assigned to receive a private or social reward, group leaders were randomly allocated to either a gain or loss frame.[8] Half of the group leaders in the private reward arm were promised the sickle as a *bonus* in case of sufficient performance. In what follows, we refer to this as the "private incentive framed as a gain" (*PIGF*). The other half of the leaders received the sickle as an *up-front* reward, but were told they would have to return it if they failed to share knowledge with at least half of their follower farmers. These group leaders were in the "private incentive framed as a loss" group (*PILF*). Leaders in both *PIGF* and *PILF* were informed about the conditions for earning the sickle in an instruction session that was specific to their experimental arms. They received their reward in the privacy of their house. If a leader under-performed in the *PILF*-arm then the sickle would be re-claimed in private during a visit to the leader's house. Information about leader performance, relative to others, was kept private.

Leaders from the social prestige arm were also allocated to two sub-groups: a "social incentive framed as a gain" group (*SIGF*) and a "social prestige incentive framed as a loss" group (*SILF*). Certificates of recognition were offered during an official village visit at which the village chief was also invited—either at baseline (*SILF*) or at endline (*SIGF*). The recall of certificates in case of insufficient performance, in the *SILF* arm, was organized as a "semi-public" event. Specifically, leaders who failed to meet the threshold were visited at their houses (to avoid the outright embarrassment of having to return their certificate during a public gathering), but village chiefs were informed about the outcomes during a separate visit. Group leaders were also informed about this sequence of events during separate instruction sessions organized for the leaders from the different experimentral arms. Hence, we have four treatment groups and one control group. The experimental design is summarized in Table 1.

[8] This means the number of groups in each of the 4 treatment arms is about half the number of groups in the control arm (see Table 1). We could have assigned fewer groups to the control arm to increase power for the comparisons that we make across the four incentive regimes. However, we are sufficiently powered to pick up small effects. Moreover, differences in intervention costs between participants from the treatment and control groups suggested an uneven ratio of participants from the 4 treatment arms to the control participants. The current assignment of groups to treatment is a compromise between these considerations.

Table 1: Summary of experimental design

Sample frame: 114 sub-kebelles (*K*), 571 lead farmers (*N*), 1684 follower farmers (*n*)				
Training only: *K*=28, *N*=190, *n*=565	*Private incentive*: *K*=29, *N*=190, *n*=566		*Social prestige incentive*: *K*=29 *N*=191, *n*= 554	
	Private incentive framed as a gain (*PIGF*) K=14, N=90, n=271	Private incentive framed as a loss (*PILF*) K=15, N=100, n=295	Social prestige framed as a gain (*SIGF*) K=14, N=91, n=265	Social prestige framed as a loss (*SILF*) K=15, N=100, n=290

Hypotheses

Based on the theory in Section 2, our predictions about the experimental outcomes are as follows.

1. The private and social prestige incentive will increase knowledge diffusion from the group leader to group members relative to the control group. We have no clear ex-ante predictions regarding whether the private or social prestige incentive is more powerful in promoting knowledge diffusion—this presumably is context-specific and, therefore, an empirical matter.

2. If group leaders are loss averse, then the loss frame is more successful in promoting knowledge diffusion than the gain frame when combined with the private incentive.

3. If the welfare loss due to shame is greater than the welfare gain due to pride, then the loss frame is more successful in promoting knowledge diffusion than the gain frame when combined with the social incentive.

Also, return to the challenge of implementing the loss frame "in the field" caused by the long delay between the up-front payment and the day of reckoning. An asymmetry exists between the social prestige reward and the material reward. While claiming back the sickle requires some cooperation by the group leader, cooperation by the leader is not required to destroy the symbolic value of the social prestige reward. If group leaders fear that rumors about unsatisfactory performance may spread regardless of whether they retain or return their certificates, then the symbolic value of retained certificates will be nullified. Clawing back certificates should therefore be easier than clawing back sickles. This results in the final prediction.

4. In a field setting where material rewards can be hidden so that reclaiming them is difficult, then the "clawback" design is not credible and the loss-frame should be less effective than the gain frame. This consideration is unlikely to be relevant for social prestige incentives.

We now explore the extent to which these predictions are supported by our data.

2.4. Data and identification strategy

We obtained IRB approval for the experiment, pre-registered the experiment,[9] and obtained informed consent from respondents and local officials. Well-trained enumerators and supervisors collected two waves of panel data using tablets. We collected baseline survey data in June-August 2018 before training the group leaders. We visited 114 sub-kebeles to survey selected group leaders and between 3 and 4 follower farmers per group. In total, 571 group leaders and 1684 follower farmers were interviewed. We collected information on household demographics, food security, crop and livestock production, plots owned and managed, exposure to and awareness of maize field pests (including stemborer, fall armyworm, and Striga weed), field pest controlling strategies, sources of agricultural information, and knowledge about farming practices.

Appendix Table A1 summarizes the baseline data for the group leaders and provides t-tests to check pre-existing differences across experimental arms. On average, leaders are 45 years old, have a family of 5.5 persons, and receive some 6 years of education. Nearly all leaders are male, and the main activity is crop farming. Half the leaders own a milking cow. There are few significant differences among leaders in their field pests experiences and knowledge. A sizable minority of the leaders know that stemborer, fall armyworm, and Striga weed constrain maize production. We include baseline controls to increase the precision of our impact estimates and control for any pre-existing differences. Appendix Table A2 summarizes baseline data for follower farmers. Again, assignment to treatment arms resulted in balanced groups. Joint

[9] The experiment was pre-registered at: https://www.socialscienceregistry.org/trials/5642AER.

significance tests based on a regression model with all baseline controls also indicate that treatment assignment is unrelated to group member or leader characteristics (not shown).

We undertook the second survey wave in November 2020, after the second post-experimental cropping season, to measure group leader performance. We managed to revisit 558 group leaders and 1644 follower farmers; 53 respondents were missing during the follow-up (13 group leaders and 40 follower farmers). These individuals had either moved or were not available. Regressing attrition status on baseline controls and treatment arms suggests that attrition is nearly random and not correlated with treatment status (see Appendix Table A3).

We collected data on PPT knowledge diffusion and experimentation during the follow-up survey. We use five different proxies to measure group leader effort: (i) whether or not the group leader organized at least one training event to share PPT knowledge with follower farmers (binary variable, based on whether follower farmers reported the occurrence of such an event); (ii) the number of group members that the leader reached out to individually (count variable—the number of follower farmers who reported having received PPT information from their leader during a one-on-one conversation); (iii) knowledge of follower farmers (binary variable based on whether farmers met a knowledge threshold about PPT: farmers should be able to mention at least two benefits as well as have a basic understanding of how PPT "works" in the field);[10] (iv) PPT experimentation by leaders (binary variable indicating whether the group leader experimented with PPT on his plots, based on survey question and verified in the field); and (v) PPT experimentation by followers (binary variable whether any of the group members tried out PPT on their plots, based on survey question and verified in the field).

Outcome variables are summarized in Table 2, split across the five experimental arms. In the top row we summarize outcomes for the group leader performance test. In the control group, 64% of

[10] Advantages of PPT include: (i) control of stem-borer and fall armyworm, (ii) reduction of soil erosion, (iii) improvement of soil fertility, and (iv) companion crops are an important source of animal feed. Regarding implementation and functioning, PPT involves (i) intercropping a leguminous fodder crop called desmodium with maize, (ii) desmodium pushes stem-borer away from the maize, (iii) brachiaria is sown surrounding the maize plot; and (iv) brachiaria attracts the stem-borer, "pulling" it from the maize field. If farmers are able to mention at least two advantages and two items regarding implementation and functionality, then she is defined to have "basic knowledge" about PPT and received a score of one for the binary variable (else a zero).

the leaders from the control group trained at least half of their group members about the benefits and implementation of PPT. Leaders in the various treatment arms consistently do better, especially in the *SILF* arm where 89% of the leaders trained enough follower farmers, presumably to avoid the shame of being exposed to a certificate recall.

Next, we turn to the dependent variables used in the regression analysis. In the control group nearly half the group leaders organized a session to discuss PPT, and, on average, leaders discussed PPT with nearly half of their follower farmers. However, only 9% of the follower farmers possessed basic knowledge of PPT. In the control group, about one-third of the leaders adopted PPT on their maize plot, but uptake by their group members was low. Overall, 64% of the group leaders shared knowledge with at least 50% of their group members. The active attitude of group leaders in the control group reflects that the one-to-five group is an existing and functioning institution in our study region. Group leaders have a clear mandate and are experienced in communicating with their peers.

Table 2: Descriptive statistics of performance measures

	Control	Private incentive framed as a gain (*PIGF*)	Private incentive framed as a loss (*PILF*)	Social prestige framed as a gain (*SIGF*)	Social prestige framed as a loss (*SILF*)
	Performance measure for reward				
Group leader meets threshold (50% follower farmers “knowledgeable”)	0.64	0.81	0.75	0.71	0.89
	Dependent variables				
Leaders organized at least one training event to train group members about PPT (1=yes)	0.46	0.61	0.58	0.55	0.73
Leaders provided PPT information to group members (Number of group members)	2.5	3.6	2.9	2.9	3.7
Leaders experimented with PPT on their maize plots (1=yes)	0.29	0.51	0.34	0.42	0.44
Group members have knowledge of PPT (1= have full knowledge)	0.09	0.37	0.38	0.34	0.38
Group members experimented with PPT on their maize plots (1=yes)	0.04	0.04	0.06	0.04	0.03
Number of group leaders	190	90	100	87	91
Number of follower farmers	557	268	293	240	286

However, it is obvious from descriptive statistics (Table 2) that various incentive and framing regimes outperform the control group. Treated group leaders appear more active and achieve higher knowledge scores among their followers than group leaders from the control group. But we also document that leaders did not manage to convince many follower farmers to adopt PPT on their plots — neither in the control group nor in the intervention groups. While adoption rates for leaders are in the 30-50% range (depending on treatment), the adoption rate among follower farmers hovers around 5% only. Adoption by followers is low and not significantly different across arms. This suggests that some of the barriers to adoption discussed above prevented farmers from taking up the new technology.[11]

We use the following equations to estimate the impact of incentives and loss-framed messages on knowledge diffusion (effort):[12]

$$Y_{is} = \alpha + \beta_1 Private\ reward_{is} + \beta_2 Social\ prestige_{is} + \gamma X_{is} + \delta_v + \varepsilon_{is}, \tag{2}$$

$$Y_{is} = \xi + \psi_1 Gain\ frame_{is} + \psi_2 Loss\ frame_{is} + \rho X_{is} + \delta_v + \varepsilon_{is}\ , \ \ \& \tag{3}$$

$$Y_{is} = \pi + \varphi_1 PIGF_{is} + \varphi_2 PILF_{is} + \varphi_3 SIGF_{is} + \varphi_4 SILF_{is} + \upsilon X_{is} + \delta_v + \varepsilon_{is}\,. \tag{4}$$

In equation (2), we estimate the effect of private incentive and social prestige incentive on our five performance measures Y, conditional on the ith respondent baseline characteristics (X_{is}) in sub-kebele s, and kebele-level fixed effects, δ_{v}. The incentive treatment effect estimated in (2) is not conditional on the loss-framed messages. Equation (3) estimates the effect of the loss-frame on the same outcome variables. This treatment effect is not conditional on the underlying incentive (private or social prestige reward). Finally, equation (4) introduces all four experimental groups separately via four dummy variables: *PIGF*, *PILF*, *SIGF*, and *SILF*. This model allows exploring whether incentives and loss-framed messages interact, for example, if group leaders are more sensitive to losing social prestige than the private reward. In models (2-4), the control

[11] In case of heterogeneity among follower farmers, it is conceivable that leader farmers tried to target the subsample of farmers who they expected to gain from their training. However, we do not have data to test this. Anyway, leaders were incentivized to train follower farmers regardless of whether they were likely to adopt. Perhaps outcomes would have been different if we had incentivized leader farmers to promote actual adoption or experimentation by leader farmers (as in BenYishav and Mobarak 2019). This is left for future research.

group is the omitted category. Since the unit of randomization was the sub-kebele level, we also cluster the standard errors at the sub-kebele level. To control for multiple hypotheses testing and reduce the False Discovery Rate (4 treatments × 5 dependent variables) we also report Anderson sharpened *q*-values, in addition to regular *p*-values, following Benjamini and Hochberg (1995).

Since we randomly assigned groups to treatment, the identifying assumptions are satisfied unless there are spillover effects from the treated groups to the control group (or from one treated group to another). This would violate the stable unit treatment value assumption, SUTVA. To minimize this risk, treatments are block randomized at the sub-kebele level—one treatment per sub-kebele. To further check whether the SUTVA is satisfied, we focus on the behavior of group leaders in the control group. We distinguish between two types of leaders—those living in kebeles without incentivized leaders (52 leaders) and those with leaders from any of the other treatment arms (138 leaders). We compare the diffusion effort of the control group leaders across these two groups to explore whether the nearby presence of incentivized leaders affects behavior.

2.5. Regression results

Table 3 summarizes the regression results of equation (2) and describes the average effect of private and social prestige incentives on PPT knowledge diffusion. Below the coefficients we report *p*-values and FDR sharpened *q*-values (observe that *q*-values can be smaller or greater than *p*-values).

Both types of incentives induce group leaders to increase their diffusion effort—the propensity to organize a training event, one-on-one outreach to individual members, and experimentation by the leaders. We also find that both types of incentives promoted learning among group members though this did not translate into significant uptake of PPT during the time interval of our study. Table 3, therefore, supports prediction 1. Private and social incentives are equally successful in promoting knowledge diffusion when measured at the level of follower farmers (columns 4-5). However, leaders incentivized by a certificate are significantly more likely to organize a training event for group members (column 1, p=0.04) and to reach out to individual members (column 2, p=0.07).

Table 3: Private and social prestige incentives

	Knowledge diffusion effort by leaders			*Follower outcomes*	
	Leader organized at least one training event	Number of group members received PPT information from leaders	Leaders experiment with PPT	Group members have knowledge of PPT	Group members experiment with PPT
Private reward	0.12** (0.03) (0.02)	0.54*** (0.01) (0.01)	0.16*** (0.02) (0.02)	0.24*** (0.00) (0.00)	-0.00 (0.80) (1.00)
Social prestige	0.24*** (0.00) (0.00)	0.94*** (0.00) (0.00)	0.15** (0.02) (0.02)	0.26*** (0.00) (0.00)	-0.01 (0.70) (1.00)
Village fixed effects	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes
Mean of dependent variables	0.56	2.9	0.38	0.27	0.04
R-squared	0.15	0.16	0.22	0.21	0.06
Number observations	558	558	558	1644	1644
Test of coefficients: Private vs social incentives (*p*-values)	0.04	0.07	0.93	0.53	0.80

Note: All models are estimated with O.L.S., standard errors clustered by sub-village (the unit of randomization). The private reward is a sickle, and the social recognition is a framed certificate of good performance. Controls are *Household members* (number), *Household head age* (years), *Household head literacy* (1= yes), *Primary activity farming* (1=yes), *Lived in the village* (years), *Farm size* (hectare), *Grazing land* (hectares), *Knowledge of fall armyworm* (1= yes), *FAW major constraint in maize production* (1=yes), *Knowledge of stemborer* (1= yes), *Stemborer (SB) major constraint in maize production*? (1=yes), *Knowledge of Striga* (1=yes), *Striga is major constraint in maize production* (1=yes), *Extension* (*Number of visits last 12 months*), *Credit constrained* (1=yes), and *Household owns milking cow* (1=yes). In parentheses, p-values and Anderson's sharpened q-values. ***p < 0.01, ** p < 0.05, and * p < 0.1 indicate significance levels**.**

Table 4 presents the average impact of gain- and loss-framed incentives, not controlling for the type of incentive. Both gain- and loss-framed incentives raise diffusion effort and knowledge diffusion, and leaders in a loss-framed regime are more likely to organize a training and information session with group members than their counterparts from the gain frame (*p*=0.09). The long delay between up-front payment and performance measurement, on average, did not negate the effectiveness of loss-framed messaging. Table 4 supports hypotheses 2 and 3.

Table 4: Gain and loss-framed incentives

	Knowledge diffusion effort by leaders			*Follower outcomes*	
	Leader organized at least one training event	Number of group members received PPT information from leaders	Leaders experiment with PPT	Group members have knowledge of PPT	Group members experiment with PPT
Gain frame	0.13** (0.02) (0.01)	0.65*** (0.00) (0.00)	0.19*** (0.00) (0.01)	0.28*** (0.00) (0.00)	0.01 (0.77) (1.00)
Loss frame	0.23*** (0.00) (0.00)	0.83*** (0.00) (0.00)	0.12* (0.07) (0.04)	0.24*** (0.00) (0.00)	-0.01 (0.39) (1.00)
Village fixed effects	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes
Mean of dependent variables	0.56	2.91	0.38	0.27	0.04
R-squared	0.15	0.16	0.22	0.21	0.06
Number observations	558	558	558	1644	1644
Test of coefficients: Gain vs loss frame (*p*-values)	0.09	0.38	0.25	0.36	0.21

Note: All models are estimated with O.L.S., standard errors clustered by sub-village (the unit of randomization). The private reward is a sickle, and the social recognition is a framed certificate of good performance. Controls are *Household members* (number), *Household head age* (years), *Household head literacy* (1= yes), *Primary activity farming* (1=yes), *Lived in the village* (years), *Farm size* (hectare), *Grazing land* (hectares), *Knowledge of fall armyworm* (1= yes), *FAW major constraint in maize production* (1=yes), *Knowledge of stemborer* (1= yes), *Stemborer (SB) major constraint in maize production*? (1=yes), *Knowledge of Striga* (1=yes), *Striga is major constraint in maize production* (1=yes), *Extension* (*Number of visits last 12 months*), *Credit constrained* (1=yes), and *Household owns milking cow* (1=yes). In parentheses, p-values and Anderson's sharpened q-values. ***p < 0.01, ** p < 0.05, and * p < 0.1 indicate significance levels.

Table 5 contains our main results, summarizes results for all combinations of incentive types and loss-framed messages separately, and compares the various coefficients to each other. The aggregate results in Tables 3 and 4 conceal quite a bit of heterogeneity. First, consider the impact of incentive-and loss-frame combinations on diffusion effort of group leaders (columns 1-3)—the key variables directly under the leader's control. All coefficients have the expected sign, and all-but-two are significantly different from zero. Social prestige framed as a loss wins the "horse race" for our two training variables – the propensity to organize a training event and the number of group members informed individually. The probability of organizing a training event increases by more than 60% (compared to the control group), and the number of group members reached by leaders increases by 44%. The former result is significantly greater than any of the other treatment effects (*p*=0.00), and the latter is greater than the effect of the private reward framed as a loss (*p*=0.00) or a gain (*p*=0.09), and the social reward framed as a gain (*p*=0.00).

The regression results in Table 5 suggest the following narrative. Leaders do not prefer a framed certificate of recognition over a sickle. If anything, leaders work harder to earn a material reward when a gain frame is used than a symbolic one: *PIGF* coefficients in columns (1-3) in Table 5 are consistently larger than *SIGF* coefficients, but not significantly different. But *losing* social prestige is considered worse than losing the sickle. Leaders work harder to retain their certificate than their sickle (p=0.00 in columns 1 and 2). Social prestige, therefore, affects leaders' behavior asymmetrically—while leaders invest modest levels of effort to earn it, they work quite hard to prevent losing it. This is consistent with the notion that (avoiding) shame is a powerful stimulus for behavior, or that the social prestige function is highly non-linear (as in Butera et al. 2022). The coefficients associated with the social prestige incentive framed as a gain (*SIGF*) are relatively small. Weaker patterns emerge for the variable measuring whether the leaders experimented with the technology on one of their plots (column 3). While the leaders from the *PILF* group cannot be statistically distinguished from the control group, and are less likely to experiment with the new technology than the leaders from the *PIGF* group, the differences across groups are small. Across all groups we find that about half the leaders experimented with PPT on their farm.

Table 5: Incentives, gain-loss frames, and the diffusion of knowledge

Treatments	*Knowledge diffusion effort by Leaders*			*Follower outcomes*	
	Leaders organized at least one training event	Number group members received PPT information from leaders	Leaders experiment with PPT	Group members have knowledge of PPT	Group members experiment with PPT
Private incentive framed as gain (*PIGF*)	0.15** (0.03) (0.04)	0.81*** (0.00) (0.00)	0.23*** (0.01) (0.03)	0.20*** (0.00) (0.00)	-0.03 (0.05) (0.25)
Private incentive framed as loss (*PILF*)	0.14** (0.04) (0.04)	0.44* (0.08) (0.04)	0.10 (0.14) (0.12)	0.27*** (0.00) (0.00)	0.01 (0.52) (1.00)
Social prestige framed as a gain (*SIGF*)	0.12 (0.11) (0.05)	0.46** (0.03) (0.03)	0.14* (0.08) (0.10)	0.26*** (0.00) (0.00)	-0.00 (0.96) (1.00)
Social prestige framed as a loss (*SILF*)	0.34*** (0.00) (0.00)	1.29*** (0.00) (0.00)	0.14* (0.09) (0.10)	0.27*** (0.00) (0.00)	-0.01 (0.69) (1.00)
Village fixed effects	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes
Mean of dependent variables	0.56	2.91	0.38	0.27	0.04
R-squared	0.16	0.17	0.22	0.21	0.06
Number observations	558	558	558	1644	1644
Test of coefficients:					
PIGF versus *PILF*	0.89	0.22	0.09	0.21	0.06
PIGF versus *SIGF*	0.77	0.25	0.40	0.28	0.11
PIGF versus *SILF*	0.02	0.10	0.34	0.17	0.20
PILF versus *SIGF*	0.84	0.94	0.68	0.88	0.40
PILF versus *SILF*	0.00	0.00	0.62	0.90	0.18
SIGF versus *SILF*	0.01	0.00	0.97	0.80	0.67

Note: All models are estimated with O.L.S., standard errors clustered by sub-village (the unit of randomization). The private reward is a sickle, and the social recognition is a framed certificate of good performance. Controls are *Household members* (number), *Household head age* (years), *Household head literacy* (1= yes), *Primary activity farming* (1=yes), *Lived in the village* (years), *Farm size* (hectare), *Grazing land* (hectares), *Knowledge of fall armyworm* (1= yes), *FAW major constraint in maize production* (1=yes), *Knowledge of stemborer* (1= yes), *Stemborer (SB) major constraint in maize production*? (1=yes), *Knowledge of Striga* (1=yes), *Striga is major constraint in maize production* (1=yes), *Extension* (*Number of visits last 12 months*), *Credit constrained* (1=yes), and *Household owns milking cow* (1=yes). In parentheses, p-values and Anderson's sharpened q-values. ***p < 0.01, ** p < 0.05, and * p < 0.1 indicate significance levels.

Overall, our regression results also suggest that the credibility of the loss-frame is imperfect and varies with the nature of the incentive—consistent with prediction 4 (and therefore in contrast to prediction 2). To earn the private reward (sickle), leaders invest more effort in a gains frame (*PIGF*) than in a loss frame (*PILF*, column 3, p=0.09). This is consistent with the assumption that some leaders believe that they can keep "their sickle" regardless of performance—for example, they may try to hide it from the experimenter. For none of the other leader coefficients do we find that the loss frame extracts greater effort than the gain frame. The reverse seems true (even if other differences are not significant). But, consistent with expectations, the same is not true for the social prestige incentive. The gains from prestige are not embodied in the certificate but the opinion of others. If a leader successfully hides the sickle, he can benefit from it in the future. The gains from social prestige are not embodied in the certificate. If knowledge spreads that the leader failed to meet the threshold, social prestige dissipates and shame is the result, regardless of whether the leader owns a certificate.

This intuition is consistent with follow-up data we collected. In our experiment, 25 leaders from the *PILF* group failed to meet the threshold. The same was true for only ten leaders from the *SILF* treatment. After the end-line, we tried to give complying leaders from the gain-frame treatment arms their rewards, and tried to “claw back” the rewards from under-performing leaders from the two loss-frame arms. We were able to reach *all* leaders from the gain frame (100%). However, we could only contact 23 leaders from the *PILF* group and 8 from the *SILF* group—an attrition rate of more than 11%. Perhaps some under-performing leaders from loss frame arms avoided us? More importantly, while all leaders from the *SILF* group voluntarily returned their reward, this was not true for leaders from the *PILF* group. Here, 24% of the leaders (or 6 leaders) refused to return their sickle, using a range of excuses. If some leaders from the *PILF* group anticipated that they would not comply with the experiment’s rules, this explains the difference in the effectiveness of the incentive regimes. An equal proportions test indicates that the share of group leaders refusing to return their reward is significantly greater for private rewards than social prestige rewards ($p<0.05$).

Our analysis yields an additional result. While the extra effort by leaders in treatment arms translated into more knowledge by followers (column 4), it did not invite more experimentation by followers (column 5). We speculate that this reflects the existence of barriers to adoption, as discussed above (e.g. labor, space, seed). If follower farmers do not want to adopt innovations, the returns to diffusion effort will be negligible. While group leaders can be incentivized to work harder, reaching out to more farmers, the social return of these efforts is small if the context within which adoption should occur does not match the requirements of the innovation. Of course we should also consider that we incentivized leaders to promote knowledge diffusion among group members—not to promote experimentation by group members. Lead farmers did what was necessary to obtain (or keep) their reward, and not much more. Using an experimental design with rewards contingent on experimentation by others would plausible produce different impacts (e.g., BenYishay and Mobarak, 2019), and would be a useful extension in future research.

Finally, we checked the existence of spillover effects. Appendix Table A4 reveals that the presence or absence of incentivized group leaders in the kebele does not change leaders' diffusion effort from the control group. We interpret this as weak evidence that inter-sub-kebele spillover effects are relatively unimportant. As mentioned, intra sub-kebele spillover effects are ruled out by our design in which we assign all leaders from the same sub-kebele to the same experimental arm.

2.6. Discussion and conclusions

Many development experts believe that African smallholder farmers should adopt a greater range of productivity-enhancing inputs and practices if Africa is to reach the Sustainable Development Goals with respect to poverty, food security, and sustainable resource management. However, adoption levels are lagging behind expectations in many countries. Imperfect information flows may explain part of this—farmers may be unaware of innovations or be uncertain about how to implement them. But one lesson from our study is that adoption of push-pull technology is not just limited by a lack of information—other constraints impede uptake as well. Information is a necessary but nog a sufficient condition for adoption.

We use an experimental design to analyze the effectiveness of two types of incentives—a private material reward and a social prestige reward—and examine how incentives interact with gain-loss framed messaging. We consider an existing institution for farmer extension in rural Ethiopia. In these so-called one-to-five group system, a group leader is responsible for sharing information with his five follower farmers. The loss-framed messaging we apply provides leaders with an up-front privare or social reward that is "clawed back" in case of under-performance. This should make leaders work harder. In the case of a private reward, the loss frame leverages loss aversion, as leaders want to avoid losing their material endowment (if the threat of losing the up-front reward is perceived as credible). In the case of social prestige, leaders work harder because they want to avoid shame caused by losing social prestige.

Our main results are consistent with predictions. Group leaders who receive an incentive work harder than their peers from the control group to share information. Introducing a loss frame combined with a social incentive causes leaders to work even harder. This points to the salience of shame as a motivational emotion. The loss frame did little to improve the performance when combined with a material reward. This is consistent with the idea that several leaders did not believe that they would lose their up-front reward in case of under-performance. If clawing back material rewards after a long delay is difficult, the loss-frame loses its credibility. Since hiding and keeping a material reward is "relatively easy", as proven by some of our respondents, we indeed find that the effectiveness of standard loss-framed messaging combined with material rewards is compromised in the field. [12]

This implies an important lesson for policy makers. While loss-framed messaging can be used to increase effort, the "space" for applying them in a welfare-enhancing fashion may be narrower than heretofore expected. Clawing back private rewards may be difficult in real-life settings in low-income settings where incomplete contracting is the rule (rather than the exception). While clawing back social prestige may be easier to accomplish, it is not obvious that designing and implementing such incentive systems will improve welfare (Butera et al. 2022). The prospect of shame associated with losing prestige may more than offset potential welfare gains due to improved efficiency and productivity. For example, workers may be reluctant to work for firms that leverage shame to make people work harder as part of their incentive regime. Much more work on the welfare effects of loss-framed social rewards is needed.

In this paper we report the outcomes of linear models. However, we obtain qualitatively similar results when we estimate (non-linear) probit and poisson models instead. (Results not shown but available on request.)

[12] Not all leaders in a loss-frame regime anticipate refusing to return their rewards. But if some leaders believe they can hide their reward, then *average* performance of the incentive regime is compromised. This is what we find.

2.7. References

Abeler, J., A. Falk, L. Götte and D. Huffman. 2011. "Reference Points and Effort Provision." *American Economic Review* 101(2):470–492.

Alem, Y. and E. Dugoua. 2022. "Passive Learning and Incentivized Communication: A Randomized Controlled Trial in India." *Economic Development and Cultural Change*. In Press

Agricultural Transformation Agency (ATA) and Ministry of Agriculture (MoA). 2014. *National Strategy for Ethiopia's Agricultural Extension System: Vision, Systemic Bottlenecks and Priority Interventions*. Accessed at http://extwprlegs1.fao.org/docs/pdf/eth190334.pdf. (April 2021)

Bandiera, O. and I. Rasul. 2006. "Social Networks and Technology Adoption in Northern Mozambique." *The Economic Journal* 116: 869-902

Banerji, A. and N. Gupta. 2014. "Detection, Identification and Estimation of Loss Aversion: Evidence from an Auction Experiment." *American Economic Journal: Microeconomics* 6(1):91–133.

Beaman, L., A. BenYishay, M. Mobarak and J. Magruder. 2018. *Can Network Theory-Based Targeting Increase Technology Adoption?* Cambridge MA: National Bureau of Economic Research paper 24912

Benjamini, Y. and Y. Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerfiul Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B* 57(1):289-300

BenYishay, A. and M. Mobarak. 2019. "Social Learning and Incentives for Experimentation and Communication." *Review of Economic Studies* 86(3):976-1009

Brownback, A. and S. Sadoff. 2020. "Improving College Instruction through Incentives**."** *Journal of Political Economy* 128(8):2925-2972

Bulte, E. J. List and D. van Soest. 2020. "Toward an Understanding of the Welfare Effects of Nudges: Evidence from a Field experiment in Uganda." *Economic Journal* 130(632):2329-2353

Butera, L., R. Metcalfe, W. Morrison and D. Taubinsky. 2022. "Measuring the Welfare Effects of Shame and Pride.". *American Economic Review* 112(1):122-168

Carpenter, J. and C. Myers. 2010. "Why Volunteer? Evidence of the Role of Altruism, Image, and Incentives." *Journal of Public Economics* 94:911-920

Crawford, V. and J. Meng. 2011. "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependence Preferences with Rational Expectations Targets for Hours and Income." *American Economic Review* 101(5):1912–32

Conley, T. and C. Udry .2010. "Learning About a New Technology: Pineapple in Ghana". *American Economic Review* 100 (1):35-69.

Dar, M., A. de Janvry, K. Emerick, E. Sadoulet and E. Wiseman. 2021. "Private Input Suppliers as Information Agents for Technology Adoption in Agriculture." Unpublished manuscript. University of California at Berkeley

DeJanvry, A., E. Sadoulet and M. Rao. 2016. "Adjusting Extension Models to the Way Farmers Learn." In: *Learning for adoption: Technology adoption in developing country agriculture*. FERDI Discussion Paper

DellaVigna, S. and D. Pope. 2016. "What Motivates Effort? Evidence and Expert Forecasts." *Review of Economic Studies* 85(2):1029-1069

De Quidt, J., 2018. "Your Loss is My Gain: A Recruitment Experiment with Framed Incentives." *Journal of the European Economic Association* 16(2):522-559

Duflo, E., M. Kremer and J. Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *American Economic Review* 101(6):2350-2390

Dupas, P., 2014. "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment." *Econometrica* 82(1):197–228.

Ericson, K. and A. Fuster. 2011. "Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments." *Quarterly Journal of Economics* 126(4):1879–1907.

Feder, G., R. Just and D. Zilberman. 1985. "Adoption of Agricultural Innovations in Developing Countries: A Review." *Economic Development and Cultural Change* 33(2):255-298.

Fetene, G., S.Balew, Z. Abro, M. Kassie and T. Tefera.(2021. *Push-Pull Technology as a Climate-Smart Integrated Pest Management Strategy in Southern Ethiopia*. 2021 Conference, August 17-31, Virtual 315246, International Association of Agricultural Economists.

Foster, A. and M. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103(6):1176-1209

Fryer, R., S. Levitt, J. List, and S. Sadoff, 2018. *Enhancing the Efficacy of Teacher Incentives through Framing: A Field Experiment*. Cambridge MA: NBER Working Paper 18237

Gugissa, D., Z. Abro,T. Tefera. 2022. "Achieving a Climate-Change Resilient Farming System through Push‑Pull Technology: Evidence from Maize Farming Systems in Ethiopia." *Sustainability* 14, 2648. https://doi.org/10.3390/su14052648

Heffetz, O., and J. List. 2014. "Is the Endowment Effect an Expectations Effect?" *Journal of the European Economic Association* 12(5):1396–1422.

Hossain, T. and J. List. 2012. "The Behavioralist Visits the Factory: Increasing Productivity using Simple Framing Manipulations**."** *Management Science* 58(12):2151-2167

Imas, A., S. Sadoff, A. Samek. 2017. "Do People Anticipate Loss Aversion?" *Management Science* 63(5):1271-1284

Khan, Z., J. Pittchar, C. Midega, J. Pickett. 2018. "Push-Pull Farming System Controls Fall Armyworm: Lessons from Africa." *Outlooks on Pest Management* 29:220–224.

Kőszegi, B. and M. Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics* 121(4):1133–65.

Kassie, M., J. Stage, G. Diiro, B. Muriithi, G. Muricho, S. Ledermann, J. Pittchar, C. Midega and Z. Khan. 2018. “Push–Pull Farming System in Kenya: Implications for Economic and Social Welfare.” *Land Use Policy* 77:186-198

Kosfeld, M. and S. Neckermann. 2011. “Getting More Work for Nothing? Symbolic Awards and Worker Performance.” *American Economic Journal: Microeconomics* 3(3):86-99.

Lazear, E. 2000. “Performance Pay and Productivity**.”** *American Economic Review* 90(5):1346-1361

Levitt, S., J. List, S. Neckermann and S. Sadoff. 2016. “The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance.” *American Economic Journal: Economic Policy* 8(4):183-219

List, J., and A. Samek. 2015. “The Behavioralist as Nutritionist: Leveraging Behavioral Economics to Improve Child Food Choice and Consumption.” *Journal of Health Economics* 39:133-146

Magnan, N., D. Spielman, T. Lybbert and K. Gulati. 2015. “Leveling with Friends: Social Networks and Indian Farmers’ Demand for Technology with Heterogeneous Benefits.” *Journal of Development Economics* 116:223-251.

Midega, C., J. Pittchar, J. Pickett, G. Hailu, and Z. Khan. 2018. “A Climate-Adapted Push-Pull System Effectively Controls Fall Armyworm, Spodoptera frugiperda (J E Smith) in Maize in East Africa.” *Crop Protection* 105:10-15

Munshi, K. 2004. “Social Learning in a Heterogeneous Population: Technology Diffusion in the India Green Revolution.” *Journal of Development Economics* 73(1):185-213.

Ndayisaba, P., S. Kuyah, C. Midega, P. Mwangi, and Z. Khan. 2021. “Intercropping *Desmodium* and Maize Improves Nitrogen and Phosphorus Availability and Performance of Maize in Kenya.” *Field Crops Research* 26:108067.

Ndayisaba, P. S. Kuyah, C. Midega,P. Mwangi, and Z. Khan. 2020. “Push-Pull Technology Improves Maize Grain Yield and Total Aboveground Biomass in Maize-Based Systems in Western Kenya.” *Field Crops Research* 256:107911.

Pamuk, H., E. Bulte and A. Adekunle. 2014. “Do Decentralized Innovation Systems Promote Agricultural Technology Adoption? Experimental Evidence from Africa.” *Food Policy* 44:227-236

Pickett, J., C. Woodcock, C. Midega, Z. Khan. 2014. “Push-Pull Farming Systems.” *Current Opinions in Biotechnology* 26:125–132.

Shikuku, K., J. Pieters, E. Bulte and P. Laderach. 2018. “Incentives and the Diffusion of Agricultural Knowledge: Experimental Evidence from Northern Uganda.” *American Journal of Agricultural Economics* 101:1164-1180

Simonsohn, U. and G. Loewenstein. 2006. “Mistake #37: “The Effect of Previously Encountered Prices on Current Housing Demand.” *Economic Journal* 116(508):175–199.

Suri, T. 2011. “Selection and Comparative Advantage in Technology Adoption.” *Econometrica* 79(1):159-209

Tangney, J., J. Stuewig and D. Mashek. 2007. “Moral Emotions and Moral Behavior.” *Annual Review of Psychology* 58:345-372

Appendices

Table A1: Summary of baseline data for group leaders and balance tests

Baseline controls	(1) Control	(2) PIGF	(3) PILF	(4) SIGF	(5) SILF	t-test (1)-(2)	t-test (1-3)	t-test (1-4)	t-test (1-5)	t-test (2-3)	t-test (2-4)	t-test (2-5)	t-test (3-4)	t-test (3-5)	t-test (4-5)
household members (number)	5.5 [0.13]	5.6 [0.19]	5.6 [0.17]	5.5 [0.18]	5.8 [0.21]	-0.01	-0.01	0.00	-0.3	0.00	0.09	-0.16	0.05	-0.20	-0.3
household head age (years)	45.3 [0.7]	46.6 [1.2]	40.4 [0.83]	43.1 [1.1]	44.6 [1.0]	-1.3	4.9***	2.2*	0.75	6.2***	3.5**	2.0	-2.7**	-4.2***	-1.5
household head education (years)	6.7 [0.56]	6.4 [0.57]	5.3 [0.52]	6.4 [0.05]	5.8 [0.05]	0.36	1.4*	1.03	0.9	1.1*	0	0.6	-1.1*	-0.5	0.6
Primary activity farming (1=yes)	0.99 [0.01]	0.99 [0.01]	0.99 [0.01]	1.0 [0.00]	1.00 [0.0]	0.0	0.0	-0.01	-0.01	0.00	-0.01	-0.01	-0.01	-0.01	0.0
Lived in village (years)	11.9 [0.06]	11.98 [0.02]	11.97 [0.03]	12.00 [0.0]	11.97 [0.02]	-0.08	-0.07	-0.1	-0.07	0.01	-0.02	0.01	-0.03	0.00	0.03
Farm size (hectare)	0.06 [0.01]	0.06 [0.01]	0.06 [0.01]	0.07 [0.02]	0.07 [0.01]	0.00	0.00	-0.01	-0.01	-0.00	-0.01	-0.01	-0.01	-0.01	0.00
Knowledge of fall army worm (1= yes)	0.33 [0.03]	0.33 [0.05]	0.39 [0.05]	0.27 [0.05]	0.37 [0.05]	0.00	-0.06	0.06	-0.04	-0.06	0.06	-0.04	0.12*	0.02	-0.10
Knowledge of stem borer (1= yes)	0.06 [0.02]	0.08 [0.03]	0.05 [0.02]	0.13 [0.04]	0.11 [0.03]	-0.02	0.01	-0.07*	-0.05	0.03	-0.05	-0.03	-0.08*	-0.06	0.02
Credit constrained (1=yes)	0.36 [0.03]	0.38 [0.05]	0.27 [0.04]	0.33 [0.05]	0.33 [0.05]	-0.02	0.09	0.03	0.032	0.11	0.05	0.05	-0.06	-0.06	-0.00
Household owns milking cow (1=yes)	0.48 [0.04]	0.52 [0.05]	0.35 [0.05]	0.45 [0.05]	0.61 [0.05]	-0.05	0.13**	0.02	-0.13**		0.07	-0.08	-0.10	-0.26***	-0.15**

Note: ***p < 0.01, ** p < 0.05, and * p < 0.1 indicate significance levels.

.

Table A2: Summary of baseline data for follower farmers and balance tests

Baseline controls	(1) Control	(2) PIGF	(3) PILF	(4) SIGF	(5) SILF	t-test (1)-(2)	t-test (1-3)	t-test (1-4)	t-test (1-5)	t-test (2-3)	t-test (2-4)	t-test (2-5)	t-test (3-4)	t-test (3-5)	t-test (4-5)
Household members (number)	4.9 [0.08]	4.6 [0.13]	4.7 [0.11]	4.6 [0.12]	5.0 [0.11]	0.3*	0.2	0.3*	-0.1	-0.1	0.00	-0.4**	0.10	-0.3*	-0.4**
Household head age (years)	47.5 [0.57]	47.6 [0.9]	46 [0.75]	49 [0.8]	48 [0.8]	-0.1	1.5*	-1.5*	-0.5	1.6*	-1.4	-0.4	0.3	-0.2	-0.1
Household head education (years)	5 [0.02]	5.2 [0.03]	5.3 [0.03]	0.5 [0.03]	5.1 [0.03]	-0.2	-0.3	0.00	-0.1	-0.1	0.2	0.1	0.3	0.2	-0.1
Primary activity is farming (1=yes)	0.98 [0.01]	0.97 [0.01]	0.98 [0.01]	0.99 [0.01]	0.98 [0.01]	0.01	0.0	-0.01	0.0	-0.01	-0.02	-0.01	-0.01	0.00	0.01
Lived in the village (years)	11.98 [0.01]	11.98 [0.01]	12.0 [0.04]	12.0 [0.01]	12.0 [0.05]	0.00	-0.02	-0.02	-0.02	-0.02	-0.02	0.02	0.00	0.00	0.00
Farm size (hectare)	0.06 [0.01]	0.06 [0.01]	0.05 [0.01]	0.04 [0.01]	0.05 [0.01]	0.00	0.01	0.02	0.01	0.01	0.02	0.01	0.01	0.00	-0.01
Knowledge fall armyworm (1= yes)	0.15 [0.01]	0.18 [0.02]	0.14 [0.02]	0.15 [0.02]	0.20 [0.02]	-0.03	0.01	0.00	-0.05*	0.04	0.03	-0.02	-0.01	-0.06*	-0.05*
Knowledge stem-borer (1= yes)	0.23 [0.02]	0.24 [0.03]	0.27 [0.03]	0.26 [0.02]	0.29 [0.03]	-0.01	-0.04	-0.03	-0.06*	-0.03	-0.02	-0.05*	0.01	-0.02	-0.03
Credit constrained (1=yes)	0.25 [0.02]	0.26 [0.03]	0.23 [0.02]	0.23 [0.02]	0.28 [0.03]	-0.01	0.02	0.02	-0.03	0.03	0.03	-0.02	-0.00	-0.05	-0.05*
Household owns milking cow (1=yes)	0.36 [0.02]	0.38 [0.03]	0.36 [0.03]	0.31 [0.03]	0.34 [0.03]	-0.02	0.0	0.05	0.02	0.02	0.07*	0.04	0.05	0.02	-0.03

Note: ***p < 0.01, ** p < 0.05, and * p < 0.1 indicate significance levels. .

Table A3: Attrition analysis

Dependent variables	**Attrition status of group leaders**	**Attrition status of group members**
Models	**Probit**	**Probit**
Household members (number)	0.00 0.01	-0.01** 0.00
Household head age (years)	-0.00 0.00	0.00 0.00
Household Primary activity farming (1=yes))	-0.02 0.01	-0.01 0.01
Household head education (years)	0.01 0.05	-0.00 0.04
Lived in the village (years)	-0.06 0.07	-0.01 0.01
Agriculture land owned by household (hectare)	0.01 0.03	-0.00 0.01
Knowledge of stemborer (1= yes)	0.09 0.06	-0.01 0.02
Knowledge of fall armyworm (1= yes)	-0.16** 0.08	0.00 0.01
Household owned milking cow (1 =yes)	0.01 0.05	-0.00 0.01
Household faced credit constraint (1=yes)	0.06 0.04	-0.01 0.01
Private incentive reward framed as gain (PIGF)	0.01 0.09	0.00 0.02
Private incentive reward framed as loss (PILF)	0.01 0.07	0.02 0.03
Social prestige framed as gain (SIGF)	0.00 0.08	-0.00 0.02
Social prestige framed as loss (SILF)	0.05 0.10	0.02 0.03
Village effect	Yes	Yes
Mean dependent variable	0.07	0.04
Pseudo r-squared	0.17	0.06
Akaike crit. (AIC)	112	403
SD dependent var	0.25	0.19
Number of obs	571	1684
Bayesian crit. (BIC)	183	588
Note: ***p < 0.01, ** p < 0.05, and * p < 0.1 indicate significance levels.		

Table A4: Spillover analysis

Dependent variables	Lead farmers from kebele *without* incentivized leaders	Lead farmers from kebele with incentivized leaders	Difference	P values
Leaders organize at least one training event to train group members about PPT (1=yes)	0.38	0.48	-0.1	0.22
Group members received PPT information from leaders (Number)	2.44	2.56	-0.12	0.71
Leaders experimented with PPT on their maize plots (1=yes)	0.25	0.31	-0.06	0.41
Group members have knowledge of PPT (1= have full knowledge)	0.11	0.08	0.02	0.43
Group members experimented with PPT on their maize plots (1=yes)	0.03	0.04	-0.01	0.62

Chapter 3

Short-Run Subsidies and Long-Run Willingness to Pay: Learning and Anchoring in an Agricultural Experiment in Ethiopia

Abstract: We study how temporary provision of an agricultural innovation at zero cost affects long-run demand for that innovation. Our experimental design enables us to distinguish between an "anchoring effect" of subsidies and a "learning effect". We document large and persistent anchoring and learning effects. For the innovation that we consider, an integrated pest management (IPM) package for Ethiopian smallholder farmers, the learning effect dominates the anchoring effect, so temporary subsidized provision promotes long-run technology diffusion.

Keywords: free input provision, full subsidies, technology adoption, reference-dependent preferences

JEL codes: D91, O13

Publication Status: Balew, S., Bulte, E., & Kassie, M. (2025). "Short-run subsidies and long-run willingness to pay: Learning and anchoring in an agricultural experiment in Ethiopia." *American Journal of Agricultural Economics*, 107(2), 655-669.

3.1. Introduction

The adoption rate of technological innovations in important economic sectors lags far behind expectations in many low-income countries. This is due to various reasons (e.g. Foster and Rosenzweig 2010), one of which is the existence of liquidity constraints. Subsidies are often used to promote the adoption of innovations, and may enhance welfare if they increase access for poor beneficiaries, enabling them to learn about the benefits of the innovation (in the context of experience goods), or if adoption results in positive externalities. A large literature studies the *static* welfare effects of subsidies, some of which focuses on crowding-in or crowding-out behavior. Concerns are that subsidies may crowd in low-valuation users, which disrupts the allocative function of prices. Subsidies may also invite the "sunk-cost fallacy" effect, capturing the hypothesis that subjects use goods more carefully if they paid more to acquire them. Empirical evidence supporting these concerns is mixed and weak.[1]

Short-term subsidies can also affect long-term outcomes. They can increase future adoption if adoption relaxes liquidity constraints in follow-up periods, or if they influence beneficiaries' perception of the innovation's benefits through learning. Short-run subsidies can also affect long-term adoption via reference price dependence, or "anchoring". This occurs if deviations from a reference point invite "loss-gain utility" effects (Kőszegi and Rabin 2006). If respondents dislike paying more than they expected, based on past prices, then short-run subsidies may discourage long-run demand because follow-up prices appear "too high". A small set of studies examines such *dynamic* effects of subsidies due to learning and anchoring. Dupas (2014a) studies the long-term effect of a one-time bed net subsidy, and finds evidence of learning but not of anchoring—leading to a positive overall effect on long-run WTP. Similar findings are reported by Meriggi, Bulte and Mobarak (2021), who studied the deman[i]d for solar lamps, and Cai, de Janvry and Sadoulet (2020) who look at short-term subsidies for drought insurance premiums. Carter, Laajaj and Yang (2023) also find persistent positive impacts of temporary agricultural inputs subsidies on utilization. All these results are indicative of learning.

[1] It is not evident that subsidies undermine the screening effect of prices in a development context, and effects likely vary by product (e.g., Ashraf, Berry and Shapiro 2010; Cohen and Dupas 2010; Cohen, Dupas and Schraner 2015; Meriggi, Bulte and Mobarak 2021; Shukla, Pullabhotla, and Baylis 2022). Field experiments provide little or no empirical support for the sunk cost effect (Ashraf 2010; Meriggi, Bulte and Mobarak 2021).

The objective of this paper is to use an experimental design to study how a one-time subsidy for an agricultural innovation affects long-term willingness to pay (WTP) for that innovation and, more importantly, to measure the separate effects of the subsidy on learning and anchoring. Our two-stage randomization design enables us to obtain unbiased estimates of learning and anchoring effects for a subsample of respondents. Our study focuses on an integrated pest management (IPM) package that helps farmers to manage pests and weeds (see below). Respondents are unfamiliar with the IPM product, uncertain about its effectiveness, and have no prior experience with the product's pricing—facilitating anchoring on the subsidized offer price. We offer the IPM package free of charge to a randomly selected subsample of farmers, some of whom subsequently also receive an additional incentive to try the product on their farms. Without this extra incentive (a reward for experimenting with the technology), the actual usage of the subsidized product remains low. This reflects considerable opportunity costs of land and labor associated with adoption—the IPM package's purchase price is only part of the full private cost. Most farmers offered the package at zero cost chose not to use it. We estimate the local average treatment effect (LATE) by using the randomly assigned incentive to experiment with the IPM package as an instrumental variable for actual experimentation. We measure long-run WTP for the package through a follow-up Becker-DeGroot-Marschak (BDM) auction. Our results are based on a large randomized experiment involving 754 farmers.

Our main results are as follows. Free provision of the IPM package invites anchoring, which negatively affects WTP for the package afterward. However, hands-on experience with the technology fosters learning and updating of beliefs about the benefits of usage. The anchoring and learning effects "pull in opposite directions", but the latter dominates the former for the farmers experimenting with the package. The net effect of free provision on follow-up WTP for farmers who receive the subsidy and are predicted to use the package on their farms is positive. These findings nevertheless suggest that care should be taken when marketing new products through full subsidies if complementary costs are associated with adoption of the innovation. In the absence of additional interventions that invite experimentation, temporary full subsidies may reduce future demand. For the case of IPM management, which provides private benefits for adopters and some public benefits because of reduced pest diffusion, the case for sustained subsidies is unclear. Hence, evaluating the post-subsidy implications is important. More than three years after the intervention we still document sizable coefficients consistent with anchoring, but these are no longer significant at conventional significance levels—suggesting the anchoring effect gradually wears off or is dominated by other information.

This paper fits in a small literature that considers the impact of *full* subsidies on subsequent adoption or willingness to pay—offering new technologies at zero cost to beneficiaries. Shampanier et al. (2007) argue that zero-pricing invites a behavioral response from beneficiaries that is qualitatively different from charging very low positive prices. Three studies consider the reduced form (aggregate) effect of short-term full subsidies on long-run WTP—capturing both anchoring and learning. Fischer et al. (2019) document evidence consistent with anchoring; free provision of curative health inputs (pharmaceuticals) reduces future WTP. In contrast, Bensch and Peters (2019) study how free provision of improved wood stoves affects follow-up demand for these stoves, and find a positive effect. Similarly, Omotilewa, Ricker-Gilbert, and Ainembabazi (2019) find that full subsidies of an improved gain storage technology increases follow-up WTP. However, these studies cannot prove that anchoring does not exist. They suggest that anchoring, if it exists, is dominated by a positive learning effect so that (full) subsidies increase long-run adoption.

Our main contribution to this literature is that we use the first experimental design that enables unbiased estimation of the anchoring and learning effects associated with full subsidies for a new technology. Our two-stage design relies on fewer assumptions than two previous studies that aim to distinguish between and quantify anchoring and learning effects. Dupas (2014a) estimate the aggregate effect of subsidized bed nets against malaria and subsequently distinguished between anchoring and learning effects by making additional assumptions based on a parsimonious experience-good model. Shukla, Pullabhotla, and Baylis (2022) offer hermetic storage bags in stage 1 of their experiment and randomly assign farmers to one of three treatment arms: free provision of bags, a flat rate price, and a random strike prices (in a BDM auction). They subsequently compare willingness to pay for additional bags in stage 2. However, because farmers faced different prices across experimental arms in stage 1, adoption rates varied across these arms. It is possible that across experimental arms, “different types of farmers” purchased the bag. Any differences in WTP during stage 2 may therefore be driven by both an anchoring effect as well as a selection effect. Additional assumptions are needed to interpret the difference in bids across arms as a causal effect of the zero price. Our design requires fewer assumptions, but comes at the price that we can only make causal statements about a subsample of our respondents—so-called "complying farmers" who are induced to experiment with the new technology by the extra incentive provided (see below).

Other contributions are as follows. First, we focus on a specific type of innovation. IPM technologies like the one we study are labor intensive and require complementary inputs to implement. In that sense, the IPM package is like many agricultural innovations (e.g., think of improved seed or conservation agriculture), but unlike the other innovations that have been studied in this literature (malaria bed nets, PICS bags, solar lamps). As this literature grows, future researchers will hopefully be able to probe the robustness of anchoring and learning effects across different "types" of innovations (e.g. with and without complementary inputs or external benefits). Second, we measure follow-up WTP at two points in time, so our results also speak to the depreciation of anchoring and learning effects.

The data we use in this paper are from the same intervention reported by Balew et al. (2023). That paper focuses on knowledge diffusion and social learning among farmers. It documents that private and social (prestige) incentives, applied using gain and loss frames, are effective in motivating leader farmers to put more effort into training and instructing co-farmers (see below). The main dependent variable in the current paper—long-run WTP for the IPM package—was not part of the earlier paper. Indeed, our endline data were collected after the "knowledge diffusion" experiment was completed.

The paper is organized as follows. In section 2 we provide background information about the new technology and introduce the intervention. In section 3 we provide information about the experimental design, summarize the data, and sketch the identification strategy. Section 4 contains our results, and in section 5 we briefly touch upon a few implications for policy making. The discussion and conclusions ensue.

3.2. Background and Intervention

We study WTP for a package of agricultural inputs that helps farmers to manage pests that threaten their (maize) crops. In our study region and across Africa, farmers suffer from pests and weeds like the stemborer (SB), fall armyworm (FAW), and Striga.[2] The most common response is to

[2]The FAW emerged in 2016 in West Africa and has spread rapidly to other countries. It has developed into a major pest—causing significant crop damage and economic loss across the continent. Annual losses are estimated at 4-18

apply pesticides and herbicides, sometimes with adverse consequences for the environment and human health (Abro et al., 2021). A more ecologically sustainable approach to managing pest infestations is integrated pest management (IPM). Push-pull technology (PPT) is an example of IPM, based on intercropping the main crop (typically maize) with a combination of repelling and attracting plants. The repelling crop (*Desmodium*) is planted in strips between maize plants, producing chemical volatiles that drive out pest species.[3] An attracting crop (Napier grass or *Brachiaria*) is grown along the plot's edges to attract and "trap" pest insects and prevent stemborer larvae from developing (Khan et al., 2008). PPT provides additional benefits to adopting farmers. *Desmodium* and *Brachiaria* provide high-quality fodder for livestock and contribute to soil improvement—*Desmodium* is a leguminous species that fixes atmospheric nitrogen. However, adopting PPT requires labor to plant and manage companion crops. Companion crops also involve an opportunity cost as they reduce land available for the main crop. *Desmodium* and *Brachiaria* are perennial crops and can be used for up to 3-5 years.

Although PPT was developed some time ago, its adoption rate remains low. The external inputs are not available in most regions. Most farmers have an imperfect understanding of both the benefits and the costs of adoption, and diffusion of PPT is slow and characterized by (social) learning among the target population. Promoting its diffusion has become a priority for research organizations, policymakers and extension agents throughout Africa (Kassie et al., 2018). In rural Ethiopia, for example, extension workers aim to promote diffusion via the so-called *one-to-five* extension system (Balew et al., 2023), where farmers are encouraged to organize themselves into groups of six and select a group leader. This leader is trained by extension experts, and offered the necessary inputs when necessary (perhaps at a subsidized price). Next, the leader should share his experiences with his followers—the other five group members.

3.3. Sampling, Experimental Design, and Data

We study WTP for push-pull technology to prevent pest damage for a sample of one-to-five group leaders in the Jabi Tehnan district, in north-western Ethiopia. One-to-five groups are an existing extension institution in Ethiopia, and we worked with existing group leaders. In what follows we

---

million tons of maize, which is valued between US$2.5 and US$6 billion; sufficient to feed 40 million to 100 million people (FAO, 2020).

[3] *Desmodium* also suppresses *Striga* through allelopathic mechanisms, improving maize yields (Khan et al. 2008).

use the words (group) leader and farmer interchangeably. We obtained IRB approval for the experiment and informed consent from respondents and local officials. We implemented a cluster randomized experiment in 38 kebeles (villages), or 114 sub-kebeles. After conducting a census of the one-to-five groups in the sub-kebeles we randomly selected approximately five groups per sub-kebele for participation in the experiment. The total sample consists of 754 group leaders. All leaders participated in a two-day training session in their villages, offered by the International Centre of Insect Physiology and Ecology (icipe) and Jabi Tehnan district agricultural office experts. They received training on PPT implementation (correct spacing, row planting, timely weeding and harvesting of companion crops) and learned about its benefits.

After participating in the training, group leaders were randomly assigned to experimental arms.[4] To reduce the risk of spillovers we used a cluster randomized design, and assigned sub-kebeles to treatment arms. We used a two-stage randomization design. First, we assigned 28 sub-kebeles (183 leaders) to the "pure control" arm and 86 sub-kebeles (571 leaders) to receive treatment. Leaders from the pure control group could buy a package sufficient for 0.2 hectares for a price of 2,800 ETB (=USD50.1). They could also purchase smaller packages at proportionally reduced prices. Instead, all treatment farmers were offered a zero-price input package, enabling them to implement PPT on part of one of their maize plots. These starter packages consisted of 500 grams of *Brachiaria* and *Desmodium* seed, sufficient for intercropping 0.2 hectares (the average size of a maize plot in our sample is 0.43 hectares).

Because of labor and land opportunity costs, we expected that zero-pricing was unlikely to persuade all farmers to adopt PPT on their plots. Farmers selected for treatment were next assigned to sub-arms with different incentives to adopt and experiment with the PPT package (and share lessons learned with group members). In sub-arm 1 farmers received no additional incentive, other than the free starter package. In the four remaining sub-arms leaders received a reward in case of "good performance", where performance was measured in terms of follower knowledge about IPM (see online appendix A). To effectively disseminate knowledge to their followers, leaders should use PPT on their farms and use these experimental (sub)plots as the basis for instruction

[4] Since leaders learned about their assignment after participating in the training, learning effort during the training should be identical across arms.

and joint learning. In sub-arms 2 and 3, leaders received a private material reward for experimenting and sharing information (a sickle). In sub-arm 2 sickles were promised as bonuses in case of good performance (gain frame), and in sub-arm 3 sickles were distributed up-front and “clawed back” in case of under-performance (a loss frame to leverage loss aversion and increase performance—see Bulte, List and van Soest 2020). In sub-arms 4 and 5 leaders received a social recognition reward in case of good performance. A certificate of good performance was awarded during a public ceremony, again either as a bonus or as an up-front reward. Following Balew et al. (2023), these 4 sub-arms of the Incentive treatment are labelled PIGF (private incentive gain frame), PILF (private incentive loss frame), SIGF (social incentive gain frame) and SILF (social incentive loss frame).

Details of these sub-arms and the incentivizing effects of the various treatments on experimentation are provided in Balew et al. (2023). For the current project the distinction between sub-arms is not important. What matters is that (i) a random sample of group leaders was offered a package of agricultural inputs at zero cost (subsidy-only), and (ii) a random subsample of these leaders was subsequently incentivized to experiment with these inputs on their plots (subsidy-plus-incentive). The design and details of are summarized in Figure 1.

Figure 1: Summary of experimental design

<table>
<tr><td colspan="6">Sample frame:
114 sub-kebeles (SK), 754 lead farmers (N)</td></tr>
<tr><td>Pure control;
SK=28, N=183</td><td colspan="5">Treatment; SK=86, N=571</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">Only subsidy;
SK=28,
N=201</td><td colspan="2">Subsidy + private incentive (sickle) for experimentation with PPT:
SK=29 N=190, n=534</td><td colspan="2">Subsidy + social prestige incentive (certificate) for experimentation with PPT:
SK=29, N=191, n= 574</td></tr>
<tr><td>Sickle as a bonus
SK=14, N=91
(PIGF)</td><td>Sickle up-front
SK=15,
N=102 (PILF)</td><td>Certificate as bonus
SK=14, N=86
(SIGF)</td><td>Certificate up-front
SK=15, N=91
(SILF)</td></tr>
</table>

We used well-trained enumerators to collect three waves of panel data for leaders, using tablets. Baseline survey data were collected before the training sessions in June–August 2018. We interviewed all group leaders. We collected data on household demographics, food security, crop and livestock production, plots owned and managed, exposure to and awareness of maize field pests (including SB, FAW, and Striga), pest control management, and sources of agricultural information. Online appendix Table B1 summarizes these data and provides t-tests to verify balance across the control and treatment groups. On average, leaders are 44 years old, have a family of 5.6 persons, and attended almost 6 years of formal education. Nearly all leaders are male, and the main activity is crop farming. Half of the leaders own a milking cow. Not surprisingly, most farmers are aware of the pests that threaten their crops, and recognize them as major constraints on production (especially SB and Striga, which are established pests). Overall, the sample is balanced across arms. It appears as if the invasive weed *Striga* is more of a problem (and better known) in control than in treatment sub-kebeles, but differences are small. In some estimation models we include baseline controls to increase the precision of our estimates and control for pre-existing differences.

We collected follow-up data after the next growing season in November 2020 (midline) and again in February 2022 (endline). We measured WTP for PPT among group leaders using a Becker-DeGroot-Marschak (BDM) auction mechanism (see below). Bidding in the BDM was the only way for farmers to obtain the package—we stopped offering packages after the baseline, and inputs were not for sale via agro-dealers. Our main results are based on the bidding behavior of leaders. The subsidy could make them "anchor" on the zero price. The incentive encouraged experimentation and learning about the technology's benefits.

We managed to revisit 736 group leaders during the midline, so 18 farmers were unavailable during the follow-up. During the endline, no further attrition occurred, and we collected data for all remaining 736 leaders. We provide an attrition analysis in online appendix Table B2, regressing attrition status on baseline controls and treatment arm dummies. Attrition is not correlated with most of our variables and also not with treatment status. The *p*-value of the F-test of joint significance reveals that we cannot reject the null-hypothesis.

Our midline and endline data confirm that zero-pricing of PPT inputs is not sufficient to convince all leaders to adopt and use the package. Overall, out of 558 leaders offered the subsidized package, only 212 leaders adopted and experimented with the package (between training and midline). Adoption was low among unincentivized leaders assigned to the full subsidy. Out of 193 unincentivized leaders, 57 leaders adopted (29.5%). From the subsample of 365 incentivized leaders, 155 farmers adopted, for an adoption rate of 42.5%. These adoption rates are significantly different (p<0.01), providing us with the leverage we need to treat random assignment to the incentive as an instrumental variable for experimentation (see below). For details about how adoption rates vary across the four incentivized sub-arms, refer to online appendix A and Balew et al. (2023). The adoption rate among leaders assigned to the pure control group was 0%.

In our regression analysis we focus on explaining variation in WTP for (additional) packages of PPT inputs, at midline and endline, across the experimental groups of farmers. We elicited WTP for a *Brachiaria* and *Desmodium* seed package using the BDM auction mechanism (see Burchardi et al. 2021). Recall that *Brachiaria and Brachiaria* are perennial crops, so farmers who adopted the PPT package would purchase seed to expand the share of their plot covered by PPT, or use it elsewhere on their farm. The package of seeds that we auctioned off was enough to cover 50% of a plot with average size. For our sample, the average maize plot is 0.43 hectares, and the average farm size equals 1.27 hectares (farmers have multiple plots). Before bidding, all farmers received information about the costs and benefits of PPT, repeating information from the training.

BDM is an incentive-compatible method to measure WTP. Participants submit an offer price to purchase the product, after which a sale price (or strike price) is revealed to the participants. Bidders stating bids higher than the strike price can buy a unit of the good and pay an amount equal to the strike price. Bidding below the strike price implies that the bidder cannot obtain the good. Since a farmer's bid determines whether she has the right to buy the good, but not the price paid, farmers have an incentive to state their true value. Farmers were asked to state the area they wanted to allocate to IPM (in local measurement units called *temad*), and how much they wanted to pay for a package of inputs allowing them to do this. These units were converted into a per-hectare price bid, which we compared to the per-hectare strike price. Almost all farmers who stated a bid expressed a willingness to buy a package that enabled them to cover a fraction of one hectare (likely reflecting binding liquidity constraints).

After elaborate instructions and a (hypothetical) trial run, we collected the bids under conditions of privacy (the BDM was part of the survey instrument). The uniform strike price we used at midline was 2,800 ETB per hectare (=USD 50.1), the same price used when offering the package to control group leaders (which is not the same as the market price—there is no market for the IPM package in the study region yet). Farmers bidding more than the strike price received their packages at their houses two weeks later. Unfortunately, we did not collect data on non-compliance, or the number of farmers who changed their minds or could not produce the cash necessary to purchase the package (see also Meriggi et al. 2021). We believe there are two reasons why this number may be small. First, farmers were free to pick the area they wanted to allocate to the IPM package, and capital-constrained farmers could bid for an arbitrarily small subplot. Second, farmers had two weeks to accumulate the cash for purchasing the package.

Our main analysis is based on 2sls models. These models enable distinguishing between a learning and anchoring effect, and gauging the relative magnitudes of these effects. This requires predicting which subsample of the farmers experiment with (and learn about) the innovation. In the first stage, we use the incentive to experiment with the PPT package (*Incentive*) as an (excluded) instrument for adoption and experimentation (*Experiment*), and use assignment to the full subsidy (*Subsidy*) as an included instrument. The variable *Experiment* is defined as adopting the PPT package on (part of) the farmer's plot, which we verified in the field. To be clear, we do *not* regress experimentation on different treatment groups—we do not regress *Experiment* on subsidy-only and subsidy-plus-incentive dummies. Instead, the dummy variable *Subsidy* takes a value of one for all treated farmers—farmers from the subsidy-only group and the subsidy-plus-incentive group—so *Incentive* picks up the additional effect of the conditional reward. In the second stage, we use predicted *Experiment* and *Subsidy* to explain variation in auction bids at midline and endline. Receiving the full subsidy is an included instrument, and enters both in the first and second stage of the model. Reflecting that subsidies may encourage experimentation and induce anchoring. The model reads as follows:

$$Experiment_{jsk} = \alpha_k + \beta Incentive_{sk} + \gamma Subsidy_{sk} + \delta \boldsymbol{X}_{jsk} + e_{jsk}, \quad (1)$$

$$WTP_{jsk} = \theta_k + \mu Experiment^{*}_{jsk} + \varphi Subsidy_{sk} + \pi \boldsymbol{X}_{jsk} + \varepsilon_{jsk}. \quad (2)$$

In (1)-(2), farmers from the pure control group are the omitted category, $\boldsymbol{X}_{jsk}$ is a vector of controls, e_{jsk} and ε_{jsk} are error terms, and the subscripts *j*, *s*, and *k* refer to, respectively, leader *j* in sub-*kebele s* in *kebele k*. The terms α_k and θ_k are *kebele* fixed effects, capturing variation in agroecological, market and governance conditions. For the first stage we expect both the incentive and subsidy will promote experimentation by leaders (β>0 and γ>0, respectively). Substantively we are interested in the coefficients μ and φ estimated during the second stage/ Based on the discussion above we expect $\mu>0$ (learning, assuming the IPM package performs better than expected *ex ante*) and $\varphi<0$ (anchoring on the full subsidy). We cluster standard errors at the sub-*kebele* level.

Our identification strategy rests on the assumption that *Incentive* is a valid instrument for *Experiment*. Random assignment implies our instrument is exogenous, and we will verify below that it is sufficiently strong. The exclusion restriction should also be satisfied, or we should be confident that the incentive does not affect WTP through any other channel than the effect on the propensity to experiment with the IPM package. This assumption can be questioned, and we discuss and allay possible challenges to unbiased identification. First, the incentive to try the technology could be perceived as a negative price, thus also inducing some anchoring. Alternatively, farmers may perceive the incentive as a signal that the technology is "good" and worth trying, which could affect bid amounts directly. However, we framed the incentive as a (conditional) reward for the time and effort that the group leaders allocated to teaching other members of their one-to-five group about the innovation—not as an additional subsidy for the input package. Receipt of the reward was entirely conditional on how successfully the farmers managed to diffuse knowledge to their group members. Second, farmers who received a sickle as an incentive to experiment may be willing to pay more later because this productive input may make them better able to implement the IPM technology (thus raising bids). Again, we are not very concerned about this potential threat to identification because farmers in the other treatment groups also have access to simple technologies such as sickles. Moreover, we find qualitatively similar results if we focus on farmers incentivized with a certificate of good behavior and omit farmers incentivized with a sickle (not shown, but available on request). It is hard to imagine how the certificate could affect follow-up WTP for the IPM package, other than through the incentive it implies to experiment with the technology.

In addition to estimating local average treatment effects (LATEs), we also report intention to treat (ITT) effects. This involves regressing WTP on the two treatment groups (subsidy-only and subsidy-plus-incentive). However, partial compliance implies that these ITT estimates will confound anchoring and learning effects. The ITT effect for the subsidy-only group includes the anchoring effect and a learning effect (for the 29.5% of the farmers in this group who actually adopt).[5] Similarly, the ITT for the subsidy-plus-incentive group captures both the anchoring effect and a learning effect for the 42.5% of farmers from this group who adopt. If farmers anchor on the zero price, then the anchoring effect reduces follow-up WTP. If the IPM package works better than expected, then the learning effect will increase follow-up WTP. Capturing both opposing effects in one estimate is not necessarily very informative if the aim is to learn about learning and anchoring.

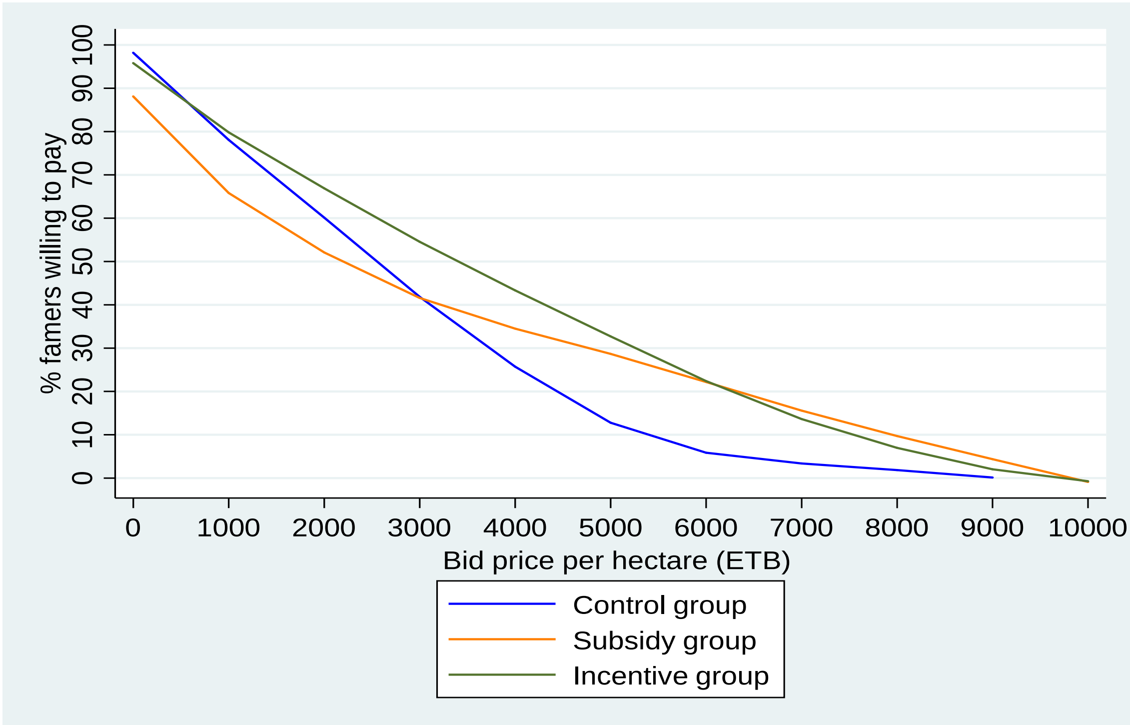
3.4. Results

Before presenting our regression results we describe our BDM-based willingness to pay data. At midline (endline), 18% (15%) of the farmers "opted out" of the BDM, and refused to state a bid—either because they feared they were unable to accumulate any cash to purchase the package, or because they believed the technology would not work for them. We did not try to persuade farmers to state a bid as the implementation protocol prescribed that farmers were free at all times to "opt out" of the experiment. The share of farmers refusing to bid does not vary across treatment arms (online appendix Table B3). In what follows we use the full sample of farmers in our analyses, imputing a zero bid value for farmers who opted out of the BDM. We also demonstrate robustness for the sub-sample of farmers who participated in the BDM—dropping the farmers who opted out and focusing on the subsample of leaders placing a bid.

Average bid amounts for the full sample varied quite a bit across treatment arms. For example, the average per-hectare bid across all respondents was 2,740 ETB per hectare (USD 49) at midline, close to the average bid of "pure control group" farmers at midline (2,793 ETB, or USD 49.4). As mentioned, this might seem like a large number, but during the auction farmers always bid for areas *smaller than one hectare* and prices paid were scaled down accordingly. This average bid hides considerable heterogeneity. Farmers from the "subsidy only" group bid only 2,449 ETB, on average (=USD 43.8), and average bids for the incentivized sub-arms ranged from 2,585 ETB (USD 46.2) to 3,135 ETB (USD 56.0), which reflects variation in the share of leaders experimenting with the package. Similar patterns in the leader data exist at the endline, where bids were typically higher.

[5] This means an aggregate effect is estimated, as in Omotilewa, Ricker-Gilbert, and Ainembabazi (2019) and Bensch and Peters (2019).

Figure 2: Demand for the PPT package of the three experimental groups (endline)



More can be learned about bidding behavior from Figure 2, which provides curves for the three experimental groups at the endline—pure control farmers ("*Control group*"), subsidy-only farmers ("*Subsidy group*"), and subsidy-plus-incentive farmers ("*Incentivize group*"). In Figure 2, WTP is on the x-axis and the percentage adoption is on the y-axis, so these curves are akin to, but not quite the same as, (inverse) demand curves. Since there are adopters and non-adopters in the subsidy-only group and in the subsidy-plus-incentive group, we cannot isolate anchoring and learning effects by eye-balling demand curves. However, because the share of adopters varies across groups, the diagram is illustrative.[6] Two features stand out. First, roughly equal shares of farmers from the subsidy-only group and subsidy-plus-incentive group are stating "high bids" (6,000-10,000 ETB, or USD 107-USD 179), and there are virtually no farmers from the pure control group bidding similar amounts. This likely reflects learning by some participants from the treatment groups who tried out the package. Second, a small share of farmers from the subsidy-

[6] Observe that we cannot distinguish between groups of farmers based on their choices (i.e. draw separate demand curves for farmers who did and did not adopt the PPT package). Because adopting farmers are likely to be different from non-adopting farmers, such an approach would introduce selection bias.

only group is bidding low or intermediate sums, compared to farmers from the subsidy-plus-incentive group (<6,000 ETB). Indeed, fewer farmers from the subsidy-only group are bidding between 500 and 3,000 ETB than from the pure control group (USD 8.9-USD 53.5). The share of farmers from the subsidy-only group bidding zero is 53%--significantly greater than from the other groups. This likely reflects anchoring by participants who were offered the subsidy but did not try out the package.

This discussion can be made more precise. Table 1 summarizes the regression results for leaders at midline and endline. Columns (1-3) and (5-7) are based on the full sample of 736 leader farmers, where we imputed a zero-bid for the leaders who opted-out and refused to state a bid. Columns (4) and (8) are based on the subsample of farmers participating in the BDM (several of which were zero bids as well). The qualitative results are rather similar across the groups, but not identical.

Bottom Panel B summarizes the results of the first stage of the 2sls model. While the subsidy invites experimentation, adoption in the subsidy-only arm is far from complete. The randomized incentive pushes up the adoption rate, and is a good instrument for experimentation. It is independent by design and plausibly exogenous—there is no reason to believe that the reward will affect follow-up bidding behavior other than through encouraging experimentation (learning). The instrument is also relevant—it enters significantly and the partial F values appear sufficiently large (close to 10, or larger, often used as a rule of thumb). We obtain qualitatively similar results when using an aggregate reward variable to capture incentivizing experimentation (columns 1-2, 5-6) and the four separate reward variables (columns 3-4 and 7-8). The incentive to experiment encourages experimentation, as intended.

More interesting results are summarized in Top Panel A. The omitted category is farmers from the pure control group—who received training about the benefits of PPT but were not offered the package at zero cost (and who, hence, did not experiment with the package). Consider the midline results, in columns (1-4), and recall that the average bid at the midline of control group leaders equaled 2,793 ETB (USD 49.9). The impact of anchoring and learning are economically substantive and statistically significant. The first insight is that leaders who received the subsidized (free) PPT package are willing to pay much less, all else equal, at midline than control group farmers. According to column (1), subsidies cause a drop in WTP equal to 40% of the WTP of control group leaders.

Table 1: Anchoring, Learning and Willingness to Pay of Group Leaders (2sls)

	Leaders at midline				*Leaders at endline*			
Panel A	*Second stage(WTP for PPT package as dependent variable)*							
Experiment	3627** (1533)	3367** (1432)	3465*** (1413)	3627* (2277)	3895** (2103)	3921** (1976)	4212** (1808)	3328 (2245)
Subsidy	-1556** (741)	-1426** (626)	-1466** (620)	-2019* (1277)	-1143 (1033)	-1137 (887)	-1253 (856)	-941 (1255)
Village FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	Yes	No	No	Yes	Yes	No
Constant	2592***(150)	2733* (1561)	2732* (1572)	3206*** (352)	2967*** (568)	5077** (2587)	5073** (2592)	3392*** (572)
Panel B	*First stage(Experiment with PPT package as dependent variable)*							
Subsidy	0.34*** (0.050)	0.28*** (0.045)	0.28*** (0.045)	0.45*** (0.054)	0.34*** (0.050)	0.28*** (0.045)	0.28*** (0.045)	0.42*** (0.050)
Incentive	0.17*** (0.049)	0.18*** (0.046)	-	-	0.17*** (0.049)	0.18*** (0.046)	-	-
PILF	-	-	0.21*** (0.065)	0.18** (0.072)	-	-	0.21*** (0.065)	0.22*** (0.074)
PIGF	-	-	0.15** (0.059)	0.11 (0.067)	-	-	0.15** (0.059)	0.14** (0.070)
SILF	-	-	0.18** (0.068)	0.10 (0.077)	-	-	0.18** (0.068)	0.15** (0.073)
SIGF	-	-	0.17*** (0.062)	0.14** (0.066)	-	-	0.17*** (0.062)	0.16** (0.065)
Village FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	Yes	No	No	Yes	Yes	No
Constant	-0.056 (0.040)	0.057 (0.30)	0.060 (0.30)	0.15 (0.54)	-0.056 (0.040)	0.058 (0.30)	0.060 (0.30)	-0.009 (0.029)
R2	0.28	0.35	0.35	0.34	0.28	0.30	0.35	0.33
Partial F	3825	25.87	22.95	2669	20027	25.87	22.95	562
N	736	736	736	602	736	736	736	629

Notes: The table reports results of a 2sls model where experimentation with the PPT package is instrumented with a random incentive (reward) to experiment. Columns (1-4) capture willingness to pay for a PPT package during midline, and columns (5-8) capture willingness to pay at endline. In columns 1, 2, 5 and 6 we use an aggregate incentive variable. In columns 3, 4, 7 and 8 we use disaggregated incentive variables (material and non-material rewards, paid as a bonus or up-front—see Balew et al. 2023). Columns (1-3) and (5-7) are based on the full sample of group leaders, where we included a zero value for farmers who refused to state a bid during the BDM. Columns (4) and (8) are based on a subsample where we dropped farmers who refused to state a bid. Covariates included: *Sex of respondent*, *Marital status*, *Family size*, *Age*, *Education*, *Primary activity is farming*, *Lived in the village*, *FAW knowledge*, *FAW major constraint*, *SB knowledge*, *SB major constraint*, *Striga knowledge*, *Striga major constraint*, *Maize farm area*, *Total land size*, *Milking cow*, *Credit constraint*. Standard errors clustered at the sub-kebele level in parentheses. *, ** and *** indicate significance at 10%, 5% and 1%.

The second insight is that experimentation with PPT induces learning and increases willingness to pay. The LATE, or the effect of experimentation for leaders tempted to experiment because of the incentive, is large—exceeding the mean WTP of control group farmers. This is consistent with evidence that the PPT technology is on average very beneficial for adopting farmers (see also Kassie et al., 2018, and the discussion below). According to our results, the technology is doing much better than farmers anticipated.

The overall effect of subsidies and incentivizing on follow-up willingness to pay is positive.[7] The learning effect dominates the anchoring effect. However, while the anchoring-on-zero effect is likely rather generic (e.g., Fischer et al. 2019; Shukla, Pullabhotla, and Baylis 2022), the learning effect must be technology- and context-specific. This implies that combining subsidies and incentives for experimentation will not generally provide an impetus to the adoption of new technologies—this will depend on conditions. If the learning effect is small (or negative, for innovations performing worse than expected), so that the anchoring effect dominates the learning effect, then short-term subsidies reduce long-term demand for the new technology.

Columns (5-8) consider the impact estimates at the endline, some 3.5 years after the PPT package was offered to participants. These results are more difficult to interpret than those in columns (1-4). First, WTP at the endline may not only be affected by the initial offer price of zero ETB, but farmers may also anchor on the strike price of 2800 ETB (USD 50.1) that we used during the midline BDM auction. It is not apparent on which price farmers should anchor here. Moreover, the omitted category now consists of "pure control group" leaders at endline, who have also participated in auctions at midline, which could affect their follow-up bidding behavior. For this reason, the coefficients in columns (5-8) should not be directly compared to those in columns (1-4). With these caveats in mind, it is interesting to observe that the patterns in the endline data are qualitatively rather similar to patterns in the midline data (even if they are less significant for

[7] We do not report the full regression models with all covariates. But some effects are noteworthy. We do not find that farmers with more land (wealthier farmers) have higher WTP, or that households with larger families (who are perhaps better able to mobilize the complementary labor input) have higher WTP. The *p*-values of total land size and family size were not even close to being significant (p=0.84 for land size and p=0.96 for family size). We believe this reflects that farmers were free to choose any quantity of IPM inputs that they wanted, so families with less land or less labor could simply make a bid for a smaller quantity (and offer a statistically indistinguishable price per hectare). We do find that households who own livestock express higher bids, reflecting that the companion crops can be used as animal feed.

anchoring). Experimentation increases WTP and we find consistently negative coefficients for the baseline subsidy variable. We interpret this as tentative evidence that the effects of anchoring and learning are relatively persistent and can carry across multiple growing seasons (albeit probably true that the anchoring effect becomes smaller over time).

ITT results are summarized in Table 2. As expected, these models capture both the learning and (opposing) anchoring effect, and produce estimates of the treatment effects that are not very different from zero. Specifically, only a quarter of the estimates of the reduced form treatment effect is significantly different from zero (7 out of 28 coefficients).

In Table B5 we follow up on the data presented in Figure 2 and report the results of quantile regressions (the 35th and 65th percentile) for the same ITT model. This allows verifying whether the effect of the subsidy on WTP varies across the distribution of bids. We document negative and positive effects for different people, depending on their position in the distribution of the population. The anchoring effect of receiving a subsidy dominates at the 35th percentile, but not at the 65th percentile (where it is offset by learning). The learning effect of incentives on WTP seems particularly large at higher percentiles of the bidding distribution, presumably reflecting that these are the farmers who started experimenting with the innovation.

Table 2: Anchoring, Learning, and Willingness to Pay of Group Leaders (ITT estimates)

	Leaders at midline(OLS)				*Leaders at endline(OLS)*			
Subsidy-only	-333 (295)	-375 (296)	-345 (295)	-399 (348)	170 (414)	115 (402)	133 (289)	432 (435)
Subsidy plus incentive	616** (245)	598** (252)	-	-	661 (401)	678* (393)	-	-
PILF	-	-	798** (333)	576 (408)	-	-	906** (463)	720 (548)
PIGF	-	-	554* (291)	479 (365)	-	-	364 (515)	519 (547)
SILF	-	-	436 (373)	479 (367)	-	-	787* (452)	533 (524)
SIGF	-	-	519 (396)	604 (374)	-	-	604 (499)	508 (544)
Village FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	Yes	No	No	Yes	Yes	No
Constant	2390*** (136)	2903** (1619)	2859** (1636)	3250** (1498)	2750*** (633)	4967*** (2808)	5009** (2807)	3351*** (536)
R2	0.070	0.133	0.135	0.099	0.140	0.210	0.212	0.173
N	736	736	736	602	736	736	736	629

Notes: The table reports results of ITT estimates. Columns (1-4) capture willingness to pay for a PPT package during midline, and columns (5-8) capture WTP at endline. In columns 1, 2, 5 and 6 we use an aggregate incentive variable. In columns 3, 4, 7 and 8 we use disaggregated incentive variables (material and non-material rewards, paid as a bonus or up-front—see Balew et al. 2023). Columns (1-3) and (5-7) are based on the full sample of group leaders, where we included a zero value for farmers who refused to state a bid during the BDM. Columns (4) and (8) are based on a subsample where we dropped farmers who refused to state a bid. The covariates included are *Family size*, *Age*, *Education*, *Primary activity is farming*, *Lived in the village*, *FAW knowledge*, *FAW major constraint*, *SB knowledge*, *SB major constraint*, *Striga knowledge*, *Striga major constraint*, *Maize farm area*, *Grazing area*, *Milking cow*, *Credit constraint*. We report standard errors clustered at the sub-kebele level in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1%, respectively.

Finally, in online appendix Table B6 we report the results of an explorative heterogeneity analysis. We are interested in exploring whether the anchoring and learning effects are different for “large” and “small” farmers, where large or small may be defined in terms of farm size or family size. We split the sample of farmers in sub-groups, based on median farm size and median family size, and repeat the 2sls analyses for the subsamples. While the first-stage results are robust across the various sub-samples of respondents, we find that the second stage results vary from sub-sample to sub-sample. Specifically, it appears as if the results reported in Table 1 are mainly driven by “large farmers”—respondents with either a large farm or large family. These farmers have more money, more labor and more land in which to experiment with and use the labor-intensive technology.

3.5. Lessons for Policy

Based on our main findings we infer two insights for policy makers. First, the combination of subsidies for IPM packages and incentives for experimentation appears to increase overall welfare. According to Kassie et al (2018), farmer net profits per acre for adopting farmers in nearby Kenya equal approximately 110 USD/acre (2019 exchange rate), taking into account changes in production cost and crop displacement. Adopting this estimate, and assuming a lifetime of 3 years for our perennial companion crops, the (undiscounted) income gain across the three year period amounts to approximately USD 825 per hectare. We estimate that the sum of production, delivery and training cost of the IPM package amounts to approximately USD 50-150 per hectare (a cost that will come down if production is scaled up). The cost of a sickle as an additional incentive to promote experimentation amounts to about USD 5. It is evident that the estimated benefits from adoption outweigh the costs. In the absence of liquidity and other constraints, social learning may foster diffusion of IPM as a market-mediated solution to mitigate the adverse effects of increasing pest pressure. In the absence of subsidies and incentives, however, the diffusion of the adoption might not occur—recall that the adoption rate in the pure control group was exactly zero percent.[8] Our results suggest subsidies would be justified to overcome the various adoption constraints that farmers face. A significant extension effort to build awareness may also increase welfare. A pre-condition is that IPM-packages should be supplied by local agro-dealers (currently the technology is not yet available in the study area), so policy makers should support this process of input market development as well.

The second insight is more generic, and concerns how policy makers should subsidize goods for poor people. Subsidies may be welfare-enhancing if poor people cannot afford to purchase essential goods, if consumption of goods involves positive externalities, or if (short-term) subsidies enable beneficiaries to experiment with and learn about certain goods. Dupas (2014b) summarizes this debate and concludes that "*for products with large social benefits, free distribution is the most cost-effective strategy for increasing coverage of essential health products and services*". An important underlying reason for this conclusion is that demand for goods tends

[8] From Figure 2 can be read off that some 55-60% of the farmers from the subsidy plus incentive group was willing to pay the price of ETB 2800 (USD 50) for the IPM package. The adoption rate in the control group was about 40%.

to be very price elastic, even at low prices. Hence, providing large subsidies that are not full subsidies (i.e., low positive prices, but no free provision) often fails to include a significant share of the target population. However, this recommendation may need to be rewritten in light of the new results on anchoring and full subsidies (Fischer et al. 2019; Shukla, Pullabhotla, and Baylis 2022; this paper).

Anchoring on full subsidies implies that providing products for free may not be the best approach to sustainably promoting the uptake of new products. While providing full subsidies may be optimal if subsidy flows are sustained for extensive periods (e.g. in the case of goods with large positive externalities), they may not be if the aim is to provide subsidies temporarily, after which allocation by markets should become dominant (e.g. in the case of learning about new goods with mainly private benefits). In the latter case, the future benefits of positive pricing (no anchoring) should be weighed against the immediate costs of positive pricing (exclusion of potential beneficiaries). Very small prices may be optimal then.[9] Of course it would also be useful to explore whether it is possible to reduce the extent to which beneficiaries anchor on subsidized goods, for example by packaging subsidized and follow-up market-mediated goods differently.

3.6. Discussion and Conclusions

To promote the diffusion of new technologies in low-income countries, policy makers often resort to temporary subsidies that enable targeted households to try out the innovation and update their beliefs about its costs and benefits. The theory of reference-dependent preferences suggests that short-run subsidies may have long-run effects if they induce anchoring on subsidized prices. The empirical literature based on experiments in the field does not provide much support for this concern (e.g., Dupas 2014a; Meriggi, Bulte and Mobarak 2021). However, a handful of studies suggest that there might be “something special” about zero-pricing, or full subsidies of innovations. Zero prices are particularly salient for beneficiaries and have been shown to induce anchoring—shifting the long-run demand curve for the good inwards (Fischer et al. 2019; Shukla, Pullabhotla, and Baylis 2022).

[9] Shukla, Pullabhotla, and Baylis (2022) demonstrate evidence of anchoring on zero prices and no evidence of anchoring on heavily-subsidized positive prices. They find that WTP in stage 2 is higher for farmers who paid a non-zero price in stage 1 than for farmers assigned to the free provision arm. In contrast, they find no differences in average bids across the range of positive prices.

In this paper, we study anchoring on zero pricing in the field, focusing on a particular agricultural innovation—a novel IPM package that protects farmers from pest and weed damages. We design and implement an experiment that enables us to identify anchoring and learning effects separately. Subsidies may induce anchoring and learning (by making the new technology more accessible), and these effects might pull in opposite directions. In case a new technology performs better than expected, learning will increase future willingness to pay after the subsidy is lifted while anchoring will decrease it. A net zero effect might thus hide two potentially large offsetting effects, impeding our understanding of the long-run welfare effects of short-run subsidies. While we speculate that anchoring may be a rather generic process, occurring in many domains (i.e., human health, cooking, agriculture), the effects of learning must be technology- and context specific. The welfare effects of subsidies therefore vary across contexts.

Our results are strong. We find that zero-pricing invites anchoring for subjects who do not experiment with the IPM package, which reduces follow-up willingness to pay for the new technology. However, zero-pricing combined with an incentive to experiment invites adoption and learning, and learning about the benefits of the IPM package that we offer dominates the anchoring effect. Hence, while anchoring has a negative effect on future demand, combining subsidies with an incentive invites sufficient experimentation in our case study to increase future demand. For the case we study, providing subsidies and add-on incentives to experiment is a policy that would likely increase welfare. An important question for future work is whether the magnitude of the anchoring effect varies across product categories. The current analysis focuses on a labor-intensive technology package that requires complementary investment (and yields a combination of private and public benefits). Additional research based on separately identified learning and anchoring effects across product categories with different sets of characteristics would be welcome.

More in general, it appears as if there is something special about zero-pricing. Zero prices, unlike small positive prices, seem to invite anchoring (Shukla, Pullabhotla, and Baylis 2022), and the price elasticity of demand is high at very low prices (Dupas 2014b). This implies a need to carefully balance the short-term benefits of full subsidies (increased inclusion) versus the longer-term costs (reduced WTP) if subsidies are temporary and eventually to be replaced by market provision. This insight is all the more relevant because we document that the effects of anchoring (and, not surprisingly, learning) appear rather persistent. Traces of anchoring can be picked up even after multiple years, suggesting care should be taken when offering innovations at zero cost if the intention is to ultimately lift these subsidies and transition towards provision based on markets.

3.7. References

Abro, Zewdu, Emily Kimathi, Hugo De Groote, Tadele Tefera, Subramanian Sevgan, Saliou Niassy, and Menale Kassie. 2021. "Socioeconomic and Health Impacts of Fall Armyworm in Ethiopia." *PLoS ONE* 16: e0257736. https://doi.org/10.1371/journal.pone.0257736.

Ashraf, Nava, James Berry and Jesse Shapiro. 2010. "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia." *American Economic Review* 100(5): 2381-2413.

Balew, Solomon, Erwin Bulte, Zewdu Abro and Menale Kassie. 2023. "Incentivizing and Nudging Farmers to Spread Information: Experimental Evidence from Ethiopia." *American Journal of Agricultural Economics* 105(3): 994-1010.

Bensch, Gunther and Jorg Peters. 2019. "One-Off Subsidies and Long-Run Adoption: Experimental Evidence on Improved Cooking Stoves in Senegal." *American Journal of Agricultural Economics* 102(1): 72-90.

Bulte, Erwin, John List and Daan van Soest. 2020. "Towards an Understanding of the Welfare Effects of Nudges: Evidence from a Field Experiment in the Workplace." *Economic Journal* 130(632): 2329-2353.

Burchardi, Konrad, Jonathan de Quidt, Selim Gulesci, Benedetta Lerva, and Stefano Tripodi. 2021. "Testing Willingness to Pay Elicitation Mechanisms in the Field: Evidence from Uganda." *Journal of Development Economics* 152: 102701. https://doi.org/10.1016/j.jdeveco.2021.102701.

Cai, Jing,, Alain de Janvry and Elisabeth Sadoulet. 2020. "Subsidy Policies and Insurance Demand." *American Economic Review* 110(8): 2422-2453.

Carter, Michael, Rachid Laajaj and Dean Yang. 2023. "Subsidies and the African Green Revolution: Direct Effects and Social Network Spillovers of Randomized Input Subsidies in Mozambique." *American Economic Journal: Applied Economics* 13(2): 206-229.

Cohen, Jessica and Pascaline Dupas. 2010. "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment." *Quarterly Journal of Economics* 125(1): 1-45.

Cohen, Jessica, Pascaline Dupas and Simone Schraner. 2015. "Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial." *American Economic Review* 105(2): 609-645.

Dupas, Pascaline. 2014a. "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment." *Econometrica* 8291): 197-228.

Dupas, Pascaline. 2014b. "Getting Essential Health Products to their End Users: Subsidize, but how Much? *Science* 345(6202): 1279-1281.

Food and Agricultural Organization of the United Nations, 2020. *The Global Action for Fall Armyworm Control: Action framework 2020–2022. Working together to tame the global threat*. Rome. http://www.fao.org/documents/card/en/c/ca9252en.

Fischer, Greg, Dean Karlan, Margareth McConnell and Pia Raffler. 2019. "Short-Term Subsidies and Seller Type: A Health Products Experiment in Uganda." *Journal of Development Economics* 137: 110-124. https://doi.org/10.1016/j.jdeveco.2018.07.013.

Foster, Andrew and Mark Rosenzweig. 2010. "Microeconomics of Technology Adoption." *Annual Review of Economics* 2: 395-42. https://doi.org/10.1146/annurev.economics.102308.124433.

Kahn, Zeyaur, Charles Midega, David Amudavi, Ahmed Hassanali and John Pickett. 2008. "On-Farm Evaluation of the Push-Pull Technology for the Control of Stemborers and Striga Weed on Maize in Western Kenya." *Field Crops Research* 106(3): 224-233.

Kassie, Menale, Jesper Stage, Gracious Diiro, Beatrice Muriithi, Geoffrey Muricho, Samuel Ledermann, Jimmy Pitcharr, Charles Midega and Zeyaur Kahn. 2018. "Push-Pull Farming System in Kenya: Implications for Economic and Social Welfare." *Land Use Policy* 77: 186-198.

Kőszegi, Botond and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics* 121(4): 1133-11656.

Meriggi, Niccolo, Erwin Bulte and Ahmed Mushfiq Mobarak. 2021. "Subsidies for Technology Adoption: Experimental Evidence from Rural Cameroon." *Journal of Development Economics* 153: 102710. https://doi.org/10.1016/j.jdeveco.2021.102710.

Omotilewa, Oluwatoba, Jacob Ricker-Gilbert, and John Ainembabazi. 2019. "Subsidies for Agricultural Technology Adoption: Evidence from a Randomized Experiment with Improved Grain Storage Bags in Uganda." *American Journal of Agricultural Economics* 101(3): 753-772.

Shampanier, Kristina, Nina Mazar and Dan Ariely. 2007. "Zero as a Special Price: The True Value of Free Products." *Marketing Science* 26(6): 742-757.

Shukla, Pallavi, Hemant Pullabhotla, and Kathy Baylis. 2022. "Trouble with Zero: The Limits of Subsidizing Technology Adoption." *Journal of Development Economics* 158: 102920. https://doi.org/10.1016/j.jdeveco.2022.102920.

Appendices

Appendix A: Incentivizing leaders and measuring their performance

Randomly selected group leaders were incentivized to experiment with the IPM package on their form by providing material and non-material rewards in case of “good performance”. Leader performance was measured by assessing the knowledge level of the follower farmers in their one-to-five group. To receive the reward, at least 50% of these follower farmers should demonstrate “basic knowledge” about push-pull technology. This was tested by questions about (i) the benefits of adopting PPT, and (ii) the implementation and functionality of PPT (this is why experimenting with the technology on the leader’s plot was important). Advantages of PPT include: (i) control of stem-borer and fall armyworm, (ii) reduction of soil erosion, (iii) improvement of soil fertility, and (iv) companion crops are an important source of animal feed. Regarding implementation and functioning, PPT involves (i) intercropping a leguminous fodder crop called *desmodium* with maize, (ii) *desmodium* pushes stem-borer away from the maize, (iii) *brachiaria* is sown surrounding the maize plot; and (iv) *brachiaria* attracts the stem-borer, “pulling” it from the maize field. Followers who were able to mention at least two advantages and two items regarding implementation and functionality were defined to have “basic knowledge” about PPT.

Balew et al. (2023) offered 4 different types of rewards to incentivize group leaders to experiment with the IPM package in their plots: (i) PIGF, a private incentive (sickle) as a bonus in case of good performance, (ii) PILF, the same private incentive as an upfront payment that had to be returned in case of insufficient performance, (iii) SIGF, a social prestige reward (certificate of good performance) as a bonus in case of good performance, and (iv) SILF, the same social incentive as an upfront reward that had to be returned in case of insufficient performance. Notation-wise, PI and SI represent private incentive and social incentive, respectively and GF and LF represent gain frame and loss frame, respectively. Balew et al. (2023) demonstrate that these 4 incentives were able to induce extra experimentation (by the leaders) which increased knowledge levels for their followers:

Table A1: Incentives, leader experimentation and follower knowledge

	Control	*Private incentive framed as gain(PIGF)*	*Private incentive framed as loss(PILF)*	*Social prestige framed as gain(SIGF)*	*Social prestige framed as gain(SILF)*
Leaders experimented with PPT on their maize plot(1=yes)	0.29	0.51	0.34	0.42	0.44
Group leader meets threshold (50% follower farmers "knowledgeable")	0.64	0.81	0.75	0.71	0.89

For more information about the incentives and their effects, refer to Balew et al. (2023).

Appendix B: Additional Tables

Table B1: Baseline balance test for leaders

	(1)	(2)	(3)	(1)-(2)	(1)-(3)	(2)-(3)
	Control	Subsidy-only	Subsidy + Incentive	Pairwise t-test	Pairwise t-test	Pairwise t-test
Variable	Mean/(SE)	Mean/(SE)	Mean/(SE)	Mean difference	Mean difference	Mean difference
Sex of respondent (1=male)	0.923	0.955	0.970	-0.032	-0.047**	-0.015
	(0.020)	(0.015)	(0.009)			
Marital status of head (1=married)	0.973	0.975	0.987	-0.002	-0.014	-0.012
	(0.012)	(0.011)	(0.006)			
Family size (number)	5.421	5.630	5.685	-0.209	-0.264	-0.055
	(0.126)	(0.130)	(0.096)			
Age of head (years)	43.694	45.275	44.288	-1.581	-0.594	0.987
	(0.665)	(0.719)	(0.512)			
Education of head (1=literate)	0.525	0.620	0.569	-0.095*	-0.044	0.051
	(0.037)	(0.034)	(0.026)			
Primary activity of head (1=farming)	0.984	0.990	0.995	-0.006	-0.011	-0.005
	(0.009)	(0.007)	(0.004)			
Lived in the village (Months)	11.913	11.890	11.976	0.023	-0.063	-0.086*
	(0.054)	(0.057)	(0.011)			
Fall armyworm Knowledge (1=yes)	0.650	0.695	0.712	-0.045	-0.061	-0.017
	(0.035)	(0.033)	(0.024)			
FAW major constraint (1=yes)	0.131	0.070	0.100	0.061**	0.031	-0.030
	(0.025)	(0.018)	(0.016)			
Stemborer Knowledge (1=yes)	0.907	0.900	0.900	0.007	0.007	-0.000
	(0.022)	(0.021)	(0.016)			
Stemborer major constraint (1=yes)	0.120	0.070	0.089	0.050*	0.031	-0.019
	(0.024)	(0.018)	(0.015)			
Striga Knowledge (1=yes)	0.683	0.600	0.679	0.083*	0.004	-0.079*
	(0.034)	(0.035)	(0.024)			
Striga major constraint (1=yes)	0.393	0.300	0.340	0.093*	0.054	-0.040
	(0.036)	(0.032)	(0.025)			
Maize farm area (hectare)	0.428	0.451	0.427	-0.023	0.001	0.024
	(0.023)	(0.021)	(0.015)			
Total farm area (hectare)	1.203	1.302	1.288	-0.099	-0.085	0.014
	(0.049)	(0.053)	(0.041)			

Extension visits (number)	2.350	1.965	2.248	0.385	0.102	-0.283
	(0.191)	(0.156)	(0.138)			
Have milking cow (1=yes)	0.470	0.485	0.482	-0.015	-0.013	0.003
	(0.037)	(0.035)	(0.026)			
Have credit constraint (1=yes)	0.350	0.360	0.321	-0.010	0.029	0.039
	(0.050)	(0.034)	(0.024)			
N	183	201	370	384	553	571
Notes: ***, **, and * indicate significance at the 1, 5, and 10 percent level						

Table B2: Attrition analysis of group leaders (OLS)

Dependent variables	**Attrition status of group leaders**
Subsidy (1= yes)	-0.007 (0.015)
Incentive (1= yes)	-0.019 (0.014)
Sex of respondent (1=male)	-0.021 (0.046)
Marital status of household head (1=married)	-0.051 (0.071)
Household members (number)	0.004 (0.003)
Household head age (years)	-0.001 (0.000)
Household Primary activity farming (1=yes)	-0.096 (0.095)
Household head education (years)	-0.002 (0.009)
Lived in the village (Months)	0.004 (0.006)
Maize plot (hectares)	-0.019 (0.022)
Total land owned by household (hectares)	0.005 (0.014)
Knowledge of stemborer (1= yes)	-0.007 (0.031)
Stemborer major problem (1= yes)	0.0255)
Knowledge of fall armyworm (1= yes)	0.002 (0.014)
Fall armyworm major problem (1= yes)	0.022 (0.023)
Knowledge of striga (1= yes)	0.001 (0.019)
Striga major problem (1= yes)	-0.026 (0.016)
Household owns milking cow (1 =yes)	0.020 (0.013)
Household faces credit constraint (1=yes)	0.016 (0.011)
Village fixed effects	Yes
Mean dependent variable	0.024
Pseudo r-squared	0.099
Akaike crit. (AIC)	-661.083
SD dependent var	0.153
Number of observations	754
Bayesian crit. (BIC)	-402.061
F test for joint significance (p-value)	0.550

Notes: Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

Table B3: Non-compliance, farmers refusing to state a bid (OLS)

Dependent variables	**No bid at midline (group leaders)**	**No bid at end-line (group leaders)**
Subsidy (1=yes)	0.021 (0.043)	0.053 (0.042)
Incentive (1=yes)	-0.037 (0.039)	-0.024 (0.031)
Sex of respondent (1=male)	0.012 (0.072)	0.030 (0.080)
Marital status of household head (1=married)	0.046 (0.076)	-0.014 (0.078)
Household members (number)	-0.015* (0.008)	-0.008 (0.007)
Household head age (years)	0.004** (0.002)	0.001 (0.001)
Household Primary activity farming (1=yes)	0.179*** (0.060)	0.198** (0.079)
Household head education (years)	-0.038 (0.032)	-0.058** (0.025)
Lived in the village (Months)	0.000 (0.022)	-0.045 (0.027)
Maize plot (hectare)	-0.209*** (0.064)	-0.211*** (0.061)
Total land owned by household (hectare)	-0.047** (0.021)	-0.068** (0.018)
Knowledge of stemborer (1= yes)	-0.006 (0.047)	0.37 (0.051)
Stemborer major problem (1= yes)	-0.055 (0.050)	0.006 (0.048)
Knowledge of fall armyworm (1= yes)	0.003 (0.036)	0.009 (0.030)
Fall armyworm major problem (1= yes)	-0.017 (0.042)	-0.057* (0.033)
Knowledge of striga (1= yes)	-0.093*** (0.033)	-0.056* (0.033)
Striga major problem (1= yes)	-0.070 (0.035)	-0.024 (0.029)
Household owns milking cow (1 =yes)	-0.048 (0.031)	-0.054** (0.024)
Household faces credit constraint (1=yes)	0.028 (0.024)	0.027 (0.024)
Village effect	Yes	Yes
Mean dependent variable	0.182	0.145
Pseudo r-squared	0.154	0.181
Akaike crit. (AIC)	676.100	518.567
SD dependent var	0.386	0.353
Number of observations	736	736
Bayesian crit. (BIC)	933.769	776.236
F test for joint significance (p-value)	0.000	0.000

Notes: Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

Table B4: Average bids of leaders for PPT package

Treatment	Midline	Endline
Arms	Average Leader bid for PPT per hectare (ETB)	Average Leader bid for PPT per hectare (ETB)
Pure control	2793	2558
Subsidy-only	2449	2780
Subsidy plus private gain (PIGF)	3135	3910
Subsidy plus private loss (PILF)	2587	3008
Subsidy plus social gain (SIGF)	2919	3363
Subsidy plus social loss (SILF)	2864	3540
Total	2740	3056

Notes: This tabulation is based on the full sample of farmers, and we imputed a bid value of zero for those leaders who opted out of the BDM. Exchange rate: 55.92 ETB = USD 1.

Table B5: Quantile regression of endline results (ITT estimates)

	Leader WTP at endline			
	35th quintile regression	**65th quintile regression**	**35th quintile regression**	**65th quintile regression**
Subsidy-only	-1540*** (581)	800 (609)	-1615*** (400)	378 (651)
Subsidy-plus-Incentive	90 (430)	1361*** (475)	243 (406)	1329*** (410)
Village F.E.	Yes	Yes	Yes	Yes
Covariates	No	No	Yes	Yes
Constant	1360*** (981)	3200*** (240)	-287 (3128)	6969 (5508)
Mean dependent var	3055	3055	3055	3055
Pseudo R2	0.15	0.11	0.19	0.16
N	736	736	736	736
*Notes; without c*ovariates			*Notes;* Covariates used are the same as in other models	

Table B6: Heterogeneity analysis

	Leaders at midline				Leaders at end line			
Panel A	*Second stage (WTP for PPT package as dependent variable)*							
	Maize plot		*Family size*		*Maize plot*		*Family size*	
	Large	Small	Large	Small	Large	Small	Large	Small
Experiment	9325* (5015)	-138 (2431)	8547** (3985)	497 (2257)	8766 ** (4414)	852 (3057)	7575** (3167)	-1569 (4614)
Subsidy	-4811* (2746)	-500 (854)	-3455** (1387)	342 (1150)	-3347 (2480)	-603 (1094)	-2903* (1588)	532 (1577)
Village F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-1358 (3703)	6657* (1572)	3623 (4127)	477 (2972)	743 (3774)	7723 (4794)	-1153 (5152)	7852** (2963)
Panel B	*First stage (Experiment with PPT package as dependent variable)*							
Subsidy	047*** (0.063)	0.18** (0.074)	0.24*** (0.066)	0.39*** (0.077)	047*** (0.063)	0.18** (0.074)	0.18** (0.074)	0.24*** (0.066)
Incentive	0.11*** (0.064)	0.20** (0.068)	0.12** (0.058)	0.16** (0.076)	0.11*** (0.064)	0.20** (0.068)	0.20** (0.068)	0.12** (0.058)
Village F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.06* (0.030)	-0.06 (0.51)	-0.01 (0.42)	0.81 (0.55)	0.06* (0.030)	-0.06 (0.51)	-0.06 (0.51)	-0.01 (0.42)
R2	0.39	0.38	0.29	0.41	0.39	0.38	0.38	0.29
Partial F	83.48	3646.00	115.41	109.69	83.48	3646.00	3646.00	115.41
N	368	368	368	368	368	368	368	368

Notes: The table reports results of a 2sls model where experimentation with the PPT package is instrumented with a random incentive (reward) to experiment. Columns (1-4) capture willingness to pay for a PPT package during midline, and columns (5-8) capture willingness to pay at endline. In columns 1, 2, 5 and 6 we use an aggregate incentive variable. In columns 3, 4, 7 and 8 we use disaggregated incentive variables (material and non-material rewards, paid as a bonus or up-front—see Balew et al. 2023). Columns (1-3) and (5-7) are based on the full sample of group leaders, where we included a zero value for farmers who refused to state a bid during the BDM. Columns (4) and (8) are based on a subsample where we dropped farmers who refused to state a bid. The covariates included are Sex of respondent, Marital status, , *Age*, *Education*, *Primary activity is farming*, *Lived in the village*, *FAW knowledge*, *FAW major constraint*, *SB knowledge*, *SB major constraint*, *Striga knowledge*, *Striga major constraint*, , *Milking cow*, *Credit constraint*. We report standard errors clustered at the sub-kebele level in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1%, respectively.

Chapter 4

The Impact of Push-Pull Technology on Livestock Productivity and Household Income in Mixed Farming Systems: Experimental Evidence from Northwest Ethiopia

Abstract : This study evaluates the impact of Push-Pull Technology (PPT), an integrated pest and forage management (IPM) system on livestock productivity and household income in mixed crop-farming systems in Northern Ethiopia. Using a cluster randomized controlled trial involving 754 households, we apply local average treatment effect (LATE) estimation to account for non-compliance. The results indicate that PPT adoption significantly enhances milk production per cow, as well as household income from feed sales and livestock sales. Adopting households gained about 2,356 ETB (USD 43) at midline and by 3,323 ETB (USD 60) at endline. These findings demonstrate the potential of PPT to address feed constraints and enhance livestock-based livelihoods in resource-constrained smallholder systems.

Keywords: Mixed farming systems, livestock feed, Randomized controlled trial, Push pull technology, livestock productivity, income

JEL Codes: C93, O33, Q12, Q18

Publication status: Balew S., (xxxx) " The Impact of Push-Pull Technology on Livestock Productivity and Household Income in Mixed Farming Systems: Experimental Evidence from Northwest Ethiopia" is submitted to a journal.

4.1. Introduction

Mixed crop-livestock farming systems play a crucial role in providing food and income for households in developing countries like Ethiopia (Guja et al., 2019). This farming approach is the predominant method of livestock production in Ethiopia, where crops and livestock are integrated to maximize resource use efficiency and productivity. Livestock contribute to household livelihoods by providing draught power, threshing, manure for soil fertility, and a source of income, while depending on crop residues as a key feed source (Duressa et al., 2014; Amejo et al., 2018). The integration of livestock into crop farming is an essential for smallholder farmers, supporting household food security, nutrition, and economic stability (Tegegne et al., 2013; Karta and Dey, 2022; Dadi et al., 2023).

Ethiopia is experiencing rapid demographic shifts, including population growth, urbanization, and increasing household incomes (FAO, 2020; Karta and Dey, 2022). These have driven an increasing demand for livestock products, such as milk and meat. This trend presents opportunities for farmers to enhance their income and improve food security (Karta and Dey, 2022). In the 2021/2022, Ethiopia had an estimated 60 million cattle, with smallholder farmers owning 96.8% of them (CSA, 2022). However, livestock productivity, especially under smallholder management, remains below potential due to various challenges (Bereda et al., 2014; Ayalew et al., 2018; Gizachew et al., 2020; Gobezie et al., 2020; Zegeye, 2023; Fetene et al., 2025).

One of the major barriers to improved livestock productivity is the shortage of high-quality and affordable feed (Mayberry et al., 2017; 2018; Abebe et al., 2023; Fetene et al., 2025). In Ethiopia, feed availability is a more critical constraint than genetic potential of livestock (Belay and Negesse,2018; Tigistu, 2021). Historically, livestock have depended on natural pastures and crop residues, but these sources of animal feed are becoming increasingly scarce due to land competition, shrinking grazing areas, low productivity, poor management, and erratic rainfall patterns (Karta & Dey, 2022). Feed shortages negatively affect farmers by reducing both livestock and crop production (Hadush, 2019, 2020; Fetene et al., 2025). A recent study on feed shortage in Ethiopia found that it restricted investment in modern inputs and heightened livestock vulnerability to diseases, causing a 14% increase in livestock mortality losses, a 77% rise in production expenses, and a 4% decline in crop production value (Fetene et al., 2025). To address these challenges, the Ethiopian government has introduced a Ten-Year Development Plan (TYDP) that prioritizes improved forage production, viewed as a strategy to boost livestock productivity (Karta & Dey, 2022).

Despite these efforts, the availability of high-quality forage seeds remains a challenge. Previous development projects, policy interventions, and technical support have attempted to address this issue, but gaps persist in forage production and utilization (Shapiro et al., 2017). Integrated pest and feed management strategies have emerged as promising solutions to address these interlinked challenges, with Push-Pull Technology (PPT) gaining significant attention(Isgren, et al., 2023).

PPT is an ecological pest, weed, and feed management strategy that enhances agricultural productivity by intercropping cereal crops with specific fodder plants that repel and attract pests while simultaneously improving soil health and providing high-quality livestock forage (Khan et al., 2014; Pickett et al.,2014 ; Kassie et al., 2018). PPT involves planting brachaira grass around cereal crops field borders to attract the pests (the "pull" component) and intercropping cereal crops with Desmodium, forage legumes that repels pest (the "push" component). In addition to pest control, PPT enhances livestock productivity by increasing the availability of nutritious forage, supporting milk production, livestock fattening, and diversifying farm income through the sale of surplus forage[1] (Khan et al.,2011,2014; Pickett et al.,2014 ; Kassie et al., 2018; Cheruiyot et al.,2020).

The effectiveness of PPT in improving livestock feed and productivity has been demonstrated in countries such as Kenya and Uganda((Khan et al., 2011, 2014). However, the technology is still relatively new in Ethiopia, and its impact on livestock productivity and household income remains underexplored. This study evaluates the effects of PPT adoption on livestock productivity and feed and livestock sales income among smallholder farmers in Northwest Ethiopia, using data from a cluster randomized controlled trial (RCT) covering 114 sub-villages and 754 households. Given that nearly 90% of Ethiopian agrarian households practice mixed crop-

[1] The advantages of PPT include: (i) effective control of stemborers and fall armyworms, (ii) control striga weed, (iii) reduced soil erosion, (iv) enhanced soil fertility, and (v) an important source of animal feed from companion crops.

livestock farming (Wakeyo and Elias, 2023; Assefa et al., 2024), understanding the impact of PPT in this context provides valuable insights for broader agricultural development efforts.

This study advances the literature on the economic impact of PPT by employing experimental data to move beyond quais-experimental analyses. Unlike prior studies, this research establishes a causal link between PPT interventions and farm households' economic returns. A further contribution is the differentiation between short-term and long-term economic effects, elucidating how the timing of PPT adoption influences farmers returns, a critical gap in the existing literature.

This paper is structured as follows: Section 2 presents the background. Section 3 outlines the experimental design and data. Section 4 describes the identification and estimation strategies. In Section 5 we present our results and discuss their implications. Section 6 concludes the study.

4.2.Background

This study assesses the impact of Push-Pull Technology (PPT) on livestock productivity and household income among the leaders of one-to-five groups in the Jabi Tehnan district of northwestern Ethiopia. These one-to-five groups are grassroots extension institutions in Ethiopia, where rural farmers organized into groups of six. One member is selected as the leader and receives training directly from government extension workers. The leader then shares this knowledge and trains the remaining five group members (Balew et al., 2023, 2025). This study focuses specifically on these leaders as critical entry points for technology dissemination and behavioral change.

PPT adoption is emerged as an innovative solution to address feed shortages in mixed crop-livestock farming systems. However, its wider adoption faces several challenges. These include labor demands for land preparation, planting, weeding, and managing pest control and forage crops (desmoidum and brachiaria) to avoid competition with staple crops. In addition , companion crops need extra space to be planted alongside food crops., and a reliable seed system the companion crops is lacking (Khan et al., 2014; Kassie et al., 2018; Balew et al., 2023, 2025 ;

Yadav et al.,2025)2. To facilitate adoption, the study seeks to provide free seeds through extension workers and deliver targeted training programs to equip leader farmers with the necessary skills[3].

In the study region, many farmers remain skeptical about the advantages of push-pull technology (PPT), resulting in a slow adoption process. This is primarily driven by social learning, in which lead farmers trained and share knowledge within their group. Because follower farmers are hesitant to adopt the technology, the overall spread of PPT is further hindered (Balew et al., 2023). This underscores the importance of evaluating and documenting the actual impacts of PPT adoption on livestock productivity and household income to inform evidence-based strategies for promoting its adoption in Ethiopia.

4.3. Sampling, experimental design, and data

4. 3.1. Sampling and experimental design

To measure the impacts of PPT, we implemented a cluster randomized controlled trail across 38 kebeles (villages) and 114 sub-kebeles. A census of groups within these sub-kebeles was conducted, and approximately five groups per sub-kebele were randomly selected for inclusion, resulting in a final sample of 754 group leaders (hereafter will use them interchangeably as farmers). After conducting a census of farmer groups within each sub-kebele, approximately five groups were randomly selected per sub-kebele, resulting in a final sample of 754 group leaders. Institutional Review Board (IRB) approval was obtained,[4] and informed consent was secured from participants and local officials.

The study applied a two-stage randomization design. 28 sub-kebeles (183 leaders) were assigned to the "pure control" group and 86 sub-kebeles (571 leaders) were assigned to the treatment group. Leaders in the treatment group received PPT seeds. Each starter package contained 500 grams each of Brachiaria and Desmodium seeds, sufficient for intercropping on 0.2 hectares. Leaders in

[2] Planting companion (desmoidum and brachiaria) crops and separating desmodium from weeds during establishment ischallenging.

[3] Balew et al., 2025, found that subsidies do not affect the long-term adoption of PPT.
[4] The experiment was pre-registered at https://www.socialscienceregistry.org/trials/5642AER.

the control group had the option to purchase a similar package for 560 ETB (USD 10) . The average size of maize plots in the sample was 0.43 hectares, with 14% allocated to PPT.

Within the treatment group, farmers were further divided into sub-arms to test different incentives aimed at promoting PPT adoption. In sub-arm 1, participants received only the starter package seeds at no cost (subsidy only group), and sub-arm 2 received additional incentives to encourage adoption (subsidy-plus-incentive group). The incentive includes either a material reward (a sickle) or a non-material reward (framed certificate with almost equivalent value of the sickle). For a detailed description of the incentives and experimental design, refer to Balew et al. (2023; 2025).

Figure 1: Summary of experimental design

<table>
<tr><td colspan="3">Sample frame:114 sub-kebeles (SK), 754 lead farmers (N)</td></tr>
<tr><td>↓
Pure control
SK=28, N=183</td><td colspan="2">↓
Treatment
SK=86, N=571</td></tr>
<tr><td></td><td>↓
Subsidy only group
SK=28, N=201</td><td>↓
Subsidy plus incentive group
SK=58, N=370</td></tr>
</table>

Before the bassline survey, all selected leaders participated in a two-day training session held in their respective villages, facilitated by experts from the International Centre of Insect Physiology and Ecology (*icipe*) and the Jabi Tehnan district agricultural office. The training focused on PPT implementation, including optimal plant spacing, timely weeding, and trimming and harvesting of companion crops, and effective use of the resulting forage for livestock feeding.

4.3.2. Data collection and descriptive statistics

This study collected three waves of data using tablet-equipped enumerators. The baseline survey, conducted between June and August 2018 before training sessions, covering household demographics (e.g., age, family size, literacy), food security status, and details about participants' crops and plots, livestock ownership and sales, forage production and sales and milk production per cow. Midline survey (November 2020) conducted after the subsequent growing season to assess early impacts, and endline (February 2022) collected final data on forage sales, herd sizes, forage and livestock selling price.

The sample data reveals the average household head is 44 years old, with an average family size of about six individuals (Table 1). The literacy rates show that around 60% of household heads are literate and 98% of group leaders are Married. Credit constraints affect about 37% of households, limiting their ability to acquire agricultural inputs. The total agricultural landholding averages 1.27 hectares per household. Sampled households cultivate an average of 0.43 hectares of maize plots where PPT is practiced, with an average yield of approximately 4,000 kg per hectare. The average annual household income was 22,081 ETB(approximately USD 401).

Livestock ownership is widespread. Approximately 98.7% of households own at least one animal, with an average of 8.4 livestock per household. Specifically, 47% own at least one milking cow, 70% had at least one heifer, 49% had at least one bull, and 91% had at least one ox. Regarding small livestock, about 67% owned at least one goat, and 69% owned at least one sheep. 25.6% of households produced and sold feed and the average annual income from feed sales was 890 ETB (USD 16).

Table 1: baseline data and descriptive statistics

Variables	Mean	Standard deviation
Covariates;-		
Marital Status of head(1=yes)	0.98	0.14
Family size(number)	5.61	1.81
Age of head (years)	44.41	9.74
Education of head (1=literate)	0.57	0.50
Primary activity of head(1=farming)	0.99	0.10
Have credit constraint(1=yeas)	0.37	0.48
Maize farm area(hectare)	0.43	0.29
Total farm area(hectare)	1.27	0.75
Outcome variables		
Cows(number)	0.82	1.04
Calves(number)	0.82	1.04
Heifers (number)	1.23	1.10
Bulls(number)	0.77	1.02
Oxen (number)	1.67	0.70
Cattle(number)	5.32	3.44
Goats (number)	1.52	1.78
Sheep (number)	1.56	1.34
Goats and sheep(number)	3.08	1.87
Milk production per cow per day(Litters)	0.93	1.11
Lactation length per cow(Months)	5.42	5.78
Feed production and sales income (ETB)	891	1604
Livestock production and sales(ETB)	4768	3019
Number of observations	754	

The follow-up data included key metrics, such as annual income generated from producing and sales of feed as well as other significant measurements used in this study (see (details in Appendix Table A4). Our midline data shows that the average income from feed sales per household was 1,149 ETB (USD 21) for PPT users, this amount rose to around 1,816 ETB (USD 33), while non-adopter households averaged 879ETB (USD 16). According to our endline data, the average income from feed sales per household increased to 1,260 ETB (USD 23) for all participants. For adopters, this figure increased to about 2,436 ETB (USD 44), whereas non-adopters reported an income of 786 ETB (USD 14)

Feed sales are strongly correlated with the total cultivated land area of adopting households. Figures 2 and 3 demonstrate that adopting households selling feed generally have larger landholdings than those that do not. This indicates that crop residues, combined with PPT forage, enable households to produce surplus feed for sale. Figure A in the appendix corroborates this finding by illustrating a negative relationship between feed sales and the ratio of livestock (TLU) to cultivated land (hectares).

Figure 2: Feed sales and land size by adopters

Figure 2A: Midline

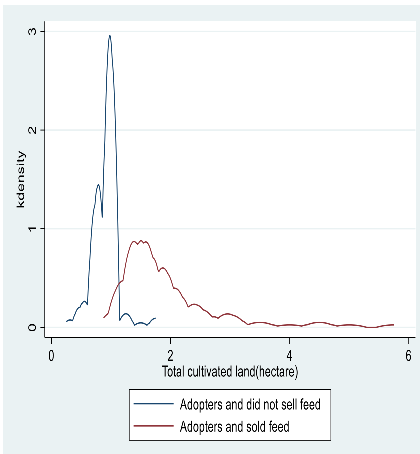


Figure 2B: Endline

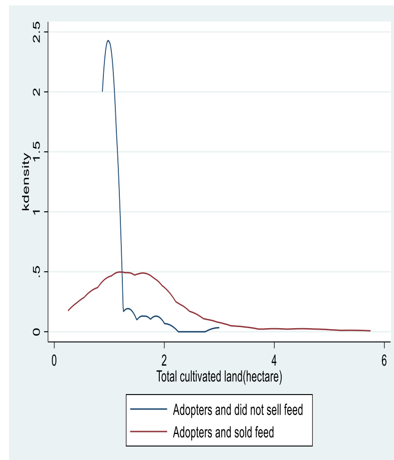
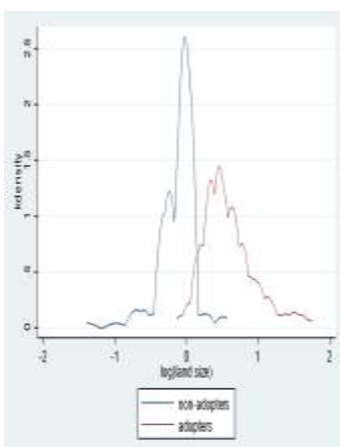
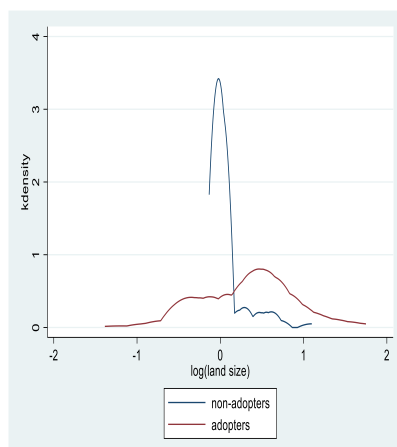


Figure 3: Feed sales and log(land size) by adopters

Figure 3A: Midline



`Figure 3B: Endline



The study assesses key outcomes, including milk production per cow, herd sizes, and household income from feed and livestock sales. The findings indicate that PPT adoption has significantly increased milk yield, livestock herd, and income. Average number of livestock per household increased by 3.2 at the midline and by 6.5 at the end line. Milk production per cow increased by

29 % at the midline and by 33% at endline. Furthermore, household income from feed[5] production and sales increase by 1,212 ETB (USD 22) at midline and by 1,840ETB (USD 33) at end line[6]. Finally, household income from livestock production and sales increase by 1,144 ETB (USD 21) at midline and by 1,483ETB (USD 27) at end line[7].

4.3.3. Baseline data balance and randomization check

Appendix Tables A1, A2, and A3 provide a comprehensive summary of the data, including t-tests designed to evaluate the balance between control and treatment groups. The results indicate that the randomization was largely effective, ensuring the difference between groups can be attributed to the treatment rather than external factors.

However, significant imbalances are evident in specific covariates and outcome variables when comparing households based on PPT adoption behavior, suggesting the presence of underlying factors influencing outcomes. These discrepancies highlight the need for a robust identification strategy to account for potential selection biases

4.3.4. Attrition analysis

The midline survey re-interviewed 736 group leaders, with a loss of 18 participants since the baseline survey. The endline survey had no further attrition, with data collected from all remaining 736 leaders. An attrition analysis (Appendix Table A5) examined attrition status against baseline controls and treatment arm dummies. The results did not show significant correlations between attrition and most variables, including treatment status. The F-test's p-value confirmed that we cannot reject the null hypothesis of no correlation, indicating that attrition bias is unlikely to affect the study's validity.

4.4. Estimation strategy

This analysis primarily employs a two-stage least squares (2SLS) approach. Experimental groups were formed through randomization, as detailed in Appendix Table A1, ensuring participant comparability. However, compliance issues regarding PPT adoption introduced some challenges.

---

[5] Feed includes crop residues, hay, and forage from PPT companion crops (desmoidum and brachiaria).
[6] These estimates are made after accounting for inflation between the years.
[7] These estimates are made after accounting for inflation between the years.

Within the treatment-assigned group, adoption was driven by personal initiative, resulting in systematic differences between adopters and non-adopters, as shown in Appendix Table A3. These differences could potentially bias estimates of the technology's impact.

To address this, we used the incentive assignment to experiment with PPT as an instrumental variable for assessing adoption effects. By applying the Local Average Treatment Effect (LATE) approach, we can isolate the impact of adoption specifically for farmers who adopted it as a result of the given incentive. This method provides a more precise estimate of adoption effects while controlling for unobservable factors that could skew our results.

In the first stage of our 2SLS analysis, we use the incentive to experiment with the PPT package (denoted as "Incentive") and assignment to the full subsidy (designated as "Subsidy") as excluded instruments. The variable "Experiment," our endogenous variable, is defined as the adoption of the PPT package on a portion of the farmer's plot, validated through field observations.

It is important to note that we do not regress experimentation on distinct treatment groups, as shown in Figure 1. Instead, the dummy variable "Subsidy" is coded as one for all treated farmers who received a free starter package, encompassing both the subsidy-only and subsidy-plus-incentive groups. The dummy variable "Incentive" is marked as one for the subset of treated households that received an incentive to trial the PPT package on their farms

In the second stage, we utilize the predicted values of "Experiment" to assess variations in several outcome variables, including total herd size, average daily milk production per cow (in liters), average lactation length per cow (in months), average income from feed and livestock sales (expressed in Ethiopian Birr, ETB), as recorded during midline and endline assessments. This relationship is represented by the following model:

$$Experiment_{jsk} = \alpha_k + \beta Incentive_{sk} + \gamma Subsidy_{sk} + \delta \boldsymbol{X}_{jsk} + e_{jsk}, \tag{1a}$$

$$Outcomes_{jsk} = \theta_k + \mu Experiment^{*}_{jsk} + \pi \boldsymbol{X}_{jsk} + \varepsilon_{jsk}. \tag{1b}$$

In equations (1a) and (1b) , farmers belonging to the pure control group serve as the omitted category. The term $\boldsymbol{X}_{jsk}$ represents a vector of control variables, while e_{jsk} and ε_{jsk} denote the error terms. The subscripts j, s, and k signify leader j in sub-kebele s in kebele k. The parameters α_k and θ_k are kebele fixed effects, capturing variation in agroecological, market, and governance condition. For the first stage, we expect that both the incentive and subsidy will promote experimentation by leaders ($\beta > 0$ and $\gamma > 0$, respectively). Substantively, we are interested in the coefficient μ estimated during the second stage. We anticipate that $\mu > 0$, suggesting that the outcomes improve with the adoption of PPT, as we assume that the PPT package for forage enhances livestock feed performance compared to that of the control group farmers. In equation 1b, the "Experiment*" variable represents predicted experimentation, where prediction for farmer j is based on the coefficients of equation (1a), including the experimental group to which farmer j was assigned.

Our identification strategy is predicated on the assumption that the Incentive and Subsidy variables serve as valid instruments for the Experiment variable. To satisfy the exclusion restriction, it is essential that the incentive and subsidy provided to farmers affects outcomes solely by increasing their willingness to experiment with PPT package. This means that any observed changes in livestock productivity or household income must be directly linked to the farmers' adoption of PPT, without interference from other channels.

If the incentive also influences farmers through alternative pathways, such as enhancing their overall motivation, improving market access, or encouraging broader agricultural improvements, it complicates the ability to attribute positive outcomes specifically to the adoption of PPT. Such confounding factors could suggest that the incentive was effective even if the true benefits stemmed from these other influences.

Of the 558 leaders who were offered a subsidized package, 212 adopted Push-Pull Technology (PPT) between the training and midline surveys, resulting in an overall adoption rate of 38%. Notably, adoption rates were particularly low among unincentivized leaders in the subsidy-only group, where just 29% (56 out of 193) opted to use the technology. In contrast, among the 365 incentivized leaders, 156 adopted PPT, yielding an adoption rate of 42%. These differences in adoption rates are statistically significant ($p < 0.01$), suggesting that random assignment to the incentive is a valid instrumental variable for the experiment. Additionally, the adoption rate among leaders in the pure control group was recorded at 0%.

Alongside the LATE effects, we also examined the intention-to-treat (ITT) effects using ordinary least squares (OLS) regressions, with the results provided in the appendix tables for reference. The ITT analysis consisted of regressing our outcome variable against two treatment groups: one receiving only subsidies and the other receiving both subsidies and incentives. To improve the precision of our estimates, we included baseline controls in our regression model. This allowed us to address any pre-existing differences among participants and ensure that our findings truly represent the impact of our interventions.

To analyze the impact of PPT adoption on milk production and lactation length per cow, we employed two regression approaches using both a full sample and a sub-sample of households. The full sample analysis encompassed all households, including those without milking cows in any given year. Milk production for these households was set to zero. To isolate the effect of PPT adoption from the effect of cow ownership, a cow presence dummy variable (1 = present, 0 = absent) was included in the regression models. A sub-sample analysis focused on only households with milking cows during the relevant period.

4.5. Results

4.5.1 PPT adoption and milk production

Table 2 presents the Local Average Treatment Effect (LATE) estimates for the full sample, while Appendix Table B1 shows the corresponding results for the sub-sample. Panel A of Table 2 reports outcomes related to daily milk production per cow and lactation length.

Although the full-sample analysis captures overall trends and offers broader representativeness, it may mask the specific effects of adoption on milk productivity. The sub-sample in Appendix Table B1 focuses on households with milking cows, enabling a more precise assessment of adoption's impact within this group.

Adopters experienced notable gains in milk production. Daily milk yield per cow increased by 0.6 liters (29%) at midline and 0.7 liters (33%) at end-line relative to mean yield among control households with milking cows (Table 2). Similarly, lactation length grew by 1.6 months (14%) at midline and 1.5 months (12.9%) at end-line. These improvements translated into total milk gains per cow of 29 liters at midline and 31 liters at end-line (Table 2).

Sub-sample results (Appendix Table B1) indicate even stronger impacts. Daily milk production per cow increased by 0.9 liters (43%) at midline and 0.8 liters (38%) at end-line. Lactation length also improved by 2 months (20 %) at midline and 2 months (15.8%) at end-line for PPT adopters relative to the mean of control group with milking cows, resulting in additional milk production of 54 liters at midline and 48 liters at end-line. While these increases imply higher potential income, milk is not sold typically directly in the study area. Instead, it is processed into butter for sale; hence the study does not assess impacts on milk sales.

A consistent pattern in both analyses shows similar milk production per cow from midline to endline among households that adopted PPT. Appendix Table C1 displays the Intent-to-Treat (ITT) results, which align generally with the LATE findings.

Table 2. Push-pull technology adoption impacts on milk production (Full sample estimations) (2SLS)

	Midline		End line	
Panel A	*Second stage*			
	Average milk production per cow in a day (liters)	Average lactation length per cow (months)	Average milk production per cow in a day (liters)	Average lactation length per cow (months)
Experiment	0.603*** (0.079)	1.579*** (0.248)	0.672*** (0.084)	1.452** (0.315)
Village Fixed effects.	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
Constant	-0.299* (0.148)	-0.437*** (0.135)	-0.474** (0.203)	-0.689*** (0.337)
Mean of dependent variable	1.209	6.539	1.280	6.997
Panel B	*First stage (Experiment with PPT package as dependent variable)*			
Subsidy	0.286*** (0.035)	0.286*** (0.035)	0.268*** (0.034)	0.268*** (0.034)
Incentive	0.141*** (0.047)	0.141*** (0.047)	0.149*** (0.047)	0.149*** (0.047)
Village F.E.	Yes	Yes	36Yes	Yes
Covariates	Yes	Yes	Yes	Yes
Constant	-0.094 (0.077)	-0.094 (0.077)	-0.088 (0.074)	-0.088 (0.074)
R2	0.361	0.361	0.367	0.367
Partial F	86.13	83.13	988.65	988.65
Number of observations	1,472	1,472	1,472	1,472

Notes: the covariates included are Marital Status of head, Family size, Age of Head, Education of Head, Primary activity of head, credit constraint, Maize farm area, and Total farm area: Standard errors clustered at the sub-kebele level and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

4.5.2 PPT adoption and herd size

Table 3 summarizes the LATE estimates for the entire sample, with Panel A illustrating the outcomes pertaining to herd size. At both midline (columns 1-3) and end-line (columns 4-6), the adoption of PPT significantly increased livestock holdings. On average, adopters owned 3.2 more animals at midline and 6.5 more at end-line compared to the control group, with these increases being statistically significant ($p = 0.00$). Specifically, at midline, adopters owned about 0.8 more cattle and 2.4 more small ruminants than non-adopter leader farmers. By the endline, these figures had further risen to 1.2 additional cattle and 5.3 more small livestock.

These patterns indicate that households adopting PPT expanded their herds over time, likely driven by improved forage availability. The notable increase in small livestock suggests that farmers may be using PPT to support livestock fattening activities. Additionally, the larger relative gains in small ruminants imply that households may find it more affordable and manageable to invest in goats and sheep rather than cattle. This could be due to financial constraints, as the income gains from PPT might not be sufficient to cover the higher costs associated with acquiring and maintaining cattle.

The ITT results in Appendix Table C2 generally agree with the LATE analysis.

Table 3. Push-pull technology adoption impacts on households' herd size (Full sample estimations) (2SLS)

	Midline			End line		
Panel A	*Second stage*					
	Total herding size (numbers)	Cattle herd size (numbers)	Goat and sheep herd size (numbers)	Total herding size (numbers)	Cattle herd size (numbers)	Average Goat and sheep herd size (numbers)
Experiment	3.186*** (0.601)	0.810* (0.421)	2.376** (0.392)	6.455*** (0.632)	1.204** (0.439)	5.256*** (0.480)
Village Fixed effects.	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.346*** (1.235)	1.932 (1.132)	1.415** (0.609)	1.639 (1.112)	1.480 (1.172)	0.165 (0.691)
Mean of dependent variable	9.036	5.537	3.499	9.900	5.936	3.953
Panel B	*First stage (Experiment with PPT package as dependent variable)*					
Subsidy	0.293*** (0.035)	0.293*** (0.035)	0.293*** (0.035)	0.293*** (0.035)	0.293*** (0.035)	0.293*** (0.035)
Incentive	0.146*** (0.048)	0.146*** (0.048)	0.146*** (0.048)	0.146*** (0.048)	0.146*** (0.048)	0.146*** (0.048)
Village F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.078 (0.081)	-0.078 (0.081)	-0.078 (0.081)	-0.078 (0.081)	-0.078 (0.081)	-0.078 (0.081)
R2	0.355	0.355	0.355	0.355	0.355	0.355
Partial F	53.63	53.63	53.63	53.63	53.63	53.63
Number of observations	1,472	1,472	1,472	1,472	1,472	1,472

Notes: the covariates included are Marital Status of head, Family size, Age of Head, Education of Head, Primary activity of head, credit constraint, Maize farm area, and Total farm area: Standard errors clustered at the sub-kebele level and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

4.5.3. PPT adoption and income from feed and livestock sales

Table 4 provides a comprehensive overview of the estimated income generated from feed and livestock sales by participating households. The implementation of Push-Pull Technology (PPT) has substantially increased income from these sales. During the midline assessment, households reported estimated earnings of 1,212 ETB (USD 22) from feed sales, which rose to 1,840 ETB (around USD 33) by the endline[8].

The increase in income can largely be attributed to the surplus feed generated through the adoption of the PPT practice. As highlighted in the descriptive statistics section, households that embraced this approach—particularly those with larger cultivated areas—showed a greater capacity for selling feed. This indicates that these households were able to produce surplus feed from a combination of crop residues, PPT forage[9], and hay. PPT can now provide adequate forage for their livestock throughout the year, especially with desmodium, which can be harvested consistently. This, along with crop residues and hay, not only satisfies the livestock's needs but also offers households a surplus feed for sale.

The rise in income from feed sales over time is linked to improved forage production as PPT plots become more established. As these plots mature, they yield greater quantities of forage, allowing households to satisfy their livestock demands while also selling surplus feed derived from a combination of PPT forage, crop residues, and hay. Indeed, Figures 2 and 3 in the follow-up data section illustrate that some households that did not engage in feed sales at the midline began doing so by the endline, as their PPT plots became more mature.

The implementation of Push-Pull Technology (PPT) has also substantially increased income from livestock sales. During the midline assessment, households reported estimated earnings of 1,144 ETB (USD 21), which rose to 1,483 ETB (around USD 27) by the endline[10].

Appendix Table C3 presents the ITT results assessing the impact of the interventions on household income generated from feed and livestock production and sales. The ITT findings consistent with the results from the LATE analysis.

---

[8] These estimates account for inflation between the periods.

[9] Surplus forage from both Brachiaria and Desmodium can be processed into hay for use during the dry seasons.

[10] These estimates account for inflation between the periods.

Table 4. Impact of PPT adoption on income from feed and livestock sales (2SLS)

	Midline		Endline	
Panel A	*Second stage*			
	Income obtained from feed sale (ETB)	Income obtained from livestock sale (ETB)	Income obtained from feed sale (ETB)	Income obtained from livestock sale (ETB)
Experiment	1212*** (283)	1144** (530)	1840*** (308)	1483*** (458)
Village F.E.	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
Constant	344 (591)	2670** (1088)	315 (561)	3660** (1623)
Mean of dependent variable	998	5108	1054	5265
Panel B	*First stage (Experiment with PPT)*			
Subsidy	0.293*** (0.035)	0.293*** (0.035)	0.293*** (0.035)	0.293*** (0.035)
Incentive	0.146*** (0.048)	0.146*** (0.048)	0.146*** (0.048)	0.146*** (0.048)
Village F.E.	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
Constant	-0.078 (0.081)	-0.078 (0.081)	-0.078 (0.081)	-0.078 (0.081)
R2	0.355	0.355	0.355	0.355
Partial F	53.63	53.63	53.63	53.63
Number of observations	1,472	1,472	1,472	1,472

Notes: The covariates included are Marital Status of head, Family size, Age of Head, Education of Head, Primary activity of head, credit constraint, Maize farm area and Total farm area : Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

4.5. Discussion and Conclusions

Smallholder farmers in economically disadvantaged countries face significant challenges within mixed crop-livestock farming systems. These challenges include pest infestations that negatively impact crop yields and insufficient feed supplies for livestock. In order to address these pressing issues and promote food security and sustainable resource management, development experts advocate for the adoption of a diverse range of productivity-enhancing inputs and practices among these farmers.

However, adoption rates in many countries have fallen short of expectations. This discrepancy can largely be attributed to inadequate information flow, leaving many farmers unaware of innovative practices or uncertain about how to implement them effectively. In response, policymakers often provide subsidies to encourage the adoption of new technologies. These subsidies allow households to evaluate the associated costs and benefits, helping to alleviate uncertainties surrounding innovative strategies such as Push-Pull Technology (PPT). Research by Balew et al. (2025) indicates that households targeted by these subsidies respond in two main ways: some trial users of PPT demonstrate increased interest in purchasing it for continued use, while others believe the technology should be offered at no cost, reflecting a reluctance to fully embrace the innovation.

In light of this context, we conducted an assessment of the effects of Push-Pull Technology on livestock productivity and household income through a cluster randomized controlled trial involving 754 households in northwestern Ethiopia. The results provide compelling evidence of PPT's beneficial impact on both livestock productivity and household income among mixed crop-livestock farming households in the region. These benefits complement the PPT gains due to reduced pest damage (or reduced pesticide expenditures) that are the initial focus of PPT.

PPT offers several significant benefits for agricultural practices beyond reducing pest damage. Notably, it provides forage for both personal use and sale, which has important economic implications for smallholder farmers. By increasing the availability of forage resources, PPT enhances livestock production and boosts household incomes, thereby strengthening farmers' livelihoods and providing a more stable economic base.

In our analysis, we found that while the Local Average Treatment Effect (LATE) estimates yield larger coefficients due to their robust methodology, the coefficients from the Intent-to-Treat (ITT) estimation are also positive and predominantly significant. This suggests that the intervention has successfully met its objectives. Moreover, the coefficients are generally higher in the group receiving subsidy- plus- incentive compared to those receiving subsidy alone. This highlights the critical role of additional experimental incentives in promoting the widespread adoption of PPT, demonstrating the effectiveness of such strategies in enhancing farmer participation.

Despite encountering initial obstacles to adoption, such as the labor demands associated with PPT implementation and the necessity for targeted training, the strategy used by this study has proven successful in overcoming these challenges. Key to this success has been the provision of free seeds, coupled with experimentation incentives and comprehensive educational support. These interventions have effectively reduced barriers to entry and facilitated smoother adoption processes.

The findings from this study provide compelling evidence for advocating the broader dissemination and adoption of PPT as a promising strategy to enhance agricultural productivity and improve food security among Ethiopian households. Moving forward, it is essential for future research to delve deeper into the impacts of PPT adoption, particularly concerning crop production and additional associated benefits. Such research will be instrumental in further fostering farmer confidence in this innovative technology. Additionally, it is crucial to thoroughly explore and quantify the associated costs of PPT implementation, as understanding these financial aspects will play a vital role in encouraging wider acceptance and maximizing the benefits for rural communities throughout Ethiopia. Through these efforts, we can significantly contribute to the resilience and prosperity of these agricultural systems.

4.6. References

Abebe, H., Alemu,T. and Gedamu,Y. (2023). "Forage Yield and Quality of Desmodium dichotomum Accessions in Eastern Amhara, Ethiopia". *Journal of Rangeland Science*, Vol. 13, No. 2.

Amejo, A. G., Gebere, Y. M., Kassa, H., & Tana, T. (2018). "Agricultural productivity, land use, and draft animal power formula derived from mixed crop-livestock systems in Southwestern Ethiopia." *African Journal of Agricultural Research*, 13(42), 2362–2381. https://doi.org/10.5897/AJAR2018.132

Assefa, S., Duguma, B., & Worku, Z. (2024). "Assessment of livestock husbandry practices and production constraints among smallholder mixed crop-livestock production systems in the Majang Zone, Southwest Ethiopia." *Heliyon*, Volume 10, Issue 19, Article e37400, published on October 15, 2024. Available at: https://doi.org/10.1016/j.heliyon.2024.e37400

Ayalew, H., Tamru, G., & Abebe, D. (2018). "Beef cattle fattening practices and marketing systems in Gondar Town, Amhara, Ethiopia." *Journal of Veterinary Science & Technology*.

Balew, S., Bulte, E., Abro, Z., & Kassie, M. (2023). "Incentivizing and nudging farmers to spread information: Experimental evidence from Ethiopia." *American Journal of Agricultural Economics*, 105, 994-1010.

Balew, S., Bulte, E., & Kassie, M. (2025). "Short-run subsidies and long-run willingness to pay: Learning and anchoring in an agricultural experiment in Ethiopia." *American Journal of Agricultural Economics*, 107(2), 655-669.

Belay, G. and Negesse,T.(2018)"Feed Resource Availability and their Nutrient Contribution for livestock Evaluated Using Feed Assessment Tool (FEAST) in Burie Zuria District, North Western Ethiopia." *Agricultural research and technology*, Volume 17 Issue 3 - August 2018 DOI: 10.19080/ARTOAJ.2018.17.556022.

Bereda, A., Yilma, Z., & Nurfeta, A. (2014). "Dairy production system and constraints in Ezha Districts of the Gurage Zone, Southern Ethiopia." *Global Veterinaria*, 12, 181–186.

Cheruiyot, Duncan; Midega, Charles A. O.; Pittchar, Jimmy O.; Pickett, John A.; and Khan, Zeyaur R. (2020). "Farmers' Perception and Evaluation of Brachiaria Grass (Brachiaria spp.) Genotypes for Smallholder Cereal-Livestock Production in East Africa." *Agriculture*.

CSA (2022). "Livestock and Livestock Characteristics." Addis Ababa; 2022, Report No. 594, pp. 1–219.

Dadi, G., & Worku, A. (2023). "Socio-Economic Importance of Sheep Production in Central Rift Valley of Oromia Regional State, Ethiopia."

American Journal of Environmental and Resource Economics, 8(3), 42-51. https://doi.org/10.11648/j.ajere.20230803.11

Duressa, D., Kenea, D., Keba, W., Desta, Z., Berki, G., Leta, G., & Tolera, A. (2014). "Assessment of livestock production system and feed resources availability in three villages of Diga district, Ethiopia." Working Paper, International Livestock Research Institute (ILRI).

FAO (2020). "Quality Declared Seed System." Plant Production and Protection Paper No. 185. Rome: FAO; 2020.

Fetene, G. M., Abro, Z., Getahun, T. D., & Kassie, M. (2025). "Impact of feed shortages on livestock and crop production in Ethiopia: Implications for rural poverty reduction." *European Review of Agricultural Economics*. https://doi.org/10.1093/erae/jbaf006

Gizachew, N., Goshu, M., Wako, A., & Berhanu, G. (2020). "Beef cattle production management practices and constraints in Arsi Negelle District, Ethiopia." *Global Veterinaria*, 22(4), 233–241.

Gobezie, E., Bekele, E., & Tadesse, I. (2020). "Constraints and opportunities of beef cattle fattening practices in three selected districts of Gamo Zone, Southern Ethiopia." *Global Veterinaria*, 22(3), 128-138.

Guja , A., Yoseph Mekasha Gebere, Habtemariam Kassa, and Tamado Tana (2019). "Characterization of smallholder mixed crop–livestock systems integrated with spatial information: A case study from Ethiopia." *Cogent Food & Agriculture*, 5(1), 1565299. doi:10.1080/23311932.2019.1565299.

Hadush, M. (2019). "The economic implications of grazing and water resource scarcity on household welfare and food security in Tigray, Ethiopia." *Agricultural and Resource Economics Review*, 48(1), 170–198.

Hadush, M. (2020). "The impact of animal resource scarcity on farm labor and production in Northern Ethiopia." *Journal of Agricultural and Applied Economics*, 52(2), 264–287.

Isgren, Ellinor; Clough, Yann; Murage, Alice; and Andersson, Elina (2023). "Are Agricultural Extension Systems Ready to Scale Up Ecological Intensification in East Africa? A Literature Review with a Focus on Push-Pull Technology (PPT)." *Food Security*, 15, 1399–1420. https://doi.org/10.1007/s12571-023-01387-z

Kalsa, K. K., & Dey, B. (2022). "Forage seed system performance of Ethiopia: An overview based on key indicators." *CABI Reviews*, 17(055). https://doi.org/10.1079/cabireviews202217055

Kassie, M., Stage, J., Diiro, G., Muriithi, B., Muricho, G., Ledermann, S., Pitcharr, J., Midega, C., & Kahn, Z. (2018). "Push-Pull Farming System in Kenya: Implications for Economic and Social Welfare." *Land Use Policy*, 77, 186–198.

Khan, Z., Midega, C., Pittchar, J., Pickett, J., & Bruce, T. (2011). "'Push–pull technology: a conservation agriculture approach for integrated management of insect pests, weeds, and soil health in Africa"' . *International Journal of Agricultural Sustainability.*

Khan, Zeyaur R.; Midega, Charles A. O.; Pittchar, Jimmy O.; and Pickett, John A. (2014). "Push–Pull: A Novel IPM Strategy for the Green Revolution in Africa." Chapter 13.

Mayberry, D., Ash, A., Prestwidge, D., Godde, C. M., Henderson, B., Duncan, A., Blum-mel, M., Ramana Reddy, Y., & Herrero, M. (2017). "Yield gap analyses to estimate attainable bovine milk yields and evaluate options to increase production in Ethiopia and India." *Agricultural Systems,* 155, 43–51.

Mayberry, D., Ash, A., Prestwidge, D., & Herrero, M. (2018). "Closing yield gaps in smallholder goat production systems in Ethiopia and India." *Livestock Science*, 214, 238–244.

Ogot NO, Pittchar JO, Midega CAO & Khan ZR.(2018). "Rearing on Advancing Household Food Security and Nutrition: A Study of Push-Pull Technology and Livestock Production". *EC Nutrition* 13.6 (2018): 374-387.

Pickett, A. J., Woodcock, C. M., Midega, C. A. O., & Khan, Z. R. (2014). "Push–pull farming systems." *Current Opinion in Biotechnology*, 26, 125–132.

Shapiro, B. I., Gebru, G., Desta, S., Negassa, A., Nigussie, K., Aboset, G., & Mechale, H. (2017). "Ethiopia livestock sector analysis." ILRI Project Report. Nairobi, Kenya: International Livestock Research Institute (ILRI).

Tegegne, A., Gebremedhin, B., Hoekstra, D., Belay, B., & Mekasha, Y. (2013). "Smallholder dairy production and marketing systems in Ethiopia: IPMS experiences and opportunities for market-oriented development." Working Paper No. 31.

Tigistu, T. (2021). "Forage Production and Utilization Training," August 2021. Addis Ababa, Ethiopia.

Wakeyo, M. B., & Elias, H. (2023). "Mixed farming system for crop yield improvement and adaptation to climate change: Evidence from smallholder farmers in Ethiopia." *Ethiopian Journal of Economics,* 32(2).

Yadav, S. K., Narwade, D.. K., Panigrahi, C.. K., Kumar, G., Rout, S., Shanker, P., Chaturvedi, K., Tiwari, A. K., & Pandey, S. K. (2025). "Push-Pull Strategies and Habitat Manipulation for Sustainable Insect Pest Management in Crops." *Uttar Pradesh Journal of Zoology*, 46(3), 1–16. ISSN: 0256-971X.

Zegeye, Y. (2023). "Imperatives and challenges of dairy production, processing, and marketing in Ethiopia." Challenges and Opportunities in Livestock Marketing in Ethiopia. Addis Ababa, Ethiopia: ESAP; 2003. pp. 1–389.

Appendices

Table A1: Baseline balance test for leaders

Variables;-	(1) Control	(2) Only Subsidy	(3) Subsidy +Incentive	(1)-(2) Pairwise t-test	(1)-(3) Pairwise t-test	(2)-(3) Pairwise t-test
	Mean/(SE)	Mean/(SE)	Mean/(SE)	Mean difference	Mean difference	Mean difference
Covariates ;-						
Marital Status of head(1=married)	0.97	0.98	0.99	-0.00	-0.01	-0.01
	(0.01)	(0.01)	(0.01)			
Family size(number)	5.42	5.65	5.68	-0.23	-0.25	-0.03
	(0.13)	(0.13)	(0.10)			
Age of head (years)	43.69	45.30	44.27	-1.60	-0.58	1.03
	(0.66)	(0.72)	(0.51)			
Education of head (1=literate)	0.52	0.62	0.57	-0.09*	-0.05	0.05
	(0.04)	(0.03)	(0.03)			
Primary activity of head(1=farming)	0.98	0.99	0.99	-0.01	-0.01	-0.00
	(0.01)	(0.01)	(0.00)			
Have credit constraint(1=yeas)	0.44	0.36	0.32	0.09	0.12*	0.04
	(0.04)	(0.03)	(0.02)			
Maize farm area(hectare)	0.43	0.45	0.43	-0.02	0.00	0.02
	(0.02)	(0.02)	(0.01)			
Total farm area(hectare)	1.20	1.30	1.29	-0.10	-0.09	0.01
	(0.05)	(0.05)	(0.04)			
Outcome variables;-						
cows(number)	0.79	0.83	0.84	-0.04	-0.04	-0.00
	(0.07)	(0.07)	(0.05)			
Calves(number)	0.79	0.83	0.84	-0.04	-0.04	-0.00
	(0.07)	(0.07)	(0.05)			
Heifers (number)	1.05	1.22	1.32	-0.17	-0.27**	-0.10
	(0.07)	(0.08)	(0.06)			
Bulls(number)	0.63	0.76	0.85	-0.13	-0.22*	-0.09
	(0.07)	(0.07)	(0.05)			
Oxen (number)	1.69	1.66	1.67	0.03	0.02	-0.01
	(0.05)	(0.05)	(0.04)			
Cattle(number)	4.95	5.30	5.50	-0.35	-0.55*	-0.20
	(0.25)	(0.24)	(0.18)			
Goats (number)	1.55	1.31	1.61	0.23	-0.06	-0.30*
	(0.13)	(0.10)	(0.10)			
Sheep (number)	1.52	1.54	1.60	-0.02	-0.08	-0.06
	(0.09)	(0.10)	(0.07)			
Goat and sheep(Number)	3.07	2.86	3.21	0.21	-0.15	-0.36**
	(0.13)	(0.12)	(0.10)			
Milk production per cow per day(litter)	0.93	0.96	93	-0.03	0.00	0.03
	(0.09)	(0.08)	(0.06)			
Lactation length per cow(months)	5.31	5.43	5.41	-0.12	-0.16	0.02
	(0.43)	(0.41)	(0.30)			

Feed production and sales income (ETB)	876	981.43	832.10	-113.79	35.54	149.34
	(122.29)	(111.45)	(81.20)			
Livestock production and sales income (ETB)	4653	4781	4818	-128	-165	-37
	(2971)	(3102)	(3003)			
Number of observations	183	201	370	384	553	571

Notes: Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

Table A2: Baseline balance test for leaders (control group vs households assigned to treatment group)

Variables;-	(1)	(2)	(1)-(2)
	Control	treatment group	Pairwise t-test
	Mean/(SE)	Mean/(SE)	Mean difference
Covariates ;-			
Marital Status of head(1=married)	0.97	0.98	-0.01
	(0.01)	(0.01)	
Family size(number)	5.42	5.67	-0.24
	(0.13)	(0.08)	
Age of head (years)	43.69	44.63	-0.94
	(0.66)	(0.42)	
Education of head (1=literate)	0.52	0.59	-0.06
	(0.04)	(0.02)	
Primary activity of head(1=farming)	0.98	0.99	-0.01
	(0.01)	(0.00)	
Have credit constraint(1=yeas)	0.46	0.36	0.10*
	(0.04)	(0.02)	
Maize farm area(hectare)	0.43	0.44	-0.01
	(0.02)	(0.01)	
Total farm area(hectare)	1.20	1.29	-0.09
	(0.05)	(0.03)	
Outcome variables;-			
Cows(number)	0.79	0.83	-0.04
	(0.07)	(0.04)	
Calves(number)	0.79	0.83	-0.04
	(0.07)	(0.04)	
Heifers (number)	1.05	1.29	-0.24**
	(0.07)	(0.05)	
Bulls(number)	0.63	0.81	-0.19*
	(0.07)	(0.04)	
Oxen (number)	1.69	1.67	0.02
	(0.05)	(0.03)	
Cattle(number)	4.95	5.43	-0.48
	(0.25)	(0.14)	
Goats (number)	1.55	1.51	0.04
	(0.13)	(0.07)	
Sheep (number)	1.52	1.58	-0.06
	(0.09)	(0.06)	
Goats and sheep(number)	3.07	3.09	-0.02
	(0.13)	(0.08)	
	(66.63)	(36.61)	
Milk production per cow per day(litter)	0.93	0.94	-0.01
	(0.09)	(0.05)	
Lactation length per cow(months)	5.31	5.46	-0.15
	(0.43)	(0.24)	
Feed production and sales income (ETB)	876	895.79	-20
	(122.29)	(66.34)	
Livestock production and sales income (ETB)	4653	4805	-152
	(2971)	(3035)	
Number of observations	183	571	754

Notes: Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

Table A3: Baseline balance test for leaders (non-experimenters vs households experimenting with PPT)

Variables;-	(1)	(2)	(1)-(2)
	Non experimenters	Experimenters	Pairwise t-test
	Mean/(SE)	Mean/(SE)	Mean difference
Covariates ;-			
Marital Status of head(1=married)	0.99	0.98	0.01
	(0.01)	(0.01)	
Family size(number)	5.48	5.97	-0.49***
	(0.09)	(0.13)	
Age of head (years)	44.70	44.51	0.19
	(0.54)	(0.65)	
Education of head (1=literate)	0.56	0.63	-0.07*
	(0.03)	(0.03)	
Primary activity of head(1=farming)	0.99	1.00	-0.00
	(0.00)	(0.00)	
Have credit constraint(1=yeas)b	0.29	0.41	-0.11***
	(0.02)	(0.03)	
Maize farm area(hectare)	0.42	0.47	-0.05**
	(0.02)	(0.02)	
Total farm area(hectare)	1.23	1.39	-0.16**
	(0.04)	(0.05)	
Outcome variables;-			
Cows(number)	0.78	0.92	-0.14
	(0.05)	(0.07)	
Calves(number)	0.78	0.92	-0.14
	(0.05)	(0.07)	
Heifers (number)	1.21	1.42	-0.21**
	(0.06)	(0.08)	
Bulls(number)	0.84	0.77	0.06
	(0.06)	(0.07)	
Oxen (number)	1.63	1.72	-0.09
	(0.04)	(0.05)	
Cattle(number)	5.24	5.76	-0.53*
	(0.19)	(0.22)	
Goats (number)	1.29	1.87	-0.58***
	(0.08)	(0.15)	
Sheep (number)	1.69	1.40	0.28**
	(0.07)	(0.09)	
Goats and sheep(number)	2.97	3.27	-0.30*
	(0.09)	(0.15)	
Milk production per cow per day(litter)	0.91	0.99	-0.08
	(0.06)	(0.07)	
Lactation length per cow(months)	5.25	5.80	-0.55
	(0.30)	(0.40)	
Feed production and sales income (ETB)	932.26	833.32	98.94
	(85.70)	(104.37)	
Livestock production and sales income (ETB)	4668	5037	-369*
	(2983)	(3115)	
Number of observations	359	212	571

Notes: Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively

Table A4: Follow up data and descriptive statistics

	Midline		Endline	
Outcome Variables;-	Mean	Sta. deviation	Mean	Sta. deviation
Cows(number)	0.94	0.86	1.10	0.84
Calves(number)	0.98	0.88	1.13	0.86
Heifers (number)	1.29	1.03	1.42	0.72
Bulls(number)	0.93	0.99	1.14	0.91
Oxen (number)	1.58	0.83	1.72	0.57
Cattle(number)	5.71	2.35	6.50	2.65
Goats (number)	1.90	1.68	2.49	2.27
Sheep (number)	2.11	1.35	2.32	1.55
Goats and sheep(number)	4.01	2.07	4.81	2.52
Milk production per cow per day(Litters)	1.44	1.17	1.47	1.10
Lactation length per cow(Months)	7.58	5.85	8.50	5.62
Feed production and sales income (ETB)	1149	2017	1260	1998
Livestock production and sales(ETB)	5377	2724	5690	3385
Number of observation	736		736	

Table A5: Attrition analysis of group leaders (OLS)

Dependent variables	Attrition status of group leaders
Subsidy (1= yes)	-0.007 (0.015)
Incentive (1= yes)	-0.019 (0.014)
Sex of respondent (1=male)	-0.021 (0.046)
Marital status of household head (1=married)	-0.051 (0.071)
Household members (number)	0.004 (0.003)
Household head age (years)	-0.001 (0.000)
Household Primary activity farming (1=yes)	-0.096 (0.095)
Household head education (years)	-0.002 (0.009)
Lived in the village (Months)	0.004 (0.006)
Maize plot (hectares)	-0.019 (0.022)
Total land owned by household (hectares)	0.005 (0.014)
Knowledge of stemborer (1= yes)	-0.007 (0.031)
Stemborer major problem (1= yes)	0.0255)
Knowledge of fall armyworm (1= yes)	0.002 (0.014)
Fall armyworm major problem (1= yes)	0.022 (0.023)
Knowledge of striga (1= yes)	0.001 (0.019)
Striga major problem (1= yes)	-0.026 (0.016)
Household owns milking cow (1 =yes)	0.020 (0.013)
Household faces credit constraint (1=yes)	0.016 (0.011)
Village fixed effects	Yes
Mean dependent variable	0.024
Pseudo r-squared	0.099
Akaike crit. (AIC)	-661.083
SD dependent var	0.153
Number of observations	754
Bayesian crit. (BIC)	-402.061
F test for joint significance (p-value)	0.550

Notes: Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

Figures A: The proportion of livestock (TLU) to cultivated land (hectares) among adopting households

Figure A.1: Midline

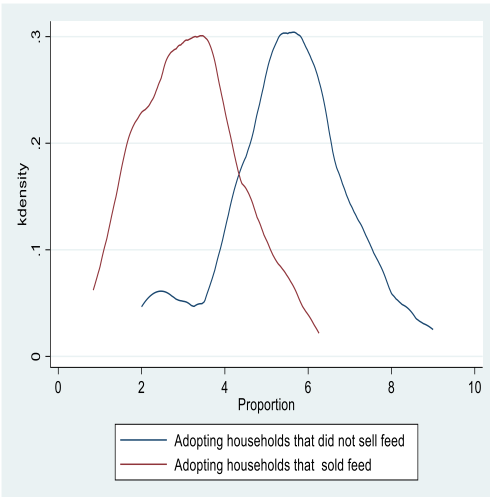


Figure A.2: Endline

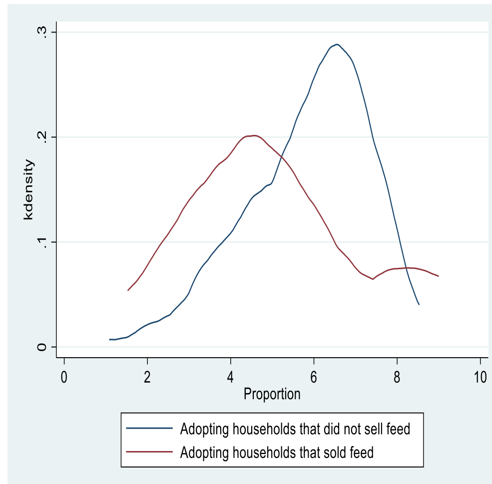


Table B1: PPT adoption, total herding size and milk production(Sub- sample estimations, 2SLS)

	Midline			End line		
Panel A	*Second stage*					
	Total herding size (numbers)	Average milk production per cow in a day (liters)	Average lactation length per cow (months)	Total herding size (numbers)	Average milk production per cow in a day (liters)	Average lactation length per cow (months)
Experiment	3.140*** (0.604)	0.872*** (0.089)	2.404*** (0.131)	6.633*** (0.640)	0.824*** (0.100)	1.779*** (0.313)
Village F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.231*** (1.239)	1.523*** (0.210)	10.432*** (0.273)	2.036* (1.111)	1.208*** (0.312)	7.856*** (1.037)
Mean of dep.var	9.043	2.11	11.795	9.968	2.134	11.270
Panel B	*First stage (Experiment with PPT package as dependent variable)*					
Subsidy	0.297*** (0.037)	0.368*** (0.050)	0.368*** (0.050)	0.293*** (0.035)	0.348*** (0.041)	0.348*** (0.041)
Incentive	0.142*** (0.049)	0.104* (0.064)	0.104* (0.064)	0.147*** (0.048)	0.136*** (0.050)	0.136*** (0.050)
Village F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.080 (0.081)	0.002 (0.135)	0.002 (0.135)	-0.077 (0.081)	0.020 (0.117)	0.020 (0.117)
R2	0.356	0.460	0.460	0.355	0.397	0.397
Partial F	48.62	175.11	175.11	53.30	265.090	265.090
Number of observation	1,470	816	816	1,471	883	883

Notes: the covariates included are Marital Status of head, Family size, Age of Head, Education of Head, Primary activity of head, credit constraint, Maize farm area and Total farm area : Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

Table B2: PPT adoption and Cattle herd size by class

(Full- sample estimations, 2SLS)

Panel A	Second stage			
	Midline		Endline	
	Total herding size big cattle/cows plus oxen (numbers)	Total herding size small cattle/caves, heifers plus bulls (numbers)	Total herding size big cattle/cows plus oxen (numbers)	Total herding size small cattle/caves, heifers plus bulls (numbers)
Experiment	-0.026 (0.248)	0.836*** (0.279)	0.213 (0.184)	0.989*** (0.302)
Village F.E.	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
Constant	1.255*** (0.431)	0.676 (0.851)	1.090** (0.492)	0.387 (0.776)
Mean of dependent variable	2.523	3.014	2.665	3.271
Panel B	First stage (Experiment with PPT package as dependent variable)			
Subsidy	0.293*** (0.035)	0.293*** (0.035)	0.293*** (0.035)	0.293*** (0.035)
Incentive	0.146*** (0.048)	0.146*** (0.048)	0.146*** (0.048)	0.146*** (0.048)
Village F.E.	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
Constant	-0.078 (0.081)	-0.078 (0.081)	-0.078 (0.081)	-0.078 (0.081)
R2	0.355	0.355	0.355	0.355
Partial F	53.63	53.63	53.63	53.63
Number of observations	1,472	1,472	1,472	1,472

Notes: the covariates included are Marital Status of head, Family size, Age of Head, Education of Head, Primary activity of head, credit constraint, Maize farm area and Total farm area : Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

Table C1: PPT adoption and milk production (ITT estimations)

	Midline		Endline	
	ITT estimations(OLS)			
Midline	Average milk production per cow in a day (litters)	Average lactation length per cow (months)	Average milk production per cow in a day (litters)	Average lactation length per cow (months)
Only Subsidy	0.165** (0.035)	0.472*** (0.0.65)	0.167** (0.057)	0.058 (0.175)
Subsidy plus incentive	0.256*** (0.042)	0.677*** (0.068)	0.279*** (0.047)	0.554*** (0.126)
Village F.E.	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
R2	0.858	0.989	0.812	0.949
Mean of dep.var	1.209	6.539	1.280	6.997
Coefficient test (p-value)	0.046	0.019	0.109	0.018
Number of observation	1,472	1,472	1,472	1,472

Notes: the covariates included are Marital Status of head, Family size, Age of Head, Education of Head, Primary activity of head, credit constraint, Maize farm area and Total farm area : Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

Table C2: PPT adoption and households' herd size (ITT estimations)

	Midline			End line		
	ITT estimations(OLS)					
	Total herding size (numbers)	Cattle herd size (numbers)	Goat and sheep herd size (numbers)	Total herding size (numbers)	Cattle herd size (numbers)	Goat and sheep herd size (numbers)
Only Subsidy	1.111** (0.289)	0.314* (0.186)	0.798** (0.167)	2.450*** (0.303)	0.657** (0.202)	1.789*** (0.162)
Subsidy plus incentive	1.420*** (0.250)	0.365** (0.185)	1.055** (0.137)	2.902*** (0.306)	0.667** (0.219)	2.337*** (0.167)
Village F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.270	0.223	0.217	0.305	0.220	0.339
Mean of dep.var	9.036	5.537	3.499	9.900	5.936	3.953
Coefficient test (p-value)	0.340	0.807	0.212	0.249	0.736	0.015
Number of observation	1,472	1,472	1,472	1,472	1,472	1,472

Notes: the covariates included are Marital Status of head, Family size, Age of Head, Education of Head, Primary activity of head, credit constraint, Maize farm area and Total farm area : Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

Table C3: Impact of PPT adoption on households' income from forage and livestock sales (ITT estimations)

	Midline		Endline	
	ITT estimations(OLS)			
	Income obtained from feed sale (ETB)	Income obtained from livestock sale (ETB)	Income obtained from feed sale (ETB)	Income obtained from livestock sale (ETB)
Only Subsidy	443*** (153)	151 (233)	481*** (134)	522** (232)
Subsidy plus incentive	544*** (111)	448** (215)	799*** (126)	675*** (190)
Village F.E.	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
R2	0.289	0.200	0.280	0.208
Mean of dep.var	998	5108	1054	5265
Coefficient test (p-value)Coefficient test (p-value)	0.546	0.252	0.049	0.855
Number of observation	1,472	1,472	1,472	1,472

Notes: the covariates included are Marital Status of head, Family size, Age of Head, Education of Head, Primary activity of head, credit constraint, Maize farm area and Total farm area : Standard errors clustered at the sub-kebele level, and reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level, respectively.

Chapter 5

A Tale of Framing and Screening: How Health Messaging and House Screening Affect Malaria Transmission in Ethiopia

Abstract: Malaria is a major public health problem in Africa. Traditional methods of controlling malaria no longer provide adequate protection against transmission, and future approaches likely require a combination of technical solutions and behavioral change. We use a cluster randomized controlled trial to study the impacts of an intervention that combines house screening with a behavioral intervention based on health messaging. While house screening provides modest positive benefits, these benefits can be leveraged if it is combined with health messaging. We provide tentative evidence that the impact of messaging varies with the design of the choice architecture: loss-framed health messages seem to do better than gain-based messages—our data suggest they may have larger and more durable effects on behavior and health outcomes.

Keywords: House screening, improved housing, behavioral interventions, nudging, loss-framed messaging, vector control.

JEL Codes: D10, I15

‘

Publication Status: Balew, S., Bulte, E., Abro, Z., Asale, A., & Mutero, C. (2025). “A tale of framing and screening: How health messaging and house screening affect malaria transmission in Ethiopia.” *Journal of Development Economics*, *172*, 1–20. https://doi.org/10.1016/j.jdeveco.2024.103407

5.1. Introduction

According to the World Health Organization, malaria affects nearly 250 million people annually and causes over 600 thousand deaths —especially young children, mainly in sub-Saharan Africa (WHO 2021a). There is a correlation between malaria and numerous underdevelopment indicators (Sachs and Malaney 2022), and malaria's geographical distribution partly accounts for the income gap between industrialized and tropical countries (Gallup and Sachs 2000; Bleakley 2009, Cutler et al. 2010; Sarma et al. 2019; Bloom et al. 2022). Malaria reduces labor supply and productivity in the short-term via worker morbidity and increased care responsibilities. Long-term effects emerge via reduced educational outcomes (e.g., Lucas 2010; Barofsky et al. 2015; Bleakley 2003, 2010). Despite successful malaria elimination in large parts of the world, its prevalence in Africa is increasing. Previous health gains attributed to combined effects of (subsidized) insecticide-treated bed nets (ITNs), antimalarials, and indoor residual spraying (IRS) (e.g., Dupas 2014, Tarozzi et al. 2014, Cohen et al. 2015, Pinder et al. 2016, Pryce et al. 2018) have stalled and are now slowly being reversed, largely due to increasing insecticide resistance in mosquitoes (see Lindsay et al. 2021 for a discussion). Although RNA vaccines offer hope for the future, widespread accessibility for target populations will take years (Piper 2021).[1]

Additional vector control measures are needed to combat malaria transmission (Killeen et al. 2019), and housing improvements are seen as a promising yet underutilized approach (Tusting et al. 2015). Evidence suggests that fewer mosquitos enter "modern houses" and that living in modern houses is associated with lower odds of malaria infection than living in "traditional houses".[2] A practical shortcut to (costly) improvement of houses in malaria-prone areas is house screening—the placement of fine-mazed screens in front of houses' doors, windows, and eaves. Several studies evaluate the effects of house screening on mosquito densities inside houses and the prevalence of malaria.[3] While the effectiveness of house screening varies with the behavioral responses of the target population and context (e.g., other vector control measures in place), the World Health Organization recommends it in areas with a high disease burden (WHO 2021b). Untreated screens are environmentally friendly and do not contribute to resistance. They are more durable than bed nets, less susceptible to damage due to wear and tear, and provide protection during a longer period of the day (when people are indoors but not in bed). They also provide equal protection to all household members inside the house. A disadvantage is that untreated screens do not kill mosquitos. In what follows we use the terms screening and house screening interchangeably and don't use "screening" to mean anything else than a barrier to keep mosquitos out.

In this paper we use experimental data to analyze the health and economic impacts of house screening and examine how these can be leveraged by information and design of the choice architecture. We analyze how house screening and health messaging affect malaria outcomes,

[1] Currently, the WHO recommend two malaria vaccines for use in children living in moderate to high malaria transmission areas. These vaccines provide partial protection, and reduce uncomplicated malaria by approximately 40%, severe malaria by some 30%, and all-cause mortality by 13%. The WHO also recommend that malaria vaccines are delivered in conjunction with other control interventions, such as house screening and ITNs (CDC 2024).
[2] Features of modern houses are closed eaves, brick walls, tiled or metal roofs, and ceilings (Tusting et al. 2015).
[3] Examples include: Lindsay et al. (2003); Kirby et al. (2009); Massebo and Lindtjorn (2013); Mburu et al. (2018); Killeen et al. (2019); Ng'ang'a et al. (2020); Pinder et al. (2021); and Chisanga et al. (2023).

labor supply, and household income, and ask whether screening survives a simple cost-benefit analysis. We unpack household-level effects by considering effects on adult males, females, and children separately, reflecting that household members may display different behaviors and benefit unequally from house screening. Importantly, the study investigates whether the impact of the technical intervention (house screening) can be increased when it is combined with low-cost information or behavioral interventions. Screens do not protect households if doors are left open, household members spend time outdoors after dark, or if damaged screens are not repaired. Our behavioral interventions use either a gain-framed or loss-framed health message to "crowd in" complementary health-producing efforts from households. Gain-framed messages emphasize the benefits of complying with recommended behavior and loss-framed messages focus on the cost of not-complying with recommended behavior—a subtle distinction that can be traced back to prospect theory (Tversky and Kahneman 1981). We analyze the persistence of health benefits and behavioral change over two years—speaking to whether behavioral interventions can have durable impacts "in the field".

The empirical analysis is based on a cluster randomized field experiment (RCT) in north-western Ethiopia involving 914 households in 98 sub-villages (*kebeles*). Ethiopia accounts for 1.7% of all malaria cases and 1.5% of malaria deaths globally (WHO 2022). Randomly selected sub-villages are assigned to a house screening intervention or the control group. The subsample of screened sub-villages is randomly assigned to one of three sub-arms: (i) house screening only, (ii) house screens plus a health information treatment framed as a *gain* (specifying expected health benefits after compliance with prescribed behavior), or (iii) house screens plus the same health information but now framed as a *loss* in case of non-compliance with behavioral prescriptions. In the information interventions, households are informed about the expected infection rate and morbidity reduction that follows from adopting specific behaviors. Compared to the gain-framed information treatment, the loss frame exploits the insight that loss aversion can induce behavioral change (see below). All participants in the study, in treatment and control arms, received simple bed nets that were not treated with insecticides. The control group is therefore not a pure control group, and our treatment effects are conditional on all households owning at least one bed net.[4]

Our main results indicate that house screening reduces malaria transmission. On average, across treatment arms, the total number of malaria sickness episodes for treated households per malaria season goes down by 0.70 episodes (73%) in the short term (one year after house screening) and 0.32 episodes in the long term (33%, two years after house screening). At the household level, house screening leads to 6.8 (4.4) fewer malaria sick days in the short (long) term during the main malaria season. Combining the house screening intervention with information treatments results in potentially even more promising outcomes. After one year, the screen-only intervention reduces the number of malaria episodes at the household level by 0.63 cases (66%). The results for the gain-frame are very similar (minus 0.64 events), but households in the loss frame suffered 0.86 fewer malaria episodes. Similarly, while the number of malaria sick days falls by 5.5 due to house screening alone, this number goes down by 6.2 days when combined with gain-framed health

[4] The bed net we provided was: DuraNet® (Shobikaa Impex pvt Ltd, Karur, Tamil Nadu 639006, India). We did not measure actual bed net usage.

messaging and 9.3 days when combined with a loss frame. We also document that these changes are due to changes in behavior, which appear to be most pronounced in the loss frame treatment (even if differences across information treatments are not always significantly different). The added value of combining house screens with information treatments diminishes over time, but the impact of the loss frame persists after two years. This matters for welfare and policy. Our model suggests that house screening interventions increase social welfare when combined with (loss-framed) health messaging. This is less evident for the screening-only treatment.

This paper extends the literature along multiple dimensions. First, we investigate the complementarity between a technical intervention (house screening) and information treatments (health messaging) in the field—distinguishing between gain-framed and loss-framed messaging. Our findings show that the impact of messaging, which has a near-zero cost in our case, can substantially leverage the impact of a costly house screening intervention. We provide tentative evidence that the impact varies depending on the choice architecture, which speaks to the literature on the persuasiveness of framed health messages. Second, we provide much-needed evidence on the longevity of the impact of behavioral interventions—a crucial dimension that is almost missing from the literature on loss-framed incentive contracts, which is dominated by two-hour lab experiments (e.g., Ferraro and Tracy 2022). A one-time framing intervention at the project onset had long-lasting impacts on health-producing behavior—throughout multiple malaria seasons. This suggests habit formation or learning among the target population (but we also document that the effect is attenuated slowly over time, perhaps reflecting that screens depreciate over time). Finally, this study is one of the first to examine the *economic* effects of house screening at the level of individual household members and look beyond mosquito densities or the incidence of malaria.[5]

The paper is organized as follows. In section 2, we introduce the basic theory of behavioral nudging. Section 3 presents the experimental design. Section 4 summarizes our data, outlines our identification strategy, and sketches our empirical approach—which is simple in light of the randomized design. Section 5 contains our main results for health outcomes. We present the results of models explaining variation in our measures of the prevalence of malaria and the number of sick days. In section 6 we focus on the economic outcomes of house screening, including the effects on labor supply and household income. We also present the results of a simple cost-benefit analysis based on regression results. The conclusions and a discussion ensued in Section 7.

[5] Chisanga et al. (2023) consider effects on labor supply and household income in Zambia, but their analysis suffered from extensive non-compliance. About one-third of the houses in their sample were not treated in accordance with their (random) assignment. Their local average treatment effect (LATE) therefore picks up the treatment effect for a specific and non-representative subgroup of the population. Instead, we suffer from no attrition and relatively limited non-compliance, and obtain more meaningful average treatment effects for the population of interest. Other papers that look at malaria and labor focus on alternative technologies. Fink and Masiye (2015) consider a malaria prevention technology (bed nets) among cotton farmers in Zambia and Dillon et al. (2021) study a curative technology (testing and treatment) among sugar cane cutters in Nigeria. Both studies document sizable treatment effects in terms of farmer output value, with an order of magnitude varying between 5-15%.

5.2. Behavioral insights and hypotheses

This study addresses whether the impact of a house screening intervention can be enhanced when screening is combined with an information intervention. A gain-framed health message may change behavior if the information content is new. In our context, such a message explains potential health benefits that may be obtained if household members comply with recommendations regarding closing and opening doors and windows, inspecting and maintaining screens, and time spent indoors. The loss-framed intervention contains the same informational content but is framed in terms of health benefits *foregone* if households fail to comply with these recommendations. Specifically, it aims to leverage loss aversion to induce a behavioral response. There is a large body of literature on different types of "nudging" or "choice architecture interventions" aiming to promote a variety of desirable behaviors (and which received ample attention in policy circles). The idea is that carefully designing choice environments can promote desirable behaviors without restricting people's freedom of choice.

The theory behind loss-framed messaging is based on a key tenet from behavioral science—reference-dependent preferences. Utility is assumed to not only depend on absolute consumption levels but also on changes in levels relative to a reference point. Loss-framed incentives can imply stronger incentives than gain-framed ones if the frame shifts the reference point (Tversky and Kahneman 1991). Consider a model of reference-dependent utility where utility consists of two components: standard consumption utility and so-called "gain-loss utility" (Kőszegi and Rabin 2006). We specify a simple model where an individual's utility depends on her health situation, h, and a reference point, RP, for health outcomes:

$$u(h|e,S,RP) = U(h(e,S)) + \mu(h(e,S) - RP) - c(e). \qquad (1)$$

The first term on the right-hand side captures conventional utility from health situation h, which varies with health producing effort e and a material input (say house screening), S. Behavioral insights enter through the second term, which introduces reference-dependence and captures gain-loss utility. Define a simple linear but kinked value function μ so that $\mu(x) = \eta x$ for $x \geq 0$ and $\mu(x) = \eta\lambda x$ for $x < 0$, where (i) parameter η is the (idiosyncratic) weight attached to gain-loss utility, (ii) parameter λ is the consumer's coefficient of loss aversion, and (iii) x measures the difference between the actual health situation h and reference point RP. For $\lambda>1$, utility losses associated with health outcomes h below reference value RP are greater than utility gains from equal-sized realizations in excess of that same reference value. For simplicity, we assume gain-loss utility to be linear (kinked at the reference point) but it is plausible that utility is non-linear, especially for large values of x. The third term captures the cost of health-producing effort ($c_e>0$, $c_{ee}>0$, where subscripts denote partial derivatives).

A key assumption is that information treatments can shift reference point RP and, specifically, that loss-framing can shift RP upwards, Treated households, then, believe they are endowed with a superior health status (which they can "lose" when they fail to comply with certain behavioral recommendations). This speaks to the question of how people form reference points, which is a topic of considerable interest.[6] Several papers show that the timing of conditional rewards may matter. Loss averse agents seem to work harder to retain an up-front reward that they may only keep in case of sufficient performance than to earn a bonus of the same size that they will receive after meeting the same performance criteria (e.g., Hossain and List 2012; Imas et al. 2017; Fryer et al. 2018; Bulte et al. 2020).[7] Loss-framed messaging tries to achieve the same and move reference points away from the *status quo* (initial endowment) by increasing the salience of alternative outcomes and creating the perception that the endowment has already moved. In our context, gain-framed health messaging emphasizes positive deviations from the *status quo* health status (with screening but without behavioral change), and loss-framed messaging emphasize negative deviations from an alternative benchmark (the superior health status *after* screening and adopting desirable behavior).

We solve the model in Appendix A, and develop several testable hypotheses about messaging to promote health-producing behavior, based on the assumption that a fraction of the population receiving the intervention is loss averse and that loss-framing shifts the reference point which is defined in terms of health status (and not effort). The size of aggregate effects, measured at the sample level, varies with the weight attached to gain-loss utility, the intensity of loss aversion, and the share of loss averse individuals in the sample. Motivated by the model, we derive the following hypotheses.

H1: *House screening reduces the prevalence of malaria and reduces malaria sick days for all household members*;

H2: *House screening will "crowd in" health-producing behaviors that are complements in the production of a malaria-free home environment* (h_{eS}>0);

H3: *Providing information about how health-producing effort improves health outcomes* (h_e>0) *will increase the supply of effort*;

H4: *Assuming that a fraction of the respondents are loss-averse: loss-framed health messaging crowds in more health-producing effort than gain-based health messaging, on average, and further improves health outcomes*.

The effect of gain- and loss-framed health messaging may become smaller over time, as the message loses its salience or individuals forget about it. Alternatively, the depreciation of screens over time implies that the marginal product of health-producing effort falls, which should also result in a smaller behavioral response. The rate of decline, if any, is an empirical question.

How do these hypotheses compare to the existing empirical literature? Recent meta-analyses are rather skeptical about the effectiveness of nudging in promoting behavioral change. While Mertens et al. (2021) argue that choice architecture interventions have small or medium effects on behavior, most of their findings do not survive correcting for publication bias (Maier et al. 2022). Szaszi et al. (2022) conclude that "nudge interventions may work under certain conditions but their

[6]Following conjectures of Kahneman and Tversky (1979), a literature has developed that examines the role of endowments and expectations in reference point formation (e.g., Abeler et al. 2011, Ericson and Fuster 2011, Banerji et al. 2014, Heffetz and List 2014).
[7]As mentioned, this statement is contested (Ferraro and Tracy 2022). Other studies provide much weaker or no evidence to support this theory (Levitt et al. 2016; de Quidt 2018; DellaVigna and Pope 2018).

effectiveness can vary to a great degree, and the conditions under which they work are barely identified in the literature". Ferraro and Tracy (2022) evaluate the evidence supporting productivity gains due to loss-framed contracts and find sizable differences in impact between lab studies (documenting positive impacts, but typically within a time frame of only several hours) and field experiments (on average documenting near zero impact). They conclude that the literature suffers from underpowered (lab) studies and publication bias. Our study aims to speak to these concerns; it is decently powered, organized in the field, and focuses on durable impacts—instead of fleeting ones.

Zooming in on loss-framed interventions to crowd in health-producing behavior, two meta-analytic reviews that focus on studies in high-income countries suggest that (i) gain-framed messages are more effective when encouraging behavior that is perceived as safe and unlikely to result in undesirable outcomes (e.g. applying sunscreen) while (ii) loss-framed messages are more effective when encouraging behaviors that involve some risk of an undesirable outcome (e.g. self-screening against cancer) (O'Keefe and Jensen 2007; Gallagher and Updegraff 2012). This goes back to the distinction between preventive behavior and detection behavior proposed by Rothman and Salovey (1997). Since the behaviors that are encouraged by health messaging in our experiment are safe and unlikely to involve risk for the household, the health-messaging literature suggests that the gain-framed intervention may outperform the loss-framed intervention—a perspective that stands in contrast to the predictions of prospect theory outlined above. However, context likely matters, and the literature on framed health messaging in low-income countries is in its infancy—much remains to be learned. Bekalu and Eggermont (2014) discuss health messaging in Ethiopia (about promoting HIV testing) and found that the loss frame worked better than the gain frame for rural populations with low concern about the health issue—arguably the population of interest for our interest[8].

In short, it appears as if the evidence supporting the potential for loss-framed messaging to promote health-producing behavior is mixed and contradictory or domain-specific. Whether the gain- or loss frame is more effective in promoting behavioral change in a specific context is ultimately an empirical matter. To this, we now turn.

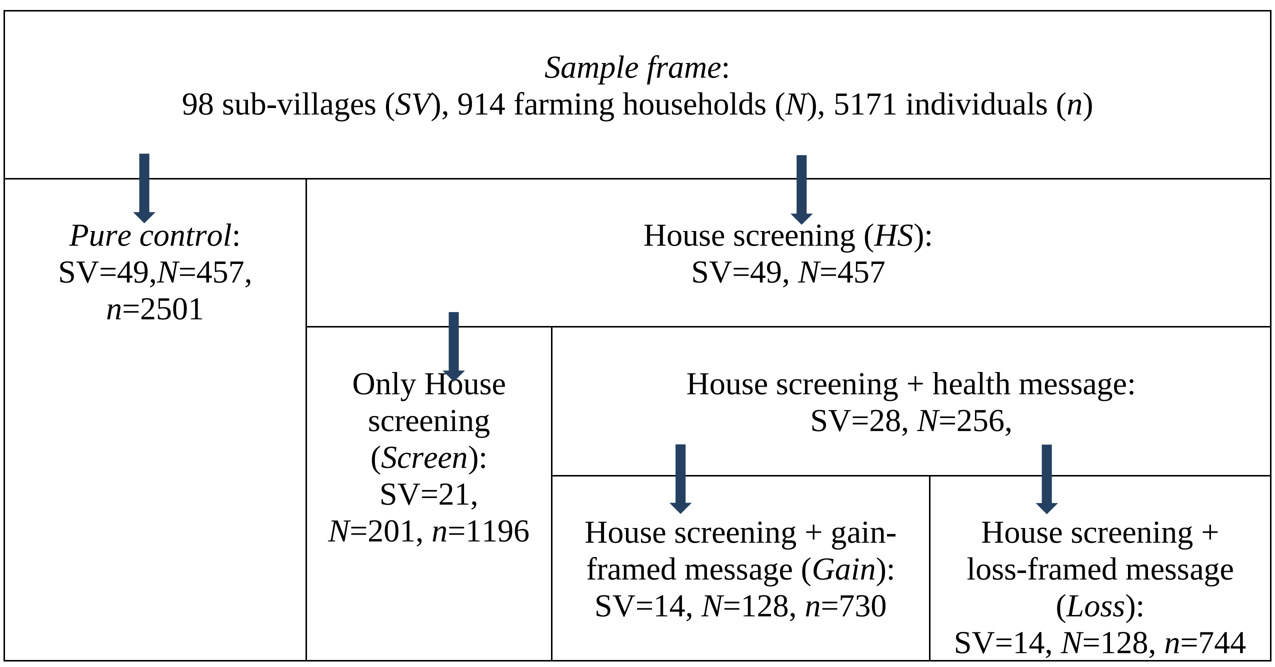
5.3. Experimental design

We started with a sample frame of 98 *kebeles* (sub-villages) from 13 *woredas* (districts), randomly drawn from a census of *woredas* in the study region characterized by a heavy malaria disease burden.[9] The local mosquito population consists mainly of *Anopheles arabiensis*—a crepuscular or nocturnal species (i.e., mostly active during the evening and night). Within each sub-village we aimed to randomly select 10 households to be included in the experiment from a census list provided by the sub-village leader (in some villages, due to logistical errors, we selected fewer households). The screening intervention was only conducted with these sampled households—we did not screen *all* houses in the sub-village (which would have been too costly).[10] All participating

[8] Instead, the gain frame outperformed the loss frame among urbanites with high concern about HIV/AIDS.
[9] A *kebele* is the smallest administrative unit in Ethiopia, similar to a neighborhood or ward. It is a subdivision of a *woreda*.
[10] Community effects are known to be important in malaria-reducing interventions, so the coverage level likely matters for outcomes. In principle, had we covered all houses in the cluster we would expect to see larger benefits. In that sense

households in the study received bed nets and a one-day training session discussing malaria symptoms and highlighting the importance of proper bed net use and seeking prompt treatment in case of symptoms. One bed net was allocated per two people in the household, so we estimate the impact of screening conditional on the household owning at least one bed net (we did not measure whether and how the net was used). The training also emphasized the importance of environmental management in reducing the availability of breeding places for mosquitos. Next, we assigned half of the villages to the control group and half to the treatment groups. Sub-villages selected for screening were further (randomly) assigned to one out of three information treatments.[11] We have 4 groups in our experiment: (i) control group houses, (ii) houses receiving screening only, (iii) houses receiving screening plus health information framed as a gain, and (iv) houses receiving screening plus health information framed as a loss.

Figure 2: Summary of experimental design



The overall design is provided in Figure 2. After screening, all groups receiving screens also received additional training and advice on how to operate their screens, the need for maintenance in case of damage, the need to keep doors and windows screens closed after dark, and the need to stay indoors after dark and sleep inside the screened house. In addition to this information, we informed households in the two information treatment arms about the expected health benefits of complying with recommended behavior. We varied the choice architecture to study whether framing matters. Experts of the International Centre of Insect Physiology and Ecology (*icipe*) estimated that under favorable conditions screening could result in a 50% reduction in the risk of contracting malaria and a 50% reduction in the number of sick days (see also Kirby et al. 2009).

our study provides an under-estimate of the potential effect of screening complete villages. In a footnote below we give more information about spillover effects of screening in our experiment.

[11] Since house screening is a relatively costly intervention, we "oversampled" control group *kebeles* and households, and did not assign sub-villages to all experimental arms in the same proportion. The number of screened households was determined by budgetary considerations. A power analysis based on the expected treatment effect of screening on malaria infection rates (by *icipe* experts) and the effect of loss-based framing on behavioral responses (by Bulte et al. 2020) suggested we could implement three treatment arms and be sufficiently powered to pick up meaningful effects.

During the training session conducted by the health extension workers, households in the gain-frame were verbally informed about this and received the following message:

"*House screening can reduce the risk of getting malaria and the number of sick days by half. These health benefits are obtained if household members behave in accordance with recommendations*".

In contrast, households in the loss-frame received the following message:

"*House screening reduces the risk of getting malaria and the number of sick days by half. However, these health benefits are not obtained if household members do not behave in accordance with recommendations*".

Three weeks after the training, all participants from the gain and loss-framed arms also received a letter, hand-delivered by an extension worker, re-iterating the information about health benefits either framed as a gain or as a loss. These letters were intended to increase the salience of the health message and as a reminder of the information shared during the training.[12]

To reduce spillover effects we assigned sub-villages, rather than individual houses, to treatment. Spillovers would occur if mosquitos denied entry into screened houses would enter control group houses instead—artificially inflating impact estimates. Half of the sub-villages were assigned to the control arm, 21 to the screening-only arm, and 28 to the framing treatments. We conducted a census of farming households in sub-villages for participation in the experiment.

House screening was implemented by local artisans who received training and instructions. A team of 25 local artisans was given a 1-day training (half-day theoretical and half-day practical) on using polyethylene products for doors and windows, and eaves screening. We also provided the building materials: polyethylene mesh, timber, and nails. Screening materials were not treated with insecticides. The training sessions were organized and facilitated by *icipe* staff and health experts from the Jabi-Tehnan district health office.

5.4. Data and identification strategy

The experiment was conducted in Jabi-Tehnan district in north-western Ethiopia. We obtained IRB approval from the Amhara Region Public Health Institute before starting the experiment, pre-registered the experiment,[13] and obtained informed consent from respondents and local officials. Data collection involved three waves of panel surveys administered via tablets. Baseline survey data were collected in May-June 2019, follow-up surveys in January-February 2020, and endline

[12] One can argue about the appropriate reference point: the *ex-ante* health level (before screening) or the *ex-post* health level (after screening) with or without appropriate health producing behavior. In the experiment, framing occurred *after* screening, so we assume the reference point is based on health levels that can be achieved with screening. We assume that, thanks to the framing intervention, the *Loss-frame* group used health status based on screens *plus* appropriate health behavior as the reference point, and the *Gain-frame* group used health status based on screens *without* appropriate health behavior as the reference point.

[13] The implementation of this study deviates in several ways from the pre-analysis plan (PAP). First, according to the PAP we would combine the house screening experiment with another RCT focusing on diffusion of an agricultural innovation (the uptake of which requires additional household labor). However, a pilot test revealed that the uptake rate of the agricultural innovation was too low to measure interaction effects between the two projects. We therefore decided to work with two separate samples and implemented the interventions separately. Second, we had to drop one primary outcome variable (number of school days lost due to malaria, for children) because the covid pandemic caused lockdowns of schools in our study region. Third, according to the plan we would have two treatment arms (screens combined with a loss frame and screens with a gain frame), but we also included a separate arm where households were only provided with screens, without any health messaging. This enabled us to separately identify the effects of framing and information provision. Fourth, the PAP is silent on the cost-benefit analysis we report below.

surveys in January-March 2022. The survey months align with the two local malaria seasons, and follow important stages in the agricultural season; the first malaria peak occurs during the period of farm preparation, cropping, and weeding, and the second peak occurs during the harvesting season. During the baseline, we collected information on household demographics, characteristics of individual members, housing conditions, and various malaria-related variables. At follow-up and endline, we again measured these malaria variables—our main dependent variables. As detailed in the pre-registration plan, we focus on the number of malaria episodes, the number of sick days due to malaria (a count variable), the number of workdays lost due to malaria, and household cash income. The study period was the past malaria season, or the period September-December for the midline and endline wave of data collection. Trained enumerators and supervisors with experience and medical knowledge conducted the data collection, ensuring clarity and accuracy by probing respondents and asking follow-up questions to distinguish malaria from other illnesses.

An F-test of joint orthogonality of baseline variables revealed that assignment to treatment (receiving screens) was random (p=0.20). Appendix Table A1 provides baseline characteristics and balance tests. Most variables are well-balanced, although some variables have small differences (we control for baseline covariates in most of the regression models that we estimate). For example, the number of adult male malaria episodes is higher in the *Loss* arm than in the *Screen* and *Pure Control* arms. Individuals who recover from malaria may develop a degree of temporary (partial) immunity that can reduce the severity of future infections. Short-term immunity lasts for only a few months (Doolan et al., 2009) so should not affect our impact estimates. Nevertheless, we check whether results for adult males are qualitatively different than for other household members.

We have two primary health outcomes: malaria prevalence and the number of malaria sick days. We operationalized prevalence as the number of sickness episodes due to malaria. At baseline, the average number of malaria events per season at the household level was 0.96, with small differences across subject categories (0.45 for adult females in the household, 0.32 for adult males, and 0.23 for children under 13 years of age). The average number of malaria sick days per season for all family members together equals 10.1. This number comprises 3.9 days for adult females, 3.5 days for adult males, and 3.4 days for children (with differences between the aggregate number and individual member types due to weighing of members based on average household composition). We are also interested in health-related behaviors. At baseline, midline and endline we collected data on the average number of hours household members spend outdoors in the evening. On average, family members spend 2.5 hours outdoors after sunset at baseline. At endline, for the subsample of households who received screens, we also collected data on the average time when screened doors were closed in the evening and opened in the morning[14] and we introduced a proxy for maintenance effort: a count variable for the number of "holes" in screens as observed by enumerators who inspected all screens. Our households' average cash income

[14] This question replaced the question about closing doors and windows at baseline, which was mainly used to establish "balance" in behavior across arms.

during the malaria season was slightly more than 8000 Ethiopian birr or approximately US$150. This reflects the poor and rural nature of our subjects, largely dependent on subsistence agriculture.

Many of our dependent variables are self-reported. The exception is screen quality, a proxy of maintenance and repair, which is observed at endline by the enumerators. Using self-reported malaria data instead of rapid diagnostic tests enables measuring malaria infections during the full peak season (as opposed to infection status at the moment of testing), which matches the time frame of our economic variables.[15] Setting up a surveillance system based on biomarkers to track infection during the malaria season would be complex and too costly. A drawback is that self-reported variables may introduce (non-classical) measurement error, including experimenter demand effects. We return to this concern in the discussion.

During the screening intervention the local screening team sometimes ran out of materials before completing screening of all houses assigned to receive screening. This resulted in two types of outcomes (see online Appendix Table A2A for details). First, 45 houses assigned to receive treatment were *not screened* at all. Appendix Table A2B, column (1), reports the outcomes of an analysis of this type of non-compliance, where we regress non-screening status on a vector of baseline variables, including house characteristics. None of the variables enters significantly. A joint significance test also fails to reject the hypothesis that our model cannot explain selection into non-compliance (p-value of F-test = 0.23).[16] Differences in the share of non-screened households between treatment arms are also not significant.

Second, 56 houses received "partial screening". These houses typically received one screened door and one screened window, but additional doors and windows went unscreened. Column (2) of Appendix Table A2B presents the results of a regression model explaining partial compliance. Partial compliance is correlated with several baseline variables, mainly related to the quality and size of the house (e.g., building materials, number of windows). This is expected, as relatively larger houses, with multiple windows (or doors), are more likely to be partially screened—small houses with only one door are either fully screened or not screened.[17] Partial compliers are found in all treatment arms, but partial-compliance is not nicely balanced: 11 houses in the loss frame, 7 in the gain frame, and 38 in the screening-only arm (see Table A2A). While the shares of partially-screened households in the two information treatments are statistically indistinguishable, the share of partially-screened households in the screening-only arm is significantly higher. Moreover, *Screen* and *Gain* have statistically different compliance rates. There is no significant difference in compliance between *Screen* and *Loss*, or between *Loss* and *Gain*.

In what follows, we retain the two types of "non-complying households" in the sample and estimate intention to treat (ITT) effects—a lower bound for the true treatment effect on the treated.

[15] Other studies also assessed malaria prevalence and incidence using self-reported survey methods (e.g., Chisanga et al. 2023; Dhewantara et al., 2019; Ipa et al., 2020; Keating et al., 2005).

[16] We collected data on outcome variables at midline and endline, and found that we cannot reject the null that outcomes for not-screened households are identical to those for households living in control sub-villages (results not shown). This suggests that spillover effects of screening are relatively unimportant at our level of screening coverage. Spillover effects will likely be larger as coverage increases. In future work, to economize on implementation and research costs, it may therefore be possible to randomize (appropriately-spaced) individual households—rather than (sub)villages—into treatment.

[17] If we compare large houses (≥ 2 windows) and small houses (<2 windows) in terms of our main dependent variables, *Malaria Prevalence* and *Number of Sick Days*, then we find that these are not statistically different.

However, differential non-compliance rates across treatment arms imply that we will overestimate the impact of information provision. Differences in ITT between the gain and loss frames on the one hand, and the screening-only frame on the other hand, capture both a treatment effect of information as well as differential compliance, because houses in the screening-only arm are more likely to be partially screened than houses in the information arms. We will argue that the latter effect is too small to explain the patterns in our data (see Discussion section).

We now turn to our empirical approach. For ease of interpretation we estimate linear models (but qualitative results are the same when we estimate non-linear models for binary dependent variables—results not shown). Random treatment assignment implies the identifying assumption is satisfied unless there are spillover effects or if there is substantial and non-random non-compliance. Spillover effects are mitigated by design (treatment at the sub-village level), but we do have some non-compliance (22% of the households assigned to screening). We cluster robust standard errors at the sub-village level and provide *p*-values in parentheses and Anderson's *q*-values to control for multiple hypotheses testing in a separate row (with the number of tests determined for each table separately).

Our main dependent variables are the number of malaria episodes, the number of malaria sickness days (health outcomes) as well as workdays lost and cash household income (economic outcomes). All these variables are measured for the previous malaria season. As mentioned, we also consider behaviors that may moderate the impact of screening. The basic ITT model we use to explore the impact of screening is as follows:

$$y_{isv} = \alpha + \beta_1 T1 + \beta_2 T2 + \beta_3 HS_{sv} + \beta_4 T1 \times HS_{sv} + \beta_5 T2 \times HS_{sv} + \gamma X_{ivs} + V_v + \varepsilon_{ivs}, \quad (2)$$

where y_{isv} represents an outcome for household i (or household member i) in sub-village s in village (*kebele*) v; $T1$ is a binary variable taking the value of one for midline values, $T2$ is a binary variable taking the value of one for endline values, HS is a binary variable indicating whether the household lived in a sub-village assigned to screening; $\boldsymbol{X}$ is a vector of individual-level, household-level or sub-village-level co-variates; $\boldsymbol{V}$ is a vector of village fixed effects capturing higher-level governance issues and agroecological conditions; and ε is an error term. The binary variables $T1$ and $T2$ pick up overall trends in malaria transmission, the impact of the bed nets we provided, and control for seasonal differences (recall the baseline was organized in May-June, and the mid- and endline took place in January-March). The variable HS, not interacted with time dummies, captures pre-existing differences between control and treatment sub-villages. Coefficients β_4 and β_5 measure causal effects of assignment to screening on our outcomes of interest, averaged across information treatments, at midline and endline respectively.

To probe the effects of information and changes in the choice architecture, we "unpack" the screening variable and estimate the following model:

$$y_{isv} = \alpha + \beta_1 T1 + \beta_2 T2 + \beta_3 Screen_{sv} + \beta_4 Gain_{sv} + \beta_5 Loss_{sv} + \beta_6 T1 \times Screen_{sv} + \beta_7 T2 \times Screen_{sv} + \beta_8 T1 \times Gain_{sv} + \beta_9 T2 \times Gain_{sv} + \beta_{10} T1 \times Loss_{sv} + \beta_{11} T2 \times Loss_{sv} + \gamma X_{ivs} + V_v + \varepsilon_{ivs}, \quad (3)$$

where $Screen_{sv}$ is a binary variable indicating whether households are from sub-villages assigned to the treatment arm receiving only screens; *Gain* is a binary variable indicating assignment to the gain frame (health information); and *Loss* is a binary variable indicating assignment to the loss-framed incentive. The remaining variables are as before. The omitted category in models (2) and (3) is the sub-sample of households in control sub-villages at baseline.

If the impact of treatments diminishes over time, because screens deteriorate or behavioral changes are not sustained, then $|\beta_6| > |\beta_7|$, $|\beta_8| > |\beta_9|$, and $|\beta_{10}| > |\beta_{11}|$. If information about the link between effort and health benefits affects behavior, then $|\beta_8| > |\beta_6|$ and $|\beta_{10}| > |\beta_6|$. If the behavioral effect is sustained, then also $|\beta_9| > |\beta_7|$ and $|\beta_{11}| > |\beta_7|$. Finally, if changes in the choice architecture affect the reference point of loss-averse subjects, so that loss framing is more effective than gain framing, then $|\beta_{10}| > |\beta_8|$. If the effect is sustained, then also $|\beta_{11}| > |\beta_9|$. Finally, as a sanity check we will verify that partially-screened houses provide less protection than fully-screened ones. If this is true, then the theoretical model predicts that the "crowding in" effect of screens on effort should be smaller for partially than fully-screened houses. We will also explore this issue empirically.

5.5. Results

5.5.1. house screening and malaria

We first consider how screening affects the prevalence of malaria. The ITT results for malaria prevalence and the number of sick days are summarized in Table 1, Panel A. Column (1) captures the impact of screening on the number of malaria episodes for all household members during the past malaria season, columns (2) and (3) capture adult females and males, respectively, and column (4) is about malaria events for children below 13 years. All regression analyses are done at the household level. We report coefficients, *p*-values (middle row) and *q*-values (bottom row). The stars indicating whether coefficients are significantly different from zero are based on the *p*-values, but it is easy to verify which results remain significant after adjustment for multiple hypotheses testing by looking at the *q*-values—overall they are rather similar.[18]

The treatment variables are mostly insignificant at baseline, indicating small pre-existing differences between control and treatment sub-villages. The time variable indicates there was no strong trend in malaria transmission in our study region (keep in mind that we provided bed nets to all households). The placement of screens affects health outcomes. The screening results are substantively and statistically significant across all subject categories. The total number of malaria episodes decreases by 0.70 cases compared to households living in unscreened houses in the short term and by 0.32 cases in the long term. This means the number of malaria episodes, on average across our treatment arms, is reduced by 73% one year after screening and by 33% two years after the intervention. The number of malaria episodes of adult females (males) is reduced by 0.31 (0.28) after one year, amounting to a risk reduction of 69% (88%). After two years, impacts fall to 0.13 events and 0.10 events, but these impacts are no longer significant. For young children, the

[18] Exceptions are two of our estimated endline coefficients for females in the loss-framed arm (*Malaria prevalence* in Table 2 and *Time spent outdoors* in Table 3). These are significant without adjusting for multiple hypotheses testing, but no longer when we make the adjustment. Other results for females remain significant after the adjustment.

number of malaria episodes decreases by 0.17 (74%) in the short and long term. While all household members benefit from screening, the coefficients slightly vary across members, which may reflect differences in protection rates due to differences in behavior. Overall, the results do not support the concern of strong reversion to the mean for adult males.

Table 1: House screening and malaria (ITT)

	Panel A: Malaria Prevalence				*Panel B: Malaria sick days*			
	All household members	All adult females	All adult males	All children < 13 years	All household members	Adult females	Adult males	Children < 13 years
Midline	-0.057 (0.654) (0.126)	-0.070 (0.353) (0.750)	0.035 (0.486) (0.510)	-0.010 (0.889) (0.450)	1.186 (0.373) (0.432)	0.174 (0.801) (0. 688)	0 .314 (0.550) (0.524)	-1.359** (0.025) (0.0597)
Endline	0.079 (0.455) (0.105)	-0.136** (0.026) (0.060)	0.057 (0.282) (0.053)	0.222** (0.010) (0.055)	-0.240 (0.983) (0.912)	-0.6418 (0.246) (0.550)	-0.094 (0.850) (0.949)	1.309** (0.044) (0.064)
HS	-0.156 (0.350) (0.167)	-0.050 (0.580) (0.189)	-0.060 (0.371) (0.670)	-0.026 (0.581) (0.470)	-1.793 (0.292) (0.269)	-0. 385 (0.397) (0.792)	-0. 369 (0.587) (0.676)	-0.910 (0.177) (0.266)
Midline × HS	-0.698*** (0.000) (0.001)	-0.312*** (0.001) (0.003)	-0.278*** (0.000) (0.001)	-0.174*** (0.002) (0.005)	-6.779*** (0.000) (0.001)	-3.155*** (0.000) (0.001)	-2.989*** (0.000) (0.002)	-1.210 (0.101) (0.173)
Endline × HS	-0. 317** (0.030) (0.016)	-0.133 (0.108) (0.058)	-0.101 (0.142) (0.077)	-0.170** (0.026) (0.024)	-4.416*** (0.004) (0.004)	-1.842** (0.015) (0.009)	-1.790*** (0.009) (0.011)	-1.671** (0.044) (0.058)
Co-variates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sub-village fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep var (control)	0.958	0.450	0.319	0.228	10.304	3.933	3.348	3.602
R-squared	0.074	0.054	0.049	0.082	0.083	0.052	0.062	0.076
Number of observations	2742	2677	2705	2174	2742	2677	2705	2174
p-value *Midline × HS=Endline×HS*	0.006	0.003	0.020	0.954	0.072	0.027	0.061	0.743

Notes: All models estimated with OLS. Baseline controls included: *household members* (number), *household head age* (years), *household head education* (1=literate), *household head sex* (1=male), *Primary activity farming* (1=yes), *Lived in the village* (years), *Mosquito nets* (number), *Rooms in the main house* (number), *Eaves in the main house* (number), *Main material of the wall of house is wood or mud* (1=yes), *Major material of the house roof is corrugated iron* (1=yes), *Shelter is separated from livestock* (1=yes). We cluster standard errors at the sub-village (sub-kebele) level. In parentheses we report *p*-values (middle row) and Anderson's *q*-values (bottom row). ***$p< 0.01$, ** $p< 0.05$ and *$p< 0.1$ indicate significance levels.

Panel B summarizes how house screening affects the number of sick days due to malaria. Overall, house screening reduces the number of sick days by 6.78 days compared to households at baseline––a 59% reduction in the short run. After two years, a reduction of 4.42 sick days (44%) remains. Effects for females, males, and young children are reported in columns (2-4). There are no significant differences in the reduction of the number of sick days between males and females ($p>0.1$), or between adults and children ($p>0.1$). At the bottom of the table, and other tables below, we report *p*-values of key tests (not adjusted for the multiplicity of null hypotheses being tested).

The estimated impact of screening on the number of sick days in our study area are larger than estimates by Chisanga et al. (2023) for Zambia. After one year of screening, they estimated that screening decreased the number of adult sick days by approximately 1.4 days per season. This is close to our point estimates of the impacts for adult females and males after 2 years. The effects we estimate after one year are larger, which may reflect differences in disease burden between the study regions. However, another important difference between the two studies is that a sizable share of the households in our Ethiopian study received the information treatment. We now turn to a more disaggregated analysis of separate treatment arms.

5.5.2. Information, the choice architecture and the impact of house screening

Our experimental design enables identification of the causal effect of information provision and changes in the choice architecture on behavioral variables that moderate the impact of screening on malaria outcomes. Before exploring how information and framing affect behavior, we analyze the reduced-form effects on malaria prevalence and the number of sick days. We estimate model (3) for these outcome variables and "unpack" average treatment effects in Table 1 by distinguishing between the effects of screens per se (*Screen*) and the effects of screens combined with health information (*Gain*) or health information provided using a loss frame (*Loss*). Regression results are summarized in Table 2. We suppress results for the time variables and treatment dummies (but these were included in the regressions).

Table 2: Information, house screening, and malaria outcomes

	Panel A: Malaria prevalence				*Panel B: Number of sick days*			
	All household members	Adult females	Adult males	Children < 13 years	All household members	Adult females	Adult males	Children < 13 years
Midline × Screen	-0.632*** (0.003) (0.003)	-0.260** (0.031) (0.021)	-0.293*** (0.002) (0.003)	-0.107* (0.073) (0.036)	-5.499*** (0.008) (0.012)	-2.694** (0.014) (0.075)	-2.389*** (0.008) (0.009)	-0.615 (0.469) (0.846)
Endline × Screen	-0.146 (0.439) (0.200)	-0.090 (0.402) (0.432)	0.010 (0.945) (0.460)	-0.102 (0.262) (0.120)	-2.582 (0.183) (0.192)	-1.267 (0.185) (0.343)	-0.915 (0.299) (0.131)	-0.753 (0.459) (0.613)
Midline × Gain	-0.643*** (0.003) (0.003)	-0.301** (0.025) (0.021)	-0.217*** (0.028) (0.010)	-0.222** (0.023) (0.024)	-6.236*** (0.009) (0.010)	-3.081** (0.012) (0.018)	-2.606** (0.037) (0.053)	-1.334 (0.223) (0.288)
Endline × Gain	-0.295 (0.111) (0.125)	-0.114 (0.278) (0.386)	-0.149* (0.083) (0.091)	-0.148 (0.218) (0.120)	-4.094** (0.036) (0.038)	-2.038** (0.041) (0.066)	-1.745* (0.077) (0.084)	-1.042 (0.274) (0.762)
Midline × Loss	-0.858*** (0.000) (0.001)	-0.412*** (0.001) (0.004)	-0. 319*** (0.000) (0.001)	-0.256*** (0.000) (0.001)	-9.347*** (0.000) (0.001)	-4.008*** (0.001) (0.004)	-4.356*** (0.000) (0.001)	-2.250** (0.013) (0.028)
Endline × Loss	-0.611*** (0.003) (0.010)	-0.220* (0.057) (0.207)	-0.224*** (0.006) (0.019)	-0.318*** (0.002) (0.004)	-6.657*** (0.002) (0.007)	-2.522** (0.022) (0.066)	-2.471*** (0.006) (0.019)	-3.209*** (0.000) (0.028)
Co-variates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sub-village fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep var (control)	0.958	0.450	0.319	0.228	10.304	3.933	3.348	3.602
R-squared	0.079	0.054	0.059	0.079	0.086	0.053	0.065	0.076
Number of observations	2742	2677	2705	2174	2742	2677	2705	2174
p-value: *Midline × Screen = Endline × Screen*	0.007	0.012	0.004	0.960	0.071	0.017	0.091	0.881
p-value: *Midline × Gain = Endline × Gain*	0.039	0.020	0.373	0.457	0.218	0.172	0.290	0.736
p-value: *Midline × Loss = Endline × Loss*	0.126	0.013	0.186	0.530	0.108	0.027	0.018	0.284
p-value: *Midline × Screen= Midline × Gain*	0.962	0.775	0.506	0.222	0. 657	0.826	0.780	0.093
p-value: *Midline × Screen= Midline × Loss*	0.294	0.242	0.798	0.017	0.057	0.315	0.071	0.013
p-value: *Midline × Gain = Midline × Loss*	0.335	0.435	0.315	0.717	0.196	0.487	0.271	0.360
p-value: *Endline × Screen = Endline × Gain*	0.495	0.844	0.135	0.801	0.336	0.432	0.463	0.380
p-value: *Endline × Screen = Endline × Loss*	0.049	0.317	0.021	0.471	0.031	0.273	0.105	0.004
p-value: *Endline × Gain = Endline × Loss*	0.174	0.413	0.407	0.437	0.222	0.727	0.427	0.071

Notes: All models are estimated with OLS. Baseline controls included: *household members* (number), *household head age* (years), *household head education* (1=literate), *household head sex* (1=male), *Primary activity farming* (1=yes), *Lived in the village* (years), *Mosquito nets* (number), *Rooms in the main house* (number), *Eves in the main house* (number), *Main material of the wall of house is wood or mud* (1=yes), *Major material of the house roof is corrugated iron* (1=yes), *Shelter is separated from livestock* (1=yes). We cluster standard errors at the sub-village level. In parentheses we report *p*-values (middle row) and Anderson's *q*-values (bottom row). ****p*< 0.01, ** *p*< 0.05 and **p*< 0.1 indicate significance levels.

For almost all models we find that house screening reduces the prevalence of malaria and the number of sick days. However, providing health information accentuates the impact of screening.

For example, consider the ITT capturing the total number of malaria episodes (columns 1 in Tables 1 and 2). The ITT was 0.70 malaria episodes in the short term and 0.32 episodes in the long term. When we consider the subsample of households assigned to receiving only screens, these ITT estimates are 0.63 and 0.15 episodes, respectively. The former ITT is statistically significant (p<0.01).

The first lesson from Table 2 is that providing information about health benefits associated with specific behaviors may increase the impact of screening. However, this is not true for the *Gain* treatment. The coefficients for the *Gain* treatment seem very similar as those for the *Screen* treatment. We cannot reject the null that the impact estimates for the screens-only treatment group and the group receiving screens and a gain-framed health message are the same (see results of Wald tests reported at the bottom of the Table). The promise of health benefits caused by specific behavioral changes induces only very small improvements in health outcomes for some household members—too small to be picked up in our experiment.

Instead, impact estimates of loss-framed messaging are larger, even if we are not always sufficiently powered to detect significant differences in our experiment. Compared to screening alone, screens plus loss-framed messaging reduce the number of malaria episodes at household-level by an additional 0.23 events at midline and an additional 0.47 events at endline. The latter difference is significant (p=0.05). Including the loss frame reduces the aggregate number of malaria sickness days by an extra 3.9 days at midline, and an extra 4.1 days at endline, compared to providing only screens. Both of these impacts are statistically significant (p=0.06 and p=0.03, respectively). It is of interest to note that for the *Loss* treatment we consistently find significant impacts after two years, for both prevalence and sick days, while none of the long-term *Screen* impacts differs from zero, and only one of the eight *Gain* impact estimates is marginally significant (p<0.1). While for all columns our coefficients can be consistently ranked (*Loss* > *Gain*) we are under-powered to pick up significant differences between *Loss* and *Gain*.[19] Loss-framed messaging appears to be a low-cost and promising way to leverage the effectiveness of house screening investments, but the evidence remains inconclusive. Nevertheless, the durability of the impact is striking and suggests the intervention triggered a sustainable behavior change (habit formation). To this, we now turn.

5.6. Isolating the mechanism: Nudging and behavior

How do information treatments accentuate the impact of screening? Our theory suggests that information and loss aversion provoke a behavioral response by utility-maximizing household members—crowding in additional health-producing efforts. Our survey instrument captures several dimensions of changes in behavior. Specifically, we asked (i) how much time household members spent "outdoors" after sunset unprotected by screens, and (ii) for the subsample of households receiving screens we asked what time these were opened and closed. Moreover, (iii) we counted the number of holes in screens at endline—an indication of the care with which screens

[19] In online Appendix Table A3 we report outcomes for a regression where we pool the information treatments—comparing *Screen-only* to *Screen + information*. Not surprisingly these are an "intermediate case" between the *Gain* and *Loss* treatments. Estimated impacts for the information treatment are statistically indistinguishable from those of the screening-only arm.

were inspected, maintained, and repaired. We estimate regression model (3) using these behavioral proxies as dependent variables. For the opening and closing of screens variables and the measure of screen damage we only have endline variables, so we estimate a simple cross-section model.

Results are summarized in Table 3, again displaying a very similar pattern in terms of the size of estimated coefficients (*Loss* > *Gain* > *Screen*). We again suppress time dummies and treatment dummies (but they were included in the models that we estimated). The first two columns demonstrate that the interventions changed the behavior of adult household members, at least at midline. The findings suggest that the relative safety of the screened home environment induced men and women to substitute part of their time outdoors for time indoors. Screening reduces the incidence of malaria by (somehow) encouraging adults to spend more time indoors. Assuming that time not spent outdoors is now spent indoors, then women spent an extra 0.49 hours (29 minutes) indoors and men an extra 0.97 hours (58 minutes) indoors, compared to the unscreened control group. Providing information magnifies this effect—particularly if information was provided in a loss-frame. The difference in time spent outdoors between the *Gain* and *Loss* treatment is statistically significant for men. Adult men from the *Loss* frame treatment spent 1.72 fewer hours outside than men from the control group—103 minutes. This is a large effect, considering that control group males, on average, spent 3.5 hours outdoors after sunset. The behavioral response of women is smaller, likely reflecting gender-specific outdoor tasks (such as cooking). However, even adult women manage to spend an additional 40 minutes indoors. The behavioral response by children is much smaller and not significantly different from zero. Many young children spend little time outdoors after dark anyway, but still get malaria in the absence of screens.[20]

It is clear that the effect of health messaging on behavior "wears off" over time. For the screen-only and gain-framed messaging treatments coefficients are not statistically different from zero. This likely provides a partial explanation for why long-term effects of screening on health outcomes in Tables 1 and 2 are worse than short-term outcomes. While the effects of loss-framed messaging also become smaller over time, they can still be picked up even two years after the intervention (for adults). On average, the extra time spent indoors falls from 40 minutes during the midline to 19 minutes at the endline for women and from 103 minutes to 40 minutes for men. Part of the behavioral change induced by loss-based framing is long-lasting.

[20] Chisanga et al. (2023) find that screening reduces malaria incidence by half as much for children compared to adults, and reduces children's time indoors, with no effect on adults. They attribute the result that children spend less time indoors (partly) to the fact that screening may impede the flow of air, causing houses to be hotter and less pleasant. Since we find that adults spend more time indoors at midline, even in the screen-only arm, it appears as if the behavioral impact of screening is context-specific and a topic worthy of additional research. For example, the sub-sample receiving screening in Chisanga et al. (2023) was a potentially-unrepresentative sub-sample of "wealthy households".

Table 3: Information, house screening, and behavior

	A: *Time spent outdoors*			B: *Close and open the screen of doors*		C: *Screen damage*
	Adult females	Adult males	Children under 13 years old	Closing time screen door (evening)	Opening time screen door(morning)	Number of holes in screen
Midline × Screen	-0.488*** (0.004) (0.005)	-0.967*** (0.000) (0.001)	0.134 (0.237) (0.527)			
Endline × Screen	-0.091 (0.585) (0.243)	-0.097 (0.724) (0.319)	-0.215 (0.352) (0.593)			
Midline × Gain	-0.562** (0.019) (0.007)	-1.500*** (0.000) (0.001)	0.181 (0.115) (0.527)			
Endline × Gain	0.421 (0.110) (0.124)	-0.397 (0.132) (0.153)	-0.186 (0.457) (0.593)	-0.298** (0.016) (0.009)	0.120 (0.288) (0.169)	-0.297* (0.098) (0.052)
Midline × Loss	-0.667*** (0.000) (0.001)	-1.721*** (0.000) (0.001)	-0.102 (0.595) (0.552)			
Endline × Loss	-0.317** (0.036) (0.122)	-0.661*** (0.003) (0.010)	-0.279 (0.124) (0.593)	-0.682*** (0.000) (0.001)	0.595*** (0.000) (0.001)	-0.473*** (0.002) (0.005)
Co-variates	Yes	Yes	Yes	Yes	Yes	Yes
Sub-village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var (screen group)	3.084	3.325	2.041	18.163	6.054	0.752
R-squared	0.193	0.198	0.120	0.261	0.172	0.540
Number of observations	2677	2705	2174	412	412	412
p-value: Midline × Screen = Endline × Screen	0.065	0.000	0.140			
p-value: Midline × Gain = Endline × Gain	0.001	0.002	0.114			
p-value: Midline × Loss = Endline × Loss	0.037	0.000	0.526			
p-value: Midline × Screen = Midline × Gain	0.759	0.084	0.714			
p-value: Midline × Screen = Midline ×Loss	0.286	0.038	0.241			
p-value: Midline × Gain = Midline × Loss	0.657	0.495	0.165			
p-value: Endline × Screen = Endline × Gain	0.070	0.363	0.925			
p-value: Endline × Screen = Endline × Loss	0.210	0.056	0.806			
p-value: Endline × Gain = Endline × Loss	0.007	0.357	0.736	0.003	0.000	0.167

Notes: All models are estimated with OLS. Baseline controls included: *household members* (number), *household head age* (years), *household head education* (1=literate), *Primary activity farming* (1=yes), *Lived in the village* (years), *Mosquito nets* (number), *Rooms in the main house*(number), *Eves in the main house*(number), *Main material of the wall of the house is wood or mud* (1=yes), *Major material of the house roof is corrugated iron* (1=yes), *Shelter is separated from livestock* (1=yes). We cluster standard errors at the sub-village (sub-kebele) level. In parentheses, we report *p*-values (middle row) and Anderson's *q*-values (bottom row). ***p< 0.01, ** p< 0.05 and *p< 0.1 indicate significance levels.

In columns (4-6) we only consider behavior related to screen usage based on endline data. The omitted category for these models is households receiving screens without information, so the sample is much smaller (households not receiving screens were dropped). Columns (4) and (5) summarize the screen door's average opening and closing time. Health messaging affects the time households open their screen—households receiving messages open their screens later and close them earlier than members of the omitted group (*Screen*). We again document the strongest behavioral response with loss-framed messaging, and now the differences between the *Gain* and *Loss* frame are statistically significant. Overall, the response seems large. Loss-framed messaging reduced the screen door's total opening time by 1.28 hours (0.682+0.595).

In column (6) we summarize regression results for a cross-section model where the dependent variable is screen quality, proxied by the number of unrepaired holes in the screen. Information treatments significantly improve screen quality, suggesting greater care and more effort to keep screens in good shape. The average number of holes equals 0.75 for *Screen* households, and falls to 0.45 holes for households in *Gain* and to 0.28 holes for *Loss.* The difference in screen quality between the loss and gain frame interventions is not statistically significant in our sample (p=0.17), reflecting the small sample size of the screen quality model (412 households only) and the associated low level of statistical power of this test. It is possible that the gradual deterioration of screens over time explains why the "leveraging impact" of the information treatments wears off over time.

In online Appendix Tables A4-6 we demonstrate that the results are robust to omitting the covariates. Our earlier findings carry over to models that only include survey and treatment dummies, and sub-village fixed effects as control variables. We also constructed a binary variable capturing whether *any* household member contracted malaria during the measurement period. The results of using this variable are very similar to the ones for our prevalence variable. They are available in online Appendix Tables A7.

Next, consider the subsample of houses that was partially-screened. Appendix Table A8 documents that partial screening attenuates the effectiveness of the intervention. We estimate different models, and document that full screening provides more protection than partial screening, and that partial screening does not outperform non-screening. Instead of looking at differences in the partial-screening rate across arms (relatively unbalanced) it is more informative to focus on differences in the full compliance rate (which is more balanced; see Table A2A).

If household members correctly believe that partial screening offers less protection than full screening, they should respond by applying less effort. This is confirmed by our data. In Appendix Table A9 we analyze how full and partial screening affect behavior. There are significant differences in terms of time spent indoors for adult males (an extra 20 minutes for men in fully-screened houses) and in terms of screen quality (0.4 fewer holes). While we do not document differences for the remaining behavioral variables, we tentatively conclude that imperfect screening impedes the "crowding in" of health-producing efforts. Perhaps not surprisingly, expectations and trust of the target group matter for the behavioral response and for final outcomes.

5.7. Economic analysis

We now turn to an economic analysis of house screening and compare the estimated gains to the cost. We focus on relatively *short-term* economic effects manifested by changes in household income. We ignore potential costs associated with behavioral change (i.e., spending more time indoors) and focus on income gains instead.[21] We use two approaches to measure income gains due to the interventions. First, we multiply the number of self-reported adult workdays gained thanks to screening by the area's average wage rate. The number of workdays gained may differ from the reduction in the number of sick days because some people go to work while being sick, and others are not working while healthy because of (malaria-driven) care responsibilities at home. Especially mothers may work fewer days when caring for sick children. Second, we directly ask households about their seasonal income from farming and off-farm sources. Chisanga et al. (2023, p7) argue that the latter income proxy may be considerably larger than the former, reflecting that "for landowning families, the return to working on their own farm *during the growing season* exceeds the average off-farm return to casual labor" (emphasis in original). Since household members typically work on-farm rather than off-farm during the growing season, the income gain estimate based on the average wage rate likely is a lower bound for the true income gain.[22]

[21]A full welfare analysis requires incorporating the utility losses due to mortality and morbidity, and curative costs associated with malaria. It also requires considering the long-term consequences due to potentially improved schooling outcomes, and so on. We lack the data to do this, so we provide a potentially gross underestimate of the overall welfare effects of screening (Sicuri et al. 2022). We collected data on school attendance (informative about long-term economic impacts), but endline data are uninformative because of school closures due to the COVID-19 pandemic. However, the midline results for schooling are fully consistent with the malaria results reported in the regression tables below: screening reduces the number of school days missed, especially when combined with the information treatments, and the coefficient for the *Loss* treatment is larger than the coefficient for the *Gain* treatment (details not shown).

[22] The relation between the number of sick days and agricultural production is complex, as there is no simple linear relationship between work days and yield. If people lose 5-6 days because of malaria when particularly important tasks for agricultural production should be carried out—such as sowing or harvesting—then the results might be a disproportionate loss in agricultural outcomes.

Table 4: House screening, information, and economic outcomes

	A: Workdays lost due to malaria			B: Log (Income)
	Adult members	Females	Males	
Midline × Screen	-5.700*** (0.007) (0.007)	-2.882*** (0.028) (0.020)	-2.837*** (0.006) (0.007)	0.235*** (0.000) (0.001)
Endline × Screen	-2.503 (0.167) (0.059)	-1.497 (0.169) (0.085)	-1.054 (0.300) (0.112)	0.152* (0.078) (0.027)
Midline × Gain	-6.184** (0.010) (0.007)	-3.437** (0.024) (0.020)	-3.203** (0.025) (0.010)	0.273*** (0.000) (0.001)
Endline × Gain	-4.869** (0.019) (0.020)	-2.411* (0.052) (0.055)	3.178*** (0.007) (0.008)	0.226*** (0.001) (0.002)
Midline * Loss	-9.81*** (0.000) (0.001)	-4.604*** (0.000) (0.001)	-5.408*** (0.000) (0.001)	0.406*** (0.000) (0.001)
Endline × Loss	-6.239*** (0.001) (0.004)	-3.444*** (0.006) (0.019)	-3.508*** (0.002) (0.007)	0.324*** (0.000) (0.001)
Co-variates	Yes	Yes	Yes	Yes
Sub-village fixed effects	Yes	Yes	Yes	Yes
Mean of dep var (control group)	9.166	5.031	4.201	8.807
R-squared	0.077	0.055	0.065	0.173
Number of observations	2742	2677	2705	2742
p-value Midline × Screen = Midline × Gain	0.852	0.731	0.807	0.305
p-value Midline× Screen = Midline × loss	0.105	0.207	0.050	0.001
p-value Midline × Gain = Midline × Loss	0.203	0.455	0.173	0.001
p-value Endline × Screen = Endline × Gain	0.307	0.496	0.106	0.403
p-value Endline× Screen = Endline × loss	0.084	0.147	0.045	0.052
p-value Endline × Gain = Endline × Loss	0.566	0.483	0.803	0.007

Notes: All models are estimated with OLS. Baseline controls included: *household members* (number), *household head age* (years), *household head education* (1=literate), *Primary activity farming* (1=yes), *Lived in the village* (years), *Mosquito nets* (number), *Rooms in the main house*(number), *Eves in the main house*(number), *Main material of the wall of the house is wood or mud* (1=yes), *Major material of the house roof is corrugated iron* (1=yes), *Shelter is separated from livestock* (1=yes). We cluster standard errors at the sub-village (sub-kebele) level. In parentheses, we report *p*-values (middle row) and Anderson's *q*-values (bottom row). ***$p<0.01$, ** $p<0.05$ and *$p<0.1$ indicate significance levels.

We first explore how screening and messaging affect labor supply and household income. Our estimates in Table 4 suggest strong results. The effect of screening on labor supply and income varies over time, and the treatment effect becomes smaller over time. In our cost-benefit analysis we use a 4-year time horizon (see below) and assume that labor supply and income as measured after the second year provide a reasonable proxy for *average* labor supply and income during the 4 year interval. We lack information about the impact of screening interventions after year 2 to examine how impacts vary over the full-time horizon of the screens, hence our results are tentative and rest on assumptions. Screen provision increases the number of adult workdays by 2.53 days after two years, relative to control group households, which is not significantly different from zero. When screening is complemented with a gain-framed health message, labor supply increases by 4.87 days ($p<0.05$). When complemented with a loss-framed health message, labor supply increases by 6.24 days ($p<0.01$)—significantly greater than the screens-only group effect ($p<0.1$). To translate these impacts into a measure of household income, we multiply additional labor supply by the daily wage. According to the Jabi-Tehnan agricultural office, the average daily district-level wage equals 250 ETB (US$4.7). Valuing labor at this price, average seasonal income gains vary from US$ 11.9 (for screening-alone) to US$ 28.4 (for screens plus loss-framed health messaging). These represent sizable effects for households in our sample—ranging from 8% to 19% of mean seasonal cash household income (US$ 150). These outcomes are in the same ballpark as estimates by Fink and Masiye (2105) for bed nets and Dillon et al. (2021) for malaria testing and treatment.

Next, turn to our survey-based measure of household cash income, in column (4). We transform household income by taking the natural logarithm to account for skewness in the income distribution. The overall treatment effect is derived by multiplying the percentage change with the control group's average income of 8070 ETB (US$150). Screening-alone increases household income after two years by 0.15×150=US$ 22.5. This income gain increases to US$ 34 when the screen is complemented with gain-framed messaging and to US$ 49 when complemented with loss-framed messaging. The latter income gain exceeds the gain due to screening-alone (p=0.05) or the gain from combining the screen with a gain-framed message (p=0.01). These numbers are lower than estimates by Chisanga et al (2023) for house screening in Zambia. Their screening-only intervention resulted in a US$ 27 income gain based on the average wage rate and a US$ 55 gain for the survey-based measure. This reflects a slightly higher wage in Zambia than in our study region (US$ 5.9 vs US$4.7), and also a different labor response.[23] In addition, the results in Chisanga et al. (2023) are based on a non-representative sample of respondents—those households living in houses that were suitable for screening (better constructed than the houses of non-compliers). If compliers are high-productivity households, as evidenced by superior housing construction, we would expect their income gains to be larger even if the impact of screening on labor supply was rather similar.

Assuming that screens offer protection for a period of four years (Kirby et al., 2009; Chisanga et al., 2023) the net present value of screens is defined as:

$$NPV_{ij} = \sum_{t=1}^{4} \frac{B-mc}{(1+r)^t} - C. \qquad (4)$$

where B is our measure of the income gain from screening during the main malaria season, mc are annual maintenance and repair costs, C is the up-front screening investment cost, and r is the discount rate. We carefully followed up on all construction costs associated with screening—material inputs, labor, and transport costs and estimate the total cost of house screening per household at US$ 49. This consists of screening materials (US$ 37.5), carpenter labor (US$ 10.3), and transportation (US$1.2). We assume annual maintenance costs are 5% of construction costs C.

[23] The labor response in Zambia was greater than in Ethiopia for the screening-only intervention—our information treatments caused increases in labor supply matching or exceeding those in Zambia.

Table 5: Internal rate of return of house screening interventions, returns per household

	Screen	*Gain*	*Loss*
Labor supply based estimate income of income gain	IRR < 0	IRR = 24%	IRR = 38%
Survey-based estimate of income gain	IRR = 23%	IRR= 52%	IRR = 87%

Table 5 provides the internal rates of return (*irr*) for the information treatments and for two different approaches to calculating income gains**.** Based on the labor-supply based estimate of the income gain, costs outweigh benefits when screening is not accompanied by a health message. Based on our survey-based measure of income gains, the outlook is brighter. We now obtain positive internal rates of return for all interventions, but obviously higher ones for the treatments with health messaging (especially when using a loss frame). Whether these rates of return will invite private investments in screening is not evident. Financial markets in rural Ethiopia are thin and interest rates charged to rural households may exceed the estimates of IRR in Table 5. Most respondents would struggle to finance the up-front cost of screening (e.g., Tarozzi et al. 2014). However, these internal rates of return suggest that publicly-funded investments in screening, accompanied by health messaging, are likely to be welfare-enhancing. This insight is accentuated when we consider that the cost-benefit analysis does not include all benefits and is based only on income gains in the main malaria season (Autumn). Malaria levels are also elevated during the Spring, so our seasonal estimates underestimate the annual income gains.

5.8. Discussion and conclusions

We use an experiment to study the health and economic impacts of house screening as a malaria control strategy and explore the scope for leveraging these impacts through information and carefully designing the choice architecture. Traditional vector control measures are increasingly failing, and alternative approaches—not based on pesticides—are currently being explored to reverse upward trends in malaria infections. Screening doors, windows and eaves is one approach to making houses mosquito-proof. However, screening requires considerable up-front investments, and the behavioral responses of members of the screened household might moderate its impact. We explore whether the short-term income gains from screening exceed the costs of screening and how the “profitability” of screening varies with complementary interventions that seek to incentivize appropriate behavior of household members. All households in this study received bed nets for free as well, so impacts are conditional on bed net ownership.

Take-up among randomly selected houses in treatment clusters was high in our intervention, reflecting that screening was provided free of charge to the household. Our results are strong and provide conditional support for recent endorsements of house screening by the World Health Organization. Consider the aggregate, household-level results. On average, house screening reduces the total number of malaria episodes during the peak season by 73% after one year and 33% after two years. The number of sick days falls by 59% after one year and 44% after two years. The impact increases when screening is complemented with informational messaging, highlighting how behavioral change can complement “technical solutions” like screening. We present tentative evidence that the choice architecture might matter in leveraging behavioral change to increase the impact of screening, in accordance with predictions of prospect theory. In particular, we measure sizable and durable behavioral changes in response to the loss frame—larger, more significant and more durable than in the case of a gain-framed message with essentially the same informational content. One explanation is that the loss-framed message shifts the reference point that households apply when evaluating changes in health outcomes after screening and crowds in complementary health-producing efforts. However, differences between the loss and gain frame are often not significant for the current sample size, so the evidence is inconclusive.

This finding complements an earlier empirical literature on the framing of (health) messaging. A recent literature concludes that the evidence supporting the claim that loss-framed messaging affects behavior is not strong (e.g. Maier et al. 2022). A literature that focuses on framing in the context of health messaging concludes that gain frames may do better than loss frames when promoting safe behavior (i.e., unlikely to result in undesirable outcomes for the respondent) (Rothman and Salovey 1997, O'Keefe and Jensen 2007; Gallagher and Updegraff 2012). Characteristics of the sample population may help to explain these divergent results (Bekalu and Eggermont 2014). It would be interesting to also study the effect of screening on behavior that is a substitute input to screens in the production of health (such as using bed nets), but we did not collect these data. It would also be interesting to repeat the screening and framing experiment in an urban setting, to examine whether the results are robust across contexts.

Four caveats are relevant. First, our analysis is mostly based on self-reported data so one may wonder whether the results are affected by experimenter demand effects. In particular, households who receive house screening at zero cost may wish to provide socially-desirable responses. We tried to reduce experimenter demand effects by communicating that the research team responsible for data collection was not the same as the funding or implementation team. However, individuals might wish to express their gratitude or to reciprocate by under-reporting sickness and inflating income gains, which would cause us to over-estimate the impact of the house screening interventions vis-à-vis the control group. Two considerations speak against the concern that our results are driven by experimenter demand effects. First, we believe the "light-touch" framing of health messages is unlikely to invite experimenter demand effects, in which case differences between treatment arms are not explained by non-classical measurement error. Ultimately, however, this is an empirical question and because we find that subtle differences in messaging lead to substantive changes in behavior it is possible that it can also change perceptions or reporting. Second, we find qualitatively similar results for the self-reported outcomes as for the objectively observed variable for screen quality (capturing maintenance and repair effort). Nevertheless, we support efforts to explore this issue further, and future research on health outcomes should preferably be based on biomarkers (but this would be much more costly).

Second, as mentioned, there is a difference in the compliance rate between the *Screen-only* arm (73%), the *Gain-framed* arm (84%) and the *Loss-framed* arm (79%). Non-compliance causes us to underestimate the impact of our treatments vis-à-vis the control group, and differences in non-compliance rates across treatment arms cause us to overestimate the effectiveness of our information treatments relative to *Screen-only* (this is because relatively more households are fully screened in the treatment arms with information). In particular, the impact gap between the *Gain-framed* arm and *Screen-only* is small, and likely evaporates when taking differential non-compliance into account. However, the difference in compliance between the *Screen-only* arm and the *Loss-framed* arm is much too small to explain the impact gap between these arms.[24] This provides further tentative evidence of the importance of the choice architecture on behavior.

Third, we referred to our framing interventions as "light-touch" and "low-cost". They are certainly low-cost for the experimenter, but the motivational impact suggests that perhaps the utility cost of being confronted with this type of messaging may be substantial for households—it might instill feelings of fear and anxiety. Studying the direct welfare effects of framing and nudging may also be of interest for future research. Fourth, while our analysis is consistent with the idea that health messaging matters and that loss-framed messaging outperforms gain-based messaging (because we find a consistent ranking of point estimates of impact: *Loss* > *Gain* > *Screen*), we also find that several differences between frames are not significant. A future study that is better powered than ours is needed to fully explore this issue. As mentioned earlier, much remains unknown about the effectiveness of nudge interventions and the conditions under which they work (Szaszi et al. 2022). This paper contributes to the evidence base of the differential effectiveness of choice architecture but cannot resolve it.

An important policy conclusion is that house screening passes a simple cost-benefit test, even if we only include short-term income effects, provided it is combined with (loss-framed) health messaging. This is true regardless of how we measure income gains—using an approach based on changes in labor supply or on changes in survey-based income. Since the cost-benefit analysis only captures short-term benefits, our estimates underestimate the true benefits. Exploring whether screening can take off as a market-mediated solution to curb malaria is worthwhile. However, this likely requires interventions in the credit market. Such public interventions may be justified if screening involves external benefits in addition to private gains, which is likely true.

[24] This can be made a bit more precise. Consider a comparison of the ITT for *Screening-only* and *Loss-frame*, for "all household members" at endline (column 1, Table 2). The ITT of the former is -15% and the ITT of the latter is -64%, so the difference in ITT between arms is -49%. This cannot reasonably be explained by a 6% difference in the compliance rate.

5.9 References

Abeler, J., Falk, A., Götte, L., and Huffman, D., 2011. Reference points and effort provision. *American Economic Review*, 101, 470–492.

Banerji, A., and Gupta, N., 2014. Detection, identification, and estimation of loss aversion: evidence from an auction experiment. *American Economic Journal: Microeconomics*, 6(1), 91–133.

Barofsky, J., Anekwe, T., Chase, C., 2015. Malaria eradication and economic outcomes: Evidence from Uganda. *Journal of Health Economics* 44, 118-136.

Bekalu, M. and Eggermont, S., 2014. The relative persuasiveness of gain-framed versus loss-framed HIV testing message: Evidence from a field experiment in northwest Ethiopia. *Journal of Health Communication* 19: 922-938

Bleakley, H., 2003. Disease and development: Evidence from the American south. *Journal of the European Economic Association*, 1(2-3), 376-386.

Bleakley, H., 2009. Economic effects of childhood exposure to tropical disease. *American Economic Review*, 99 (2): 218-23.

Bleakley, H., 2010. Malaria eradication in the Americas: A retrospective analysis of childhood exposure. *American Economic Journal: Applied Economics,* 2(2), p. 1-45.

Bloom, D., Kuhn, M. and Prettner. K., 2022.Modern infectious diseases: Macroeconomic impacts and policy responses. *Journal of Economic Literature*, 60 (1): 85-131.

Bulte, E., List, J. and van Soest, D., 2020. Toward an understanding of the welfare effects of nudges: Evidence from a field experiment in Uganda. *The Economic Journal*, 1–25.

CDC, 2024. Malaria vaccines. US Centers for Disease Control and Prevention. https://www.cdc.gov/malaria/php/public-health-strategy/malaria-vaccines.html

Chisanga, B., Bulte, E., and Kassie, M., 2023. The economic impacts of house screening against malaria transmission: Experimental evidence from Eastern Zambia. *Social Science & Medicine*, In Press

Cohen, J., Dupas, P., and Schaner, S., 2015. Price subsidies, diagnostic tests, and targeting of malaria treatment: Evidence from a randomized controlled trial. *American Economic Review* 105(2): 609-645.

Cutler, D., Fung, W., Kremer, M., Singhal, M. and Vogl, T., 2010. Early-life malaria exposure and adult outcomes: Evidence from malaria eradication in India. *American Economic Journal: Applied Economics*, 2 (2): 72-94.

DellaVigna, S. and Pope, D., 2016. What motivates effort? Evidence and expert forecasts. *The Review of Economic Studies*, 85(2): 1029-1069.

De Quidt, J., 2018. Your loss is my gain: A recruitment experiment with framed incentives. *Journal of the European Economic Association* 16(2), 522-559.

Dhewantara, P., Ipa, M., Widawati, M., 2019. Individual and contextual factors predicting self - reported malaria among adults in eastern Indonesia: Findings from an Indonesian community-based survey. *Malaria Journal*, 18(118).

Dillon, A., Friedman, J. and Serneels, P., 2021. Health information, treatment, and worker productivity. *Journal of the European Economic Association* 19(2), 1077-1115

Doolan, D. Dobano, C. and Baird, J., 2009. Acquired immunity to malaria. *Clinical Microbiology Reviews, 22*(1), 13-36.

Dupas, P., 2014. Short-term subsidies and long-term adoption of new health products: Evidence from a field experiment. *Econometrica*, 82(1), 197-228.

Ericson, K. and Fuster, A., 2011. Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments. *Quarterly Journal of Economics*, 126 (4), 1879–1907.

Ferraro, P. and Tracy, J., 2022. A reassessment of the potential for loss-framed incentive contracts to increase productivity: A meta-analysis and a real-effort experiment. *Experimental Economics*, 25: 1441-1466.

Fink, G. and Masiye, F., 2015. Health and agricultural productivity: Evidence from Zambia. *Journal of Health Economics* 42, 151-164

Fryer, R., Levitt, S., List, J. and Sadoff, S., 2018. Enhancing the efficacy of teacher incentives through framing: A field experiment. Cambridge MA: NBER Working Paper 18237.

Gallagher, K. and Updegraff, J., 2012. Health message framing effects on attitudes, intentions, and behavior: A meta-analytic review. *Annals of Behavioral Medicine, 43*(1), 101-116.

Gallup, J and Sachs, J., 2000. The economic burden of malaria. *American Journal of Tropical Medicine and Hygiene*, 64(2): 85-96.

Heffetz, O., and List, J., 2014. Is the endowment effect an expectations effect? *Journal of the European Economic Association*, 12(5), 1396–1422.

Hossain, T. and List, J. 2012. The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Science,* 58(12): 2151-2167.

Imas, A., Sadoff, S., and Samek, A., 2017. Do people anticipate loss aversion? *Management Science,* 63(5): 1271-1284.

Ipa, M., Widawati, M., Laksono, A. D., Kusrini, I., Dhewantara, W., 2020. Variation of preventive practices and its association with malaria infection in eastern Indonesia : Findings from community-based survey. *PLoS ONE*, 15(5). https://doi.org/10.1371/journal.pone.0232909.

Kahneman, D. and Tversky, A.,1979. Prospect analysis: An analysis of decision making under risk. *Econometrica,* 47(2): 263-292.

Keating, J., Macintyre, K., Mbogo, C., Githure, J. and Beier, J., 2005. Self-reported malaria and mosquito avoidance in relation to household risk factors in a Kenyan coastal city. *Journal of Biological Sciences*, 37, 761-771.

Killeen, G., Govella, N., Mlacha, Y., and Chaki, P., 2019. Suppression of malaria vector densities and human infection prevalence associated with scale-up of mosquito-proofed housing in Dar es Salaam, Tanzania: re-analysis of an observational series of parasitological and entomological surveys. *The Lancet Planetary Health*, 3(3), e132-143.

Kirby, M., Ameh, D., Bottomley, C., Green, C., Jawara, M., Milligan, P., Snell, P., Conway, D., and Lindsay, S., 2009. Effect of two different house screening interventions on exposure to malaria vectors and on anaemia in children in The Gambia: A randomised controlled trial. *The Lancet*, 374(9694), 998-1009.

Kőszegi, B. and Rabin, M., 2006. A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121 (4): 1133–65.

Levitt, S., List, J., Neckermann, S., and Sadoff, S., 2016. The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. *American Economic Journal: Economic Policy,* 8(4), 183-219.

Lindsay, S., Jawara, M., Paine, K., Pinder, M., Walraven, G., Emerson, P., 2003. Changes in house design reduce exposure to malaria mosquitoes.*Tropical Medicine & International Health,* 8(6), 512-517.

Lindsay, S., Thomas, M. and Kleinschmidt, I., 2021. Threats to the effectiveness of insecticide-treated bed nets for malaria control: Thinking beyond insecticide resistance. *The Lancet* 9, e1325-1332.

Lucas, A., 2010. Malaria eradication and educational attainment: Evidence from Paraguay and Sri Lanka. *American Economic Journal: Applied Economics*, 2 (2): 46-71.

Maier, M., Bartos, F., Stanley, T. and Wagenmakers, E., 2022. No evidence for nudging after correcting for publication bias. *Proceedings of the National Academy of Sciences* 119 (31), e2200300119

Mertens, S., Herberz, M., Hahnel U. and Brosch, T., 2021. The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences* 119 (1), e2107346118

Massebo, F., Lindtjørn, B., 2013. The effect of screening doors and windows on indoor density of Anopheles arabiensis in south-west Ethiopia: A randomized trial. *Malaria Journal*, 12(1), 1-6.

Mburu, M., Juurlink, M., Spitzen, J., Moraga, P., Hiscox, A., Mzilahowa, T., Takken, W., McCann, R., 2018. Impact of partially and fully closed eaves on house entry rates by mosquitoes. *Parasites & Vectors* 11 (383), 1-9.

Ng'ang'a, P., Okoyo, C., Mbogo, C., Mutero, C., 2020. Evaluating effectiveness of screening house eaves as a potential intervention for reducing indoor vector densities and malaria prevalence in Nyabondo, western Kenya. *Malaria Journal* 19(341). https://doi.org/10.1186/s12936-020-03413-3

O'Keefe, D. and Jensen, J., 2007. The relative persuasiveness of gain-framed and loss-framed messages for encouraging disease prevention behaviors: A meta-analytic review. *Journal of Health Communication, 12*(7), 623-644

Pinder, M., Conteh, L., Jeffries, D., Jones, C., Knudsen, J., Kandeh, B., Jawara, M., Sicuri, E., D'Alessandro, U., Lindsay, S., 2016. The RooPfs study to assess whether improved housing provides additional protection against clinical malaria over current best practice in The Gambia: Study protocol for a randomized controlled study and ancillary studies. *Trials* 17(1), 1-11.

Pinder, M., Bradley,J., Jawara, M., Affara, M., Conteh, L., Correa, S., Jeffries,D., Jones, C., Kandeh, B., Knudsen,J., Olatunji, Y., Sicuri, E., Alessandro, U., Lindsey, S., 2021. Improved housing versus usual practice for additional protection against clinical malaria in The Gambia (RooPfs): A household randomised controlled trial. *The Lancet Planetary Health*5(4), e220-e229.

Piper, K., 2021. Malaria is notoriously hard to vaccinate against. A new vaccine technology might change that. Vox. www.vox.com/future-perfect/22307700/malaria-rna-vaccines-covid-19.

Pryce, J., Richardson, M and C. Lengeler, 2018. Insecticide-treated nets for preventing malaria. Cochrane Database of Systematic Reviews, 11, Art, no.: CD000363. https://doi.org/10.1002/14651858.CD000363.pub3

Rothman, A. and Salovey, P., 1997. Shaping perceptions to motivate healthy behavior: The role of message framing. *Psychological Bulletin* 121: 1-19

Sachs, J. and Malaney, P., 2002. The economic and social burden of malaria. *Nature*, 415, 680-685

Sarma, N., Patouillard, E., Cibulskis, R. and Arcand, J-L., 2019. The economic burden of malaria: Revisiting the evidence. *American Journal of Tropical Medicine and Hygiene* 101: 1405-1415.

Sicuri, E., Vieta, A., Lindner, L., Sauboin, C., 2013. The economic costs of malaria in children in three sub-Saharan countries: Ghana, Tanzania and Kenya. *Malaria Journal*, 12(307). http://www.malariajournal.com/content/12/1/307.

Sicuri, E., Ramponi, F., Lopes-Rafegas, I. and Saute, F., 2022. A broader perspective on the economics of malaria prevention and the potential impact of SARS-CoV-2. *Nature Communications*, 13:2676.

Szaszi, B., A. Higney, A. Charlton, A. Gelman, I. Ziano, A. Aczel, D. Goldstein, D. Yeager and E. Topton, 2022. No reason to expect large and consistent effects of nudge interventions. *Proceedings of the National Academy of Sciences* 119 (31), e2200732119

Tarozzi, A., Mahajan, A., Blackburn, B., Kopf, D., Krishnan, L. and Yoong, J., 2014. Micro-loans, insecticide-treated bednets, and malaria: Evidence from a randomized controlled trial in Orissa, India. *American Economic Review*, 104 (7): 1909-41.

Tusting, L., Ippolito, M., Willey, B., Kleinschmidt, I., Dorsey, G., Gosling, R., Lindsay, S., 2015. The evidence for improving housing to reduce malaria: A systematic review and meta-analysis. *Malaria Journal* 14(1), Article 209. https://doi.org/10.1186/s12936-015-0724-1.

Tversky, A. and Kahneman, D., 1991. Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, 106(4): 1039-1061.

WHO, 2021a. World Malaria Report 2021. Geneva: World Health Organization https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021

WHO, 2021bb. WHO Guidelines for Malaria. Geneva: World Health Organization https://apps.who.int/iris/bitstream/handle/10665/343751/WHO-UCN-GMP-2021.01-Rev.1-eng.pdf?sequence=1&isAllowed=y

WHO, 2022. World Malaria Report 2022. Geneva: World Health Organization ; file:///Users/mkassie/Downloads/9789240064898-eng.pdf

Appendices

Appendix A: The simple theoretical model (Not for publication)

Consider a model of reference-dependent utility where utility consists of two components: standard consumption utility and so-called "gain-loss utility" (Kőszegi and Rabin 2006). An individual's utility depends on her health situation, h, and a reference point, RP, for health outcomes:

$$u(h|e,S,RP) = U(h(e,S)) + \mu(h(e,S) - RP) - c(e). \tag{A1}$$

The first term on the right-hand side captures conventional utility from health situation h, which varies with health producing effort e and a material input (say house screening), S. Behavioral insights enter through the second term, which introduces reference-dependence and captures gain-loss utility. Define a simple linear but kinked value function μ so that: $\mu(x) = \eta x$ for $x \geq 0$ and $\mu(x) = \eta\lambda x$ for $x < 0$, where (i) parameter η is the (idiosyncratic) weight attached to gain-loss utility, (ii) parameter λ is the consumer's coefficient of loss aversion, and (iii) x measures the difference between the actual health situation h and reference point RP. For λ>1, utility losses associated with health outcomes h below reference value RP are greater than utility gains from equal-sized realizations in excess of that same reference value. For simplicity we assume gain-loss utility to be linear (kinked at the reference point) but it is plausible that utility is non-linear, especially for large values of x. The third term captures the cost of health-producing effort (c_e>0, c_{ee}>0, where subscripts denote partial derivatives).

Two testable predictions follow immediately from the model if effort e and input S are complements in production (h_{eS}>0), which seems plausible for the behaviors we study—closing and maintaining screens, and staying indoors after dark. First, taking reservation point RP as given, the provision of screens S "crowds in" health-producing effort. If screens do not move the individual's health status from below to above reservation point RP (so that $\mu'(x)$ is constant), and $h \geq RP$, then:[25]

$$\frac{de}{dS} = \frac{-U_h h_{eS} - \eta h_{eS}}{U_h h_{ee} + \eta h_{ee} - c_{ee}} > 0. \tag{A2}$$

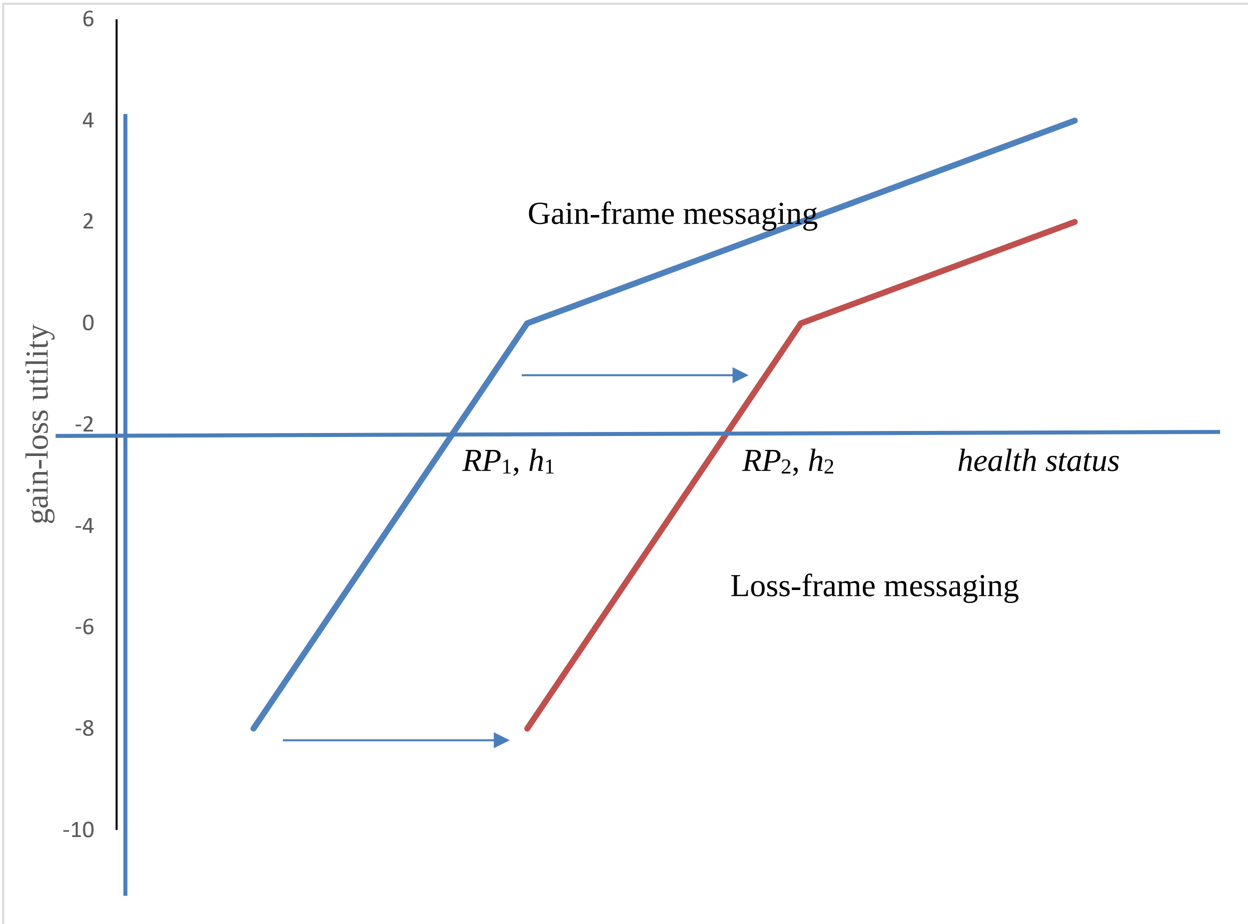
Observe that screening will "crowd out" effort if effort and screens are substitutes (h_{eS}<0). This is plausibly the case for other behaviors, such as the usage of bed nets—screens and bed nets are substitutes in the production function of bedtime protection against mosquito bites. However, we did not collect data on bed net usage, so this corollary goes untested.

The second prediction is more meaningful, and based on the idea that loss-framed messaging may shift reference point RP. Such a shift may invite an additional behavioral response. Consider the example in Figure A1. Assume an individual who is endowed with health status h_1 and treats this as her initial reference point: RP_1. After receiving a material input (screens) she is informed through a gain-framed message that supplying effort level e^* will increase her health status to h_2.

[25] Alternatively, screens may shift an individual's health status from below referenced point RP to above that same reference point. Then $\mu'(x)$ falls from $\lambda\eta$ to η. After crossing reference point RP the marginal gain-loss utility from improved health is reduced, which attenuates the extra health-producing effort that is "crowded in" by screening.

Will she voluntarily supply e=e*? Since h_2>RP_1, marginal gain-loss utility equals η. If $U_{h_2}h_{e^*} + \eta h_e \geq c_{e^*}$ then she will supply e=e* (or more). Next, assume an alternative framing approach emphasizing the potential loss of health. The same individual is informed that after receiving the material input, her health status increases to h_2, *unless she fails to supply e=e* in which case the health outcome will be lower*. If this minor change in the choice architecture implies that she now accepts RP_2 as the reference point, then marginal gain-loss utility increases to $\lambda\eta$ and the condition for supplying optimal effort e* relaxes to: $U_{h_2}h_{e^*} + \lambda\eta h_e \geq c_{e^*}$. For λ>1, she is more likely to supply at least the recommended level of effort and enjoy a better health status.

Figure A1: gain-loss utility and health messaging



Appendix B: Online tables (Not for Publication)

Table A1: Baseline characteristics and balance test

	(1)	(2)	(3)	(4)	t-test	t-test	t-test	t-test	t-test	t-test
Variables	Control	Screen	Gain	Loss	(1)-(2)	(1)-(3)	(1)-(4)	(2)-(3)	(2)-(4)	(3)-(4)
Outcome Variables										
Total number of malaria cases for all household members	0.958	1.005	0.969	1.047	-0.047	-0.010	-0.088	0.036	-0.042	-0.078
Total number of malaria cases for all adult female household members	0.450	0.443	0.496	0.595	0.007	-0.045	-0.145	-0.053	-0.152	-0.099
Total number of malaria cases for all adult male household members	0.319	0.378	0.304	0.304	-0.059	0.015	0.015	0.074	0.074	0.000
Total number of malaria cases for all children (<13)	0.228	0.213	0.297	0.270	0.016	-0.069	-0.042	-0.084	-0.057	0.027
Malaria sickness days all family members	10.30	9.31	10.61	12.33	0.99	-0.31	-2.02	-1.30	-3.02	-1.72
Malaria sickness days adult females	3.93	3.71	4.44	5.17	0.23	-0.51	-1.24	-0.74	-1.47	-0.73
Malaria sickness days adult males	3.35	3.29	4.08	4.37	0.06	-0.73	-1.02	-0.79	-1.08	-0.29
Malaria sickness days children (<13)	3.60	2.68	3.43	4.32	0.92	0.18	-0.72	-0.75	-1.64	-0.89
Days lost due to Malaria adult females	5.03	5.07	5.66	6.44	0.05	-0.55	-1.33	-0.60	-1.38	-0.78
Days lost due to Malaria adult males	4.20	4.33	5.06	5.50	-0.13	-0.86	-1.30	-0.73	-1.16	-0.43
Average time outdoor in evening females	3.08	3.23	3.01	3.17	-0.15	0.08	-0.09	0.22	0.06	-0.16
Average time outdoor in evening males	3.33	3.18	3.42	3.29	0.14*	-0.10	0.04	-0.24**	-0.11	0.14
Average time outdoor in evening children	2.04	2.12	2.17	2.05	-0.08	-0.13	-0.00	-0.05	0.08	0.13
Household yearly income (ETB)	8069	8146	8052	8039	-77.17	16.759	29.42	93.93	106.59	12.66
Control variables										
Household size (number)	5.71	5.79	5.48	5.43	-0.08	0.22	0.28*	0.31*	0.36**	0.06
Household head age (years)	43.84	42.65	45.20	45.34	1.19	-1.36	-1.50	-2.55**	-2.69**	-0.14
Marital status (1=married living with spouse)	0.99	0.99	0.98	0.98	0.00	0.01	0.01	0.01	0.00	-0.01
Household head education (1= literate)	0.54	0.62	0.44	0.56	-0.08**	0.10**	-0.02	0.18***	0.06	-0.12**
Months lived in the village (Number)	11.93	11.99	11.98	11.98	-0.06	-0.06	-0.06	0.00	0.00	0.00

Mosquito nets (number)	0.33	0.48	0.29	0.33	-0.15**	0.04	0.00	0.19**	0.15*	-0.04
Main material of wall of house is wood + mud (1=yes)	0.99	0.99	0.98	0.99	0.00	0.01	0.00	0.01	0.00	-0.01
Main material of roof of house is corrugated iron (1=yes)	0.99	0.975	0.977	0.984	0.014	0.012	0.005	-0.001	-0.009	-0.01
Rooms in main house (number)	1.62	1.57	1.56	1.63	0.05	0.06	-0.01	0.01	-0.061	-0.07
Eaves in main house (number)	2.07	2.01	2.20	2.20	0.07	-0.13*	-0.13*	-0.20**	-0.20***	0.00
Shelter is separated from livestock (1=yes)	0.58	0.64	0.59	0.50	-0.06	-0.00	0.08*	0.06	0.14**	0.09
Have knowledge of malaria (1=yes)	0.99	1.00	0.99	1.00	-0.00	0.01	-0.00	0.01	N/A	-0.01
Farm size (hectares)	1.22	1.25	1.1	1.2	-0.03	0.12	0.05	0.15*	0.07	-0.08

Note: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table A2A: Non-compliance across treatment arms

Households	*Screen-only*	*Gain-framed*	*Loss-framed*
Compliers (%)	147 (73%)	108 (84%)	101 (79%)
Non- Screened (%)	16 (8%)	13 (10%)	16 (13%)
Partially - Screened (%)	38 (19%)	7 (5%)	11 (9%)
Total	201	128	128

Table A2B: Non-compliance analysis (OLS models)

Variables	Non-compliance (1=yes)	Partially screened Households (1=yes)
Household size (number)	0.00 (0.01)	-0.01 (0.01)
Household head sex (1=male)	0.01 (0.07)	-0.02 (0.04)
Household head age (years)	0.00 (0.00)	0.00 (0.00)
Household head education (1= literate)	-0.01 (0.02)	-0.03 (0.03)
Primary activity farming (1=yes)	0.07 (0.05)	-0.40* (0.22)
Mosquito nets (number)	-0.01 (0.02)	0.00 (0.03)
Main material of the wall of the main house is wood and mud (1=yes)	0.10 (0.16)	-0.52*** (0.18)
Main material of the roof of the main house is corrugated iron (1=yes)	-0.07 (0.15)	0.09* (0.05)
Rooms in the main house (number)	-0.02 (0.02)	0.05** (0.02)
Eaves in the main house (number)	-0.00 (0.03)	0.04 (0.02)
Shelter is separated from livestock (1=yes)	0.01 (0.03)	0.02 (0.02)
Have knowledge of malaria (1=yes)	0.01 (0.04)	0.05 (0.04)
Village fixed effects	Yes	Yes
Mean dependent var	0.098	0.123
R-squared	0.095	0.392
N	457	457
F test for joint significance (p-value)	0.210	0.048

Table A3: Information, house screening, and malaria outcomes

	Panel A: Malaria prevalence				*Panel B: Number of sick days*			
	All household members	Adult females	Adult males	Children < 13 years	All household members	Adult females	Adult males	Children <13 years
Midline × Screen	-0.615*** (0.004) (0.005)	-0.253** (0.035) (0.018)	-0.282*** (0.004) (0.003)	-0.108* (0.068) (0.036)	-5.336** (0.011) (0.012)	-2.630** (0.016) (0.009)	-2.295** (0.012) (0.007)	-0.619 (0.463) (0.302)
Endline × Screen	-0.129 (0.497) (0.331)	-0.082 (0.442) (0.284)	0.017 (0.856) (0.749)	-0.102 (0.266) (0.154)	-2.419 (0.214) (0.120)	-1.194 (0.207) (0.116)	-0.819 (0.360) (0.220)	-0.742 (0.466) (0.304)
Midline × Screen + information	-0.744*** (0.003) (0.005)	-0.354*** (0.001) (0.003)	-0.267*** (0.000) (0.001)	-0.239*** (0.000) (0.001)	-7.728*** (0.009) (0.012)	-3.519*** (0.000) (0.001)	-3.481*** (0.000) (0.001)	-1.789** (0.034) (0.073)
Endline × Screen + information	-0.446*** (0.006) (0.013)	-0.165* (0.067) (0.155)	-0.183*** (0.009) (0.019)	-0.222** (0.017) (0.036)	-5.312*** (0.002) (0.005)	-2.265*** (0.007) (0.015)	-2.084*** (0.006) (0.013)	-2.249** (0.017) (0.036)
Co-variates	No	No	No	No	No	No	No	No
Sub-village fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep var (control)	0.958	0.450	0.319	0.228	10.304	3.933	3.348	3.603
R-squared	0.065	0.048	0.040	0.068	0.068	0.048	0.040	0.069
Number of observations	2742	2677	2705	2174	2742	2677	2705	2174
p-value: *Midline × Screen = Endline × Screen*	0.006	0.019	0.004	0.943	0.062	0.045	0.047	0.894
p-value: *Midline × Screen + information = Endline × Screen + information*	0.036	0.003	0.216	0.834	0.084	0.044	0.038	0.562
p-value: *Midline × Screen = Midline × Screen + information*	0.515	0.385	0.872	0.037	0. 2400	0.405	0.243	0.162
p-value: *Endline × Screen = Endline × Screen + information*	0.110	0.443	0.032	0.243	0.147	0.272	0.174	0.143

Notes: All models are estimated with OLS. We cluster standard errors at the sub-village (kebele) level. *P*-values are in reported in the middle row of each cell and Anderson's q-values in the bottom row. ***p< 0.01, ** p< 0.05 and *p< 0.1 indicate significance levels.

Table A4: Information, house screening, and malaria outcomes (without covariates)

	Panel A: Malaria prevalence				*Panel B: Number of sick days*			
	All household members	Adult females	Adult males	Children < 13 years	All household members	Adult females	Adult males	Children <13 years
Midline × Screen	-0.615*** (0.004) (0.003)	-0.253** (0.035) (0.024)	-0.282*** (0.004) (0.005)	-0.108* (0.068) (0.039)	-5.336** (0.011) (0.008)	-2.630** (0.017) (0.012)	-2.294** (0.012) (0.013)	-0.620 (0.462) (0.446)
Endline × Screen	-0.129 (0.497) (0.216)	-0.082 (0.442) (0.418)	0.017 (0.856) (0.400)	-0.102 (0.266) (0.216)	-2.419 (0.214) (0.077)	-1.194 (0.207) (0.075)	-0.819 (0.360) (0.137)	-0.741 (0.467) (0.426)
Midline × Gain	-0.638*** (0.003) (0.003)	-0.297** (0.029) (0.024)	-0.217** (0.026) (0.009)	-0.218** (0.025) (0.026)	-6.189*** (0.009) (0.008)	-3.036** (0.014) (0.012)	-2.627** (0.034) (0.019)	-1.276 (0.244) (0.323)
Endline × Gain	-0.290 (0.118) (0.134)	-0.112 (0.284) (0.397)	-0.147* (0.086) (0.095)	-0.139 (0.251) (0.216)	-4.046** (0.038) (0.040)	-2.033** (0.041) (0.066)	-1.727* (0.080) (0.087)	-1.382 (0.199) (0.249)
Midline × Loss	-0.849*** (0.000) (0.001)	-0.411*** (0.001) (0.004)	-0.315*** (0.000) (0.001)	-0.260*** (0.000) (0.001)	-9.267*** (0.000) (0.001)	-3.994*** (0.001) (0.004)	-4.315** * (0.001) (0.010)	-2.298** (0.012) (0.038)
Endline × Loss	-0.602*** (0.004) (0.013)	-0.216* (0.059) (0.216)	-0.219*** (0.007) (0.022)	-0.313*** (0.003) (0.010)	-6.557*** (0.002) (0.007)	-2.493** (0.022) (0.066)	-2.437*** (0.007) (0.022)	-3.185*** (0.005) (0.016)
Co-variates	No	No	No	No	No	No	No	No
Sub-village fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep var (control)	0.958	0.450	0.319	0.228	10.304	3.933	3.348	3.603
R-squared	0.065	0.048	0.041	0.070	0.070	0.047	0.041	0.070
Number of observations	2742	2677	2705	2174	2742	2677	2705	2174
p-value: *Midline × Screen = Endline × Screen*	0.006	0.019	0.004	0.943	0.063	0.045	0.047	0.895
p-value: *Midline × Gain = Endline × Gain*	0.038	0.016	0.349	0.437	0.193	0.175	0.255	0.914
p-value: *Midline × Loss = Endline × Loss*	0.125	0.008	0.178	0.581	0.085	0.025	0.009	0.321

p-value: *Midline* × *Screen*= *Midline* × *Gain*	0.920	0.760	0.573	0.244	0. 729	0.760	0.800	0.548
p-value: *Midline* × *Screen*= *Midline* × *Loss*	0.277	0.223	0.742	0.014	0.083	0.260	0.070	0.063
p-value: *Midline* × *Gain* = *Midline* × *Loss*	0.343	0.426	0.330	0.672	0.220	0.473	0.224	0.369
p-value: *Endline* × *Screen* = *Endline* × *Gain*	0.459	0.780	0.116	0.770	0.465	0.448	0.419	0.579
p-value: *Endline* × *Screen* = *Endline* × *Loss*	0.046	0.300	0.019	0.064	0.081	0.279	0.121	0.043
p-value: *Endline* × *Gain* = *Endline* × *Loss*	0.179	0.420	0.419	0.208	0.286	0.709	0.527	0.148

Notes: All models are estimated with OLS. Models include dummies for survey waves and treatments, but no covariates. We cluster standard errors at the sub-village (kebele) level. *P*-values are in reported in the middle row of each cell and Anderson's q-values in the bottom row. ***p< 0.01, ** p< 0.05 and *p< 0.1indicate significance levels

Table A5: Information, house screening, and behavior (no covariates)

	A: *Time spent outdoors*			B: *Close and open the screen of doors*		C: *Screen damage*
	Adult females	Adult males	Children under 13 years old	Closing time screen door (evening)	Opening time screen door (morning)	Number of holes in screen
Midline × Screen	-0.477*** (0.005) (0.006)	-0.959*** (0.000) (0.001)	0.142 (0.206) (0.370)			
Endline × Screen	-0.079 (0.639) (0.271)	-0.090 (0.743) (0.330)	-0.212 (0.359) (0.819)			
Midline × Gain	-0.557** (0.020) (0.008)	-1.497*** (0.000) (0.001)	0.189* (0.09) (0.370)			
Endline × Gain	-0.413 (0.115) (0.130)	-0.388 (0.141) (0.165)	-0.162 (0.515) (0.819)	-0.350** (0.011) (0.006)	0.074 (0.499) (0.333)	-0.288 (0.108) (0.058)
Midline × Loss	-0.667*** (0.000) (0.001)	-1.716*** (0.000) (0.001)	-0.104 (0.583) (0.448)			
Endline × Loss	-0.324** (0.031) (0.103)	-0.654*** (0.004) (0.013)	-0.259 (0.150) (0.819)	-0.677*** (0.000) (0.001)	0.588*** (0.000) (0.001)	-0.467*** (0.002) (0.005)
Co-variates	No	No	No	No	No	No
Sub-village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var (screen group)	3.084	3.325	2.041	18.163	6.054	0.752
R-squared	0.184	0.190	0.107	0.228	0.134	0.525
Number of observations	2677	2705	2174	412	412	412
p-value: Midline × Screen = Endline × Screen	0.063	0.000	0.133			
p-value: Midline × Gain = Endline × Gain	0.000	0.000	0.134			
p-value: Midline × Loss = Endline × Loss	0.041	0.000	0.576			
p-value: Midline × Screen = Midline × Gain	0.739	0.081	0.710			
p-value: Midline × Screen = Midline ×Loss	0.259	0.038	0.219			
p-value: Midline × Gain = Midline × Loss	0.642	0.500	0.146			
p-value: Endline × Screen = Endline × Gain	0.080	0.365	0.872			
p-value: Endline × Screen = Endline × Loss	0.173	0.056	0.856			

p-value: Endline × Gain = Endline × Loss	0.007	0.353	0.724	0.016	0.000	0.157

Notes: All models are estimated with OLS. Models include dummies for survey waves and treatments, but no covariates. We cluster standard errors at the sub-village (kebele) level. *P*-values are in reported in the middle row of each cell and Anderson's q-values in the bottom row. ***$p< 0.01$, ** $p< 0.05$ and *$p< 0.1$indicate significance levels

Table A6: House screening, information, and economic outcomes (no covariates)

	A: Workdays lost due to malaria			B: Log (Income)
	Adult members	Females	Males	
Midline × Screen	-5.527** (0.010) (0.008)	-2.813** (0.032) (0.020)	-2.723*** (0.009) (0.007)	0.238*** (0.000) (0.001)
Endline × Screen	-2.330 (0.199) (0.072)	-1.416 (0.189) (0.085)	-0.944 (0.361) (0.112)	0.155* (0.069) (0.027)
Midline × Gain	-6.133** (0.011) (0.008)	-3.375** (0.028) (0.020)	-3.226** (0.023) (0.010)	0.274*** (0.000) (0.001)
Endline × Gain	-4.819** (0.021) (0.022)	-2.402* (0.052) (0.055)	3.148*** (0.007) (0.008)	0.227*** (0.001) (0.002)
Midline × Loss	-9.290*** (0.000) (0.001)	-4.594*** (0.000) (0.001)	-5.370*** (0.000) (0.001)	0.407*** (0.000) (0.001)
Endline × Loss	-6.147*** (0.001) (0.004)	-3.402*** (0.006) (0.019)	-3.463*** (0.001) (0.007)	0.324*** (0.000) (0.001)
Co-variates	No	No	No	No
Sub-village fixed effects	Yes	Yes	Yes	Yes
Mean of dep var (control group)	9.166	5.031	4.201	8.807
R-squared	0.062	0.049	0.047	0.156
Number of observations	2742	2677	2705	2742
p-value Midline × Screen = Midline × Gain	0.815	0.729	0.736	0.328
p-value Midline× Screen = Midline × loss	0.098	0.192	0.043	0.001
p-value Midline × Gain = Midline × Loss	0.208	0.434	0.180	0.001
p-value Endline × Screen = Endline × Gain	0.281	0.459	0.094	0.415
p-value Endline× Screen = Endline × loss	0.079	0.138	0.041	0.054
p-value Endline × Gain = Endline × Loss	0.577	0.493	0.812	0.007

Notes: All models are estimated with OLS. Models include dummies for survey waves and treatments, but no covariates. We cluster standard errors at the sub-village (kebele) level. *P*-values are in reported in the middle row of each cell and Anderson's q-values in the bottom row. ***$p< 0.01$, ** $p< 0.05$ and *$p< 0.1$indicate significance levels

Table A7: Information, house screening and malaria (any household member sick)

	Malaria infection			
	All household members	Adult females	Adult males	Children < 13 years
Midline × Screen	-0.184*** (0.002) (0.059)	-0.135*** (0.005) (0.040)	-0.115*** (0.006) (0.040)	-0.044 (0.230) (0.038)
Endline × Screen	-0.094* (0.089) (0.055)	-0.018 (0.681) (0.040)	0.008 (0.839) (0.040)	-0.078 (0.134) (0.052)
Midline × Gain	-0.230*** (0.000) (0.051)	-0.124** (0.027) (0.055)	-0.127** (0.022) (0.055)	-0.106** (0.038) (0.050)
Endline × Gain	-0.233*** (0.000) (0.069)	-0.022 (0.601) (0.032)	-0.038 (0.243) (0.032)	-0.012 (0.713) (0.033)
Midline × Loss	-0.343*** (0.000) (0.062)	-0.224*** (0.000) (0.049)	-0.266*** (0.000) (0.049)	-0.132*** (0.005) (0.046)
Endline × Loss	-0.246*** (0.000) (0.065)	-0.104** (0.018) (0.043)	-0.0174*** (0.000) (0.046)	-0.194*** (0.000) (0.050)
Co-variates	Yes	Yes	Yes	Yes
Sub-village fixed effect	Yes	Yes	Yes	Yes
Mean dep var (control)	0.282	0.164	0.144	0.164
R-squared	0.095	0.051	0.057	0.085
Number of observations	2742	2677	2705	2174
p-value: *Midline × Screen = Endline × Screen*	0.127	0.004	0.030	0.485
p-value: *Midline × Gain = Endline × Gain*	0.961	0.089	0.397	0.916
p-value: *Midline × Loss = Endline × Loss*	0.121	0.002	0.086	0.203
p-value: *Midline × Screen= Midline × Gain*	0.484	0.860	0.652	0.027
p-value: *Midline × Screen= Midline × Loss*	0.013	0.083	0.002	0.011
p-value: *Midline × Gain = Midline × Loss*	0.135	0.090	0.048	0.700
p-value: *Endline × Screen = Endline × Gain*	0.025	0.423	0.064	0.285
p-value: *Endline × Screen = Endline × Loss*	0.037	0.066	0.000	0.011
p-value: *Endline × Gain = Endline × Loss*	0.819	0.281	0.078	0.187

Notes: Models estimated with OLS. Dependent variable: binary variable indicating whether any household member contracted malaria. Baseline controls: *household members* (number), *household head age* (years), *household head education* (1=literate), *household head sex* (1=male), *Primary activity farming* (1=yes), *Lived in the village* (years), *Mosquito nets* (number), *Rooms in the main house* (number), *Eves in the main house* (number), *Main material of the wall of house is wood or mud* (1=yes), *Major material of the house roof is corrugated iron* (1=yes), *Shelter is separated from livestock* (1=yes). Standard errors clustered at the sub-village level. *P*-values are in reported in the middle row of each cell and Anderson's q-values in the bottom row. ****p*< 0.01, ** *p*< 0.05 and **p*< 0.1 indicate significance levels.

Table A8: Full and partial screening of houses and malaria outcomes

	Fully screened and partially screened households		All households assigned to treatment (screening)	
	Number of malaria cases, all household members	*Number of malaria sick days, all household members*	*Number of malaria cases, all household members*	*Number of malaria sick days, all household members*
Full screening	-0.461*** (0.000)	-3.800** (0.019)	-0.529** (0.029)	-5.168** (0.023)
Partial screening	-	-	-0.201 (0.574)	-2.555 (0.425)
Constant	yes	yes	Yes	Yes
Co-variates	Yes	Yes	Yes	Yes
Sub-village fixed effects	Yes	Yes	Yes	Yes
R-squared	0.124	0.128	0.122	0.129
Number of observations	824	824	914	914
p-value: Full screening = partial screening	-	-	0.119	0.140

Notes: All models are estimated with OLS. The omitted category in columns (1-2) are households who were partially-screened (we dropped households assigned to screening who did not receive screens). The omitted category in columns (3-4) are households assigned to screening who were not screened. Controls included sub-kebele (village) level fixed effects, baseline covariates and baseline dependent variables (Malaria prevalence, number of malaria sick days and average time of outdoor staying): Standard errors in parentheses are clustered by sub-village (the unit of randomization). We report *p*-value. ****p*< 0.01, ** *p*< 0.05 and **p*< 0.1 indicate significance levels.

Table A9: Partial screening and behavioral outcomes

	Time spent outdoors			Close and open screen of doors		Screen damage
	Adult females	Adult males	Children under 13 years old	Average closing time of screen of the door (evening)	Average opening time of door (morning)	Number of holes in screen
Full screening	-0.234*** (0.002)	-0.591*** (0.000)	-0.056 (0.158)	-	-	-
Partial screening	-0.1800 (0.300)	-0.261* (0.065)	-0.086 (0.458)	-0.037 (0.543)	-0.147 (0.270)	0.398* (0.086)
Co-variates	Yes	Yes	Yes	Yes	Yes	Yes
Sub-village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var (control group)	3.051	3.508	1.984	20.017	6.054	0.752
R-squared	0.147	0.159	0.058	0.295	0.124	0.528
Number of observations	2677	2705	2174	412	412	412
p-value *Screen = Gain*	0.729	0.015	0.800	-	-	-

Note: All models estimated with OLS. Baseline controls included: *household members* (number), *household head age* (years), *household head education* (1=literate), *Primary activity farming* (1=yes), *Lived in the village* (years), *Mosquito nets* (number), *Rooms in the main house*(number), *Eves in the main house*(number), *Main material of the wall of the house is wood or mud* (1=yes), *Major material of the house roof is corrugated iron* (1=yes), *Shelter is separated from livestock* (1=yes). We cluster standard errors at the sub-village level. In parentheses we report *p*-values. ***$p< 0.01$, ** $p< 0.05$ and *$p< 0.1$ indicate significance levels.

Chapter 6

Synthesis

6.1 Overview

In this thesis I explore the impact of incentives and behavioral nudges on both knowledge sharing of agricultural technologies and the efficacy of health technologies in low-income rural communities. Through a combination of field experiments and analysis of real-world data, it underscores the importance of social and emotional motivators—such as recognition, reputation, and the framing of losses—in influencing behavioral change.

The primary aim is to assess how promotion efforts and informational interventions affect the uptake of push-pull technology (PPT) among smallholder farmers in Ethiopia, and how uptake affects the livelihoods of these farmers. Additionally, it examines the effectiveness of strategies designed to encourage behaviors that enhance the impact of investments in house screening for malaria prevention. A central insight emerging from this research is the pivotal role of social prestige and loss-aversion in driving these behaviors.

The subsequent sections discuss key lessons learned, policy implications, and avenues for future research, highlighting the significance of integrating behavioral insights with practical approaches to achieve sustainable and scalable development outcomes.

6.2. Lessons and insights

This thesis provides key insights into the design, implementation, and long-term impact of behavioral and technological interventions in low-income agricultural and health settings. It highlights strategic and practical considerations for policymakers and practitioners, emphasizing how motivation, social dynamics, and experiential learning can be leveraged to promote sustainable change.

6.2.1. The power of loss aversion in promoting knowledge sharing of new agricultural technologies and enhancing house screening effectiveness

Research highlights the vital role of social learning and the mechanisms through which information spreads within communities. While social networks facilitate the sharing of knowledge, active efforts are often necessary, especially when risks are involved (Munshi, 2004).

Evidence suggests that social prestige can motivate knowledge sharing as effectively as private incentives (Shikuku et al., 2018), and material rewards can also promote dissemination by enhancing motivation (BenYishay & Mobarak, 2019).

Building on this foundation, my thesis explores how framing incentives—such as social prestige and material rewards—as gains or losses influences efforts to disseminate information (Chapter 2). The findings reveal that both private rewards such as a sickle and social prestige significantly motivate lead farmers, with social prestige emerging as the more effective motivator than the sickle. Notably, loss-framed incentives that emphasize social risks, like shame, exert a stronger influence than gain framing. When combined with social prestige, loss framing further amplifies effort; however, pairing loss framing with private rewards only modestly increases activity—potentially due to limited credibility of threats. This may be attributed to the six-month delay between providing the sickle and its subsequent claw-back, which distinguishes our study from most lab-based research on claw-back effects (Chapter 2).

Although leaders increased their sharing efforts, this did not always translate into higher adoption among followers, indicating that barriers beyond mere access to information exist. Therefore, effective extension strategies could usefully incorporate social and emotional motivators—such as public recognition and loss framing—to boost participation. Nonetheless, fostering sustained behavioral change requires additional measures that promote experimentation and address other underlying barriers (Chapter 2).

Furthermore, I examined how house screening, combined with health messaging framed as losses or gains, impacts health outcomes (Chapter 5). The intervention led to a 73% reduction in malaria cases in the first year and a 33% reduction over two years, alongside decreases in sick days and improvements in household well-being. When paired with targeted messaging, loss-framed messages—highlighting risks—proved more effective in reducing malaria episodes and sick days than gain-framed messages, likely due to loss aversion. Although these effects diminished over time, the influence of loss framing persisted beyond two years.

These insights complement recent research that indicates that the impact of nudging outside controlled environments tends to be limited, with effect sizes around 0.02 in field settings and up to 0.12 in laboratory experiments (Ferraro and Tracy, 2022).

Although nudging often shows modest effects in real-world settings, my findings demonstrate that its impact "in the field" can be significantly enhanced through careful tailoring and contextual understanding. In the PPT and house screening projects, loss-framed incentives (social prestige rewards) and loss-framed health messages outperformed gain-framed approaches, emphasizing the importance of understanding the local environment and population characteristics when designing nudges—particularly in contexts with high awareness or motivation. Effective nudges rely on well-designed, targeted interventions that align with existing habits and motivations, with timing and salience playing crucial roles in their success. Incorporating incentives further reinforces sustained behavior change, while accurate measurement and high-quality data collection are essential for reliably assessing effects.

These insights suggest that combining nudges with incentives, customizing strategies to specific cultural and social contexts, and collaborating with local communities can greatly improve the relevance and longevity of behavioral interventions. Our results highlight that, when thoughtfully designed to reflect local nuances, nudging can produce meaningful and lasting change, underscoring its promising potential when implemented with cultural sensitivity.

In conclusion, combining technical solutions like house screening and PPT with behavioral incentives—particularly loss framing—can significantly improve information dissemination, health outcomes, and economic behaviors. When thoughtfully integrated, these strategies serve as powerful tools for fostering sustainable behavioral change and strengthening community resilience. The evidence highlights that leveraging emotionally salient messages and social recognition not only boosts engagement but also ensures interventions are more effective and enduring. By tailoring approaches to specific social and cultural contexts, policymakers and practitioners can maximize their initiatives' impact, ultimately leading to healthier, more resilient communities.

6.2.2. Enhancing Long-Term Adoption of Agricultural Technologies: The Critical Role of Subsidies, Experimental Incentives, and Active Engagement in Experiential Learning

While short-term subsidies can effectively incentivize the initial adoption of agricultural technologies, sustained usage over time hinges critically on how beneficiaries engage with and familiarize themselves with these innovations (Köszegi & Rabin, 2006). Experiential learning—gaining firsthand experience through active experimentation—plays a vital role in ensuring long-term benefits. Research indicates that when farmers participate in hands-on activities such as demonstrations, trial plots, and peer learning, they develop a clearer understanding of the technology's advantages, costs, and practical usability. This active involvement fosters greater confidence and a sense of ownership (Dupas, 2014; Meriggi et al., 2021; Cai et al., 2020; Carter et al., 2023).

However, simply providing full subsidies without encouraging active engagement may not be sufficient to promote initial adoption. My findings indicate that farmers who were incentivized to experiment with new technology demonstrated more effective adoption behaviors, whereas those receiving full subsidies without additional motivation tended to remain passive (Chapter 2).

I analyzed this issue in the context of a new integrated pest management approach, called push-pull technology (PPT). Adopting PPT involves leaving a portion of the maize plot for planting companion crops such as Desmodium and Brachiaria, which requires more labor during the planting year compared to conventional methods. This suggests that for technologies like PPT, which demand both land and labor, farmers need more than just subsidy; they require encouragement or incentives specifically aimed at fostering experimentation (Chapter 2).

Since adopting a new, land- and labor-intensive technology like PPT carries inherent risks—particularly if its effectiveness is uncertain—providing experimentation incentives is vital. Such incentives can motivate farmers to try the technology confidently and assess its benefits effectively (Chapter 2).

I also found that offering complete subsidies without fostering active participation can lead to anchoring effects. Farmers who receive free technology without engaging in hands-on experimentation may come to see its value as only accessible through subsidies. This perception can breed a sense of entitlement and diminish their willingness to pay once subsidies are withdrawn, thereby hindering market development. Without active involvement, farmers might also fail to fully grasp the practical benefits or limitations of the technology, leading to uncertainty and reduced motivation to invest their own resources in the future (Chapter 3).

Conversely, my findings also show that farmers who actively experiment—testing the technology, observing outcomes, and engaging with peers—are more likely to develop accurate perceptions of its value. Such experiential engagement fosters a sense of ownership, reduces uncertainty, and increases willingness to pay after subsidies end. These farmers are better positioned to integrate the technology into their routines, sustain demand, and support broader market adoption (Chapter 3).

In the case of push-pull technology (PPT), the benefits for households that experiment with it are cumulative and tend to grow over time. As farmers gain experience, refine their management practices, and as the technology matures, these advantages become more pronounced. For example, long-term adoption of PPT leads to higher livestock productivity, increased household income, and greater resilience (Chapter 4). These benefits are further amplified when farmers continuously engage with the technology, optimize its use, and seamlessly incorporate it into their routines. This increases long-term demand for PPT inputs. Without ongoing reinforcement and support, however, farmers risk reverting to old practices or discontinuing use, potentially missing out on the full economic gains PPT can offer (Chapter 4).

To promote sustained adoption, programs should incorporate activities that encourage active experimentation and learning. These include offering incentives to try new technologies, conducting field demonstrations, establishing trial plots, facilitating peer learning sessions, and organizing community engagement initiatives. Such interventions deepen farmers' understanding, mitigate anchoring effects, and foster behavioral change. Moving beyond merely providing free inputs, these strategies help beneficiaries transition from passive recipients to informed users, thereby fostering sustained demand and market-driven adoption (Chapter 3).

A carefully crafted claw-back strategy can be essential for promoting both the initial adoption and the long-term persistence of new farming practices. In my observations, some farmers who received upfront incentives for experimentation (loss-framed group) chose not to engage in experimentation and were hesitant to return the incentives when requested (Chapter 2). A well-designed claw-back mechanism can motivate farmers to adopt innovations like PPT, especially when they see the lasting benefits and durability of the technology. For example, if a technology delivers advantages over many years, farmers are more likely to continue using it even after the initial incentives end. When farmers become accustomed to new methods and see positive results, these practices tend to become ingrained, supporting sustained behavioral change. On the other hand, if incentives are viewed as temporary or if farmers are unwilling to change or they revert to traditional practices because unchanging or reverting is inexpensive and simple, initial adoption may not lead to lasting change. To maximize both adoption and sustainability, the claw-back strategy should be complemented with ongoing training and support, helping farmers understand the benefits thoroughly and encouraging independent maintenance of the new practices over time.

Overall, long-term adoption depends on active engagement and experiential learning, not just financial incentives. Farmers need firsthand experiences to build confidence and a sense of ownership. Without this, they may view the technology as only valuable when subsidized, which hampers sustained use. By combining initial support with continuous activities and strategies like claw-backs, programs can help farmers internalize the benefits and maintain their practices, ultimately leading to more resilient and productive communities.

6.3 Shortcomings and the importance of addressing them

While the thesis advocates for integrated strategies to promote behavioral change, several key shortcomings warrant attention to enhance the practical impact of these interventions.

6.3.1. Limited long-term follow-up data

Although the study demonstrates positive behavioral effects over a one-year period in Chapter 2 and a two-year period in subsequent chapters, there is limited evidence regarding the sustainability of these changes beyond these timeframes. Understanding whether behavioral improvements are maintained in the long term is essential for accurately assessing the true impact of the interventions.

Short-term gains may not lead to lasting benefits if behaviors revert once the intervention concludes or if external factors influence outcomes. Without comprehensive long-term data, policymakers and practitioners lack the confidence to justify large-scale investments or to design sustainable programs capable of producing enduring benefits. One practical approach to study this issue is to revisit the respondents in the treatment and control groups of field experiments in the future, and document their behavior. This requires additional funding.

6.3.2. Generalizability and context-specificity

The findings are primarily based on interventions conducted within specific communities in Ethiopia. This raises important questions about the extent to which these results can be generalized to other regions or cultural settings. Local norms, economic conditions, infrastructure, and social dynamics significantly influence the effectiveness of behavioral strategies. This is most likely true for interventions that leverage social prestige, but also for economic impacts of adoption of specific technologies—the livestock-related benefits studied in chapter 4 might not materialize to the same extent under different production and market conditions.

Recognizing these context-specific factors is essential to tailor policies appropriately and ensure resources are allocated efficiently. Without broader testing across diverse settings, there is a risk that strategies may be less effective or ineffective elsewhere, thereby limiting their overall impact.

6.3.3. Assessing Cost-effectiveness and Scalability

Although the thesis promotes integrated strategies, there appears to be a lack of detailed analysis regarding the costs, resource requirements, and logistical challenges involved in scaling these interventions. This gap can hinder the practical feasibility of large-scale implementation, potentially reducing their sustainability and overall effectiveness.

Evaluating cost-effectiveness helps determine whether the benefits justify the resources invested, ensuring efficient allocation of limited resources. Additionally, assessing scalability is vital to understand whether successful pilot programs can be expanded without losing their effectiveness. Without such assessments, there is a risk of adopting interventions that are too expensive or complex to sustain or replicate broadly, thereby constraining their public health impact.

6.4. Lessons for policy makers

The robust evidence presented in this thesis offers valuable insights for designing effective policies aimed at fostering agricultural innovation, improving health outcomes, and enhancing household welfare in low-income rural settings. Policymakers and development practitioners can leverage these findings to craft programs that are both impactful and sustainable. Key considerations include:

6.4.1. Promoting sustainable adoption of agricultural innovations: behavioral incentives and community engagement

Policymakers need to understand that recognizing lead farmers as community catalysts through public acknowledgment, certificates, or community awards can serve as a powerful motivator. Framing messages around potential social or reputational losses—such as shame or diminished standing—can effectively motivate leaders to organize training and outreach efforts. When combined with social prestige, these approaches can significantly amplify behavioral responses; however, attention must be paid to managing social cohesion and avoiding negative consequences like stigma.

It is important for policymakers to see the value in embedding these motivational strategies within existing community or semi-formal institutions. Leveraging trusted local structures enhances their acceptability, scalability, and sustainability by aligning with community norms and social dynamics, which are vital for lasting impact.

Understanding that simply providing information often falls short in changing farmer behavior is crucial. Incentive systems that tap into intrinsic motivators—such as social status and reputation—tend to have a more profound influence than material rewards. Social recognition and shame evoke emotional responses that are less easily manipulated, making them powerful tools; nonetheless, careful design is essential to prevent social costs like stigma or community division.

Facilitating hands-on experiences—through demonstration plots, participatory training, and trial opportunities—can deepen farmers' understanding and encourage experimentation. Recognizing that full subsidies may lead to anchoring effects and reduce willingness to pay later, policymakers should consider employing partial subsidies, performance-based incentives, and behavioral nudges to promote experimentation and belief updating, which are key to long-term adoption.

Supporting sustained adoption of technologies like push-pull systems can unlock significant economic benefits, including increased livestock productivity and household income. Therefore, it is vital for policies to include ongoing technical assistance and reinforcement to help these benefits come to fruition over time and to bolster resilience.

Given the practical difficulties of reliably delivering material rewards in resource-constrained settings, attention should be directed toward credible, sustainable incentive mechanisms. Social rewards—such as enhancing reputation or social standing—are often more feasible and impactful, provided they are designed thoughtfully to avoid excessive stigma or social friction.

Finally, adapting interventions to local social and cultural contexts is essential. For instance, in the PPT intervention, the cultural norms of the study area were considered when choosing extrinsic incentives, opting for social prestige—something highly valued by the community. Community members are eager to attain it and concerned about losing it—even considering that shameful. Similarly, sickles are economically important assets in the study region, and there is a strong desire to have them (and fear of losing them). Employing behavioral insights to craft incentive systems that effectively motivate farmers while minimizing social or welfare risks ensures better uptake. A nuanced understanding of local social dynamics and behavioral responses is key to fostering widespread adoption of innovations that promote poverty alleviation, food security, and sustainable resource management.

6.4.2. Promoting house screening coupled with loss-framed behavioral messaging

Policymakers should recognize the substantial impact of house screening in reducing malaria cases and sick days, highlighting the importance of prioritizing widespread distribution and installation in malaria-endemic regions to achieve health outcomes.

It is crucial for policymakers to understand that integrating behavioral messaging—particularly loss-framed messages emphasizing risks and potential losses—with physical interventions like house screening significantly enhances protective behaviors and supports long-lasting health benefits. Incorporating behavioral insights into public health strategies is essential for sustained success.

Adopting a comprehensive approach that combines structural measures with targeted behavioral campaigns can produce synergistic effects, leading to greater reductions in malaria incidence, sick days, and economic burdens. Policymakers should promote integrated strategies that are culturally appropriate and designed to motivate and maintain protective behaviors.

Community engagement and education are vital for fostering understanding, acceptance, and consistent use of interventions. Policymakers should invest in local outreach initiatives and establish systems for monitoring and evaluating long-term impacts, allowing for continuous refinement of strategies and reinforcement of successful practices.

It is important for policymakers to recognize that the effects of interventions may decline over time if not maintained or reinforced. Therefore, sustained efforts, periodic reinforcement of messages, and ongoing community engagement are necessary to preserve the initial benefits and prevent resurgence.

Finally, policymakers should leverage both structural and behavioral insights, understanding that a holistic, sustained approach is key to achieving lasting malaria control and improving social welfare. Implementing integrated policies that address physical and psychological determinants will be essential for building resilient, healthier communities.

6.5. Further Research

Building on the promising findings of this study, several important avenues for future research emerge.

6.5.1. Ensuring long-term sustainability of interventions

While this thesis demonstrates encouraging short- to medium-term impacts of incentive structures and behavioral strategies, understanding their durability over time is crucial. Future research should focus on whether these effects persist beyond initial engagement, considering factors such as behavioral fatigue, shifts in social norms, and the role of community reinforcement mechanisms. Longitudinal studies spanning multiple seasons or years can shed light on the sustainability of these interventions and inform strategies to maintain momentum, adapt to changing contexts, and embed behaviors into lasting social practices.

6.5.2. Evaluating transferability across contexts and technologies

The specific agricultural and health innovations examined—such as push-pull systems, house screening, and integrated pest management—may respond differently to behavioral interventions depending on local cultural, socioeconomic, and institutional landscapes. Further empirical work is needed to assess whether the positive effects of social prestige and framing observed here are generalizable to other technologies and diverse settings. Understanding contextual nuances will be vital for scaling successful strategies and ensuring their effectiveness across varied environments. This implies investments in replication across contexts and carefully designed meta-studies.

6.5.3. Unpacking the underlying behavioral and social drivers

While the effectiveness of framing and incentives is evident, a deeper comprehension of the psychological, social, and cultural mechanisms at play is essential. Future research should investigate how factors such as perceptions of risk, trust in institutions, social identity, and feedback loops influence decision-making processes. Such insights can enable the design of more targeted, culturally sensitive behavioral interventions that resonate with local motivations and beliefs, thereby enhancing their effectiveness and acceptance.

6.5.4. Assessing cost-effectiveness and scalability

For policy and programmatic adoption, it is vital to evaluate the economic viability and logistical feasibility of scaling particularly for PPT. Future studies should conduct cost-benefit analyses, explore resource requirements, and identify implementation bottlenecks. Understanding how these strategies can be integrated into existing extension services and health systems will be key to achieving large-scale, sustainable impacts in resource-constrained settings.

6.5.5. Exploring synergistic effects of push pull technology and house screening

Future research should explore the potential synergy between the push-pull system and house screening technologies to develop integrated pest and vector management strategies that maximize agricultural productivity and public health benefits. Investigating how the combined use of push-pull crops and physical house screening can simultaneously reduce pest and disease vectors—such as mosquitoes and crop pests—could lead to more sustainable and cost-effective solutions for smallholder farmers and rural communities.

Such studies should assess the compatibility, acceptance, and effectiveness of integrated interventions under diverse ecological and socio-economic contexts, as well as evaluate potential impacts on household income, health outcomes, and environmental sustainability. By understanding the complementarities and potential trade-offs of these technologies, future research can contribute to designing holistic approaches that enhance crop yields, protect households from vector-borne diseases, and promote resilience in farming systems.

6.5. References

BenYishay, A., and M. Mobarak. 2019. "Social Learning and Incentives for Experimentation and Communication." Review of Economic Studies 86(3): 976–1009.

Cai, J., A. de Janvry, and E. Sadoulet. 2020. "Subsidy Policies and Insurance Demand." American Economic Review 110(8): 2422–53.

Carter, M., R, Laajaj, and D. Yang. 2023. "Subsidies and the African Green Revolution: Direct Effects and Social Network Spillovers of Randomized Input Subsidies in Mozambique." American Economic Journal: Applied Economics13(2): 206–229

Dupas, P. 2014. "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment." Econometrica 82(1): 197–228

Ferraro, P. and J. Tracy, 2022, "A reassessment of the potential for loss-framed incentive contracts to increase productivity: a meta-analysis and a real-effort experiment." *Experimental Economics* (2022) 25:1441–1466. https ://doi.org/10.1007/s10683-022-09754-x

Koszegi, B. and M. Rabin. 2006. "A Model of Reference-Dependent Preferences." Quarterly Journal of Economics 121(4): 1133–11656.

Meriggi, N., E. Bulte, and A. M. Mobarak. 2021. "Subsidies for Technology Adoption: Experimental Evidence from Rural Cameroon." Journal of Development Economics 153: 102710. https://doi.org/10.1016/j.jdeveco.2021.102710.

Munshi, K. 2004. "Social Learning in a Heterogeneous Population: Technology Diffusion in the India Green Revolution." Journal of Development Economics 73(1): 185–213.

Shikuku, K., J. Pieters, E. Bulte, and P. Laderach. 2018. "Incentives and the Diffusion of Agricultural Knowledge: Experimental: Evidence from Northern Uganda." American Journal of Agricultural Economics 101: 1164–8

Summary

This thesis explores how incentives, behavioral nudges, and social dynamics influence the sharing of knowledge and the adoption of agricultural technologies such as push-pull pest management (PPT), as well as how framed messages affect health technology outcomes like house screening in low-income rural communities, with a particular focus on Ethiopia. Through extensive field experiments and data analysis, it underscores that emotional and social motivators—such as social prestige, recognition, and loss framing—are more influential in promoting information sharing behavior about agricultural innovations than material rewards like sickles. Additionally, the findings highlight that loss-framed health messages are more effective than gain-framed messages in enhancing the overall health outcomes of house screening.

A central insight is the effectiveness of loss aversion in fostering information dissemination and health behaviors. In the agricultural context, especially regarding PPT, framing incentives as potential social or reputational losses—like shame or diminished standing—significantly increased farmers' efforts to share knowledge among peers. Social prestige proved more motivating than material rewards, such as distributing sickles, and when loss framing was combined with social recognition, the effects were further amplified. However, achieving lasting behavioral change necessitates supplementary strategies that encourage experimentation and tackle other fundamental obstacles.

Similarly, in health interventions, loss-framed messaging—highlighting risks and potential health losses—proved more effective than gain-framed messages in reducing malaria through household screening. The intervention led to a 73% reduction in malaria cases in the first year and maintained benefits over two years, emphasizing that emotionally salient messages can reinforce health-promoting behaviors.

Nudging strategies in real-world settings often produce modest effects; however, their impact can be greatly amplified through careful adaptation to the local context. Incorporating elements such as social prestige and loss framing—aligned with local motivations—can markedly enhance both effectiveness and sustainability. Well-crafted interventions that respect and integrate social and cultural nuances are key to fostering meaningful and lasting change.

Furthermore, the findings emphasize that long-term adoption of technologies like PPT depends heavily on active engagement and experiential learning. Farmers who participate in hands-on activities—such as demonstrations, trial plots, and peer exchanges—develop greater confidence, ownership, and willingness to pay even after subsidies are withdrawn. Conversely, relying solely on full subsidies may create dependency and reduce future market demand, as farmers might perceive the technology as only accessible through continued financial support. The research demonstrates that incentives promoting experimentation, combined with ongoing reinforcement and mechanisms like claw-backs, can help farmers internalize the benefits and sustain their practices over time.

The importance of continuous engagement, technical support, and reinforcement emerges as a recurring theme. Strategies like active participation, community-based learning, and consistent messaging foster behavioral change that endures beyond initial interventions. These approaches not only improve adoption rates but also contribute to broader community resilience and economic gains, such as increased livestock productivity and household income.

Policy implications derived from the study advocate for integrating behavioral insights into development programs. Recognizing lead farmers as community catalysts through social recognition and framing messages around potential social or reputational losses can motivate engagement and dissemination efforts. Embedding these motivational strategies within trusted local institutions enhances scalability and sustainability. For health interventions, combining structural measures like house screening with loss-framed behavioral messaging and community outreach can produce synergistic effects, leading to substantial health improvements. Policymakers are encouraged to adopt holistic, culturally sensitive strategies that address both physical and psychological determinants of behavior, ensuring long-term commitment.

Despite promising results, the thesis acknowledges limitations, notably the lack of long-term follow-up data to confirm the persistence of behavioral changes. Understanding whether these effects endure over multiple seasons or years remains an open question. Additionally, the context-specific nature of the findings calls for further research across diverse settings to evaluate generalizability. Evaluating the cost-effectiveness and scalability of these interventions is also crucial—assessing resource requirements, logistical challenges, and potential for large-scale implementation will determine their practical viability. Future research should explore the potential for integrated solutions, such as combining push-pull systems with house screening, to maximize benefits across agricultural productivity and public health domains, creating holistic approaches that promote resilience and sustainability.

Overall, the thesis demonstrates that combining technological solutions with behavioral incentives—especially those leveraging loss framing and social recognition—can significantly improve information dissemination about agricultural innovations such as PPT, as well as health outcomes of health technologies like house screenings. Tailoring interventions to local social norms and promoting continuous experiential learning are crucial for long-term success. The insights offered serve as a valuable guide for policymakers and development practitioners seeking to design impactful, scalable, and sustainable programs that address poverty, food security, and health challenges in resource-constrained environments.

Acknowledgments

First and foremost, I am deeply grateful to our Heavenly Father for His grace, lovingkindness, guidance, and protection throughout this journey. All the glory be to His holy name.

This thesis would not have been possible without the support and encouragement of many individuals. I sincerely thank my supervisors—Erwin Bulte and Menale Kassie—for their unwavering support, guidance, patience, and insightful advice throughout my doctoral studies. Their critical questions and constructive feedback have been invaluable in shaping this work, and I am truly grateful for their mentorship. I extend my gratitude to the thesis assessment committee members for their time and interest in this thesis.

My gratitude goes to all the co-authors—Erwin Bulte, Menale Kassie , Zewudu Abro , Abebe Asale and Clifford M. Mutero —whose efforts contributed significantly to this thesis.

Many thanks to the faculty members within the Development Economics group, other chair groups within the Section of Economics, fellow PhD students, research participants, enumerators, government officials, and funders for their contributions and support.

Special appreciation is due to Dineke Wemmenhove for her continual support. I also thank the administrative staff at WUR, as well as my wife and brothers, for their unwavering encouragement throughout this journey.

Any errors or omissions are solely my own responsibility.

The research described in this thesis was made possible through the financial support of the Norwegian Agency for Development Cooperation (NORAD, Grant No. RAF-3058 KEN-18/0005) and the European Commission (Grant No. DCI-FOOD/2018/402-634).

Financial support from DEC group, Wageningen University for printing this thesis is gratefully acknowledged.

Cover design by: Solomon Balew

Printed by: ProefschriftMaken.nl