ISSN 2977-9669
December 2024

SPARC
Supporting Pastoralism and Agriculture in Recurrent and Protracted Crises

TECHNICAL REPORT

DO PUBLIC WORKS PROGRAMMES CREATE VALUABLE ASSETS FOR LIVELIHOODS AND RESILIENCE?

A retrospective study of the impacts of assets for natural resource management in Ethiopia and Kenya

Simon Levine, Eva Ludi and Anna McCord, with Dorice Agol, Aklilu Amsalu, Maren Duvendack, Joyce Njigua and Mulugeta Tefera

Acknowledgements

This technical report is published through the Supporting Pastoralism and Agriculture in Recurrent and Protracted Crises (SPARC) programme, which is supported by the United Kingdom's Foreign, Commonwealth and Development Office (FCDO).

This study is indebted to many people and many organisations, including: World Vision International and the World Food Programme (WFP) in Kenya, and the Government of Ethiopia who implemented the public works programmes; Carmen Leon Himmelstine for her work on the original literature review, and Neema Patel for her work in updating it; Research Guide Africa, who conducted the survey in Kenya and Samuel Gebreselassie, who conducted the survey in Ethiopia; Paulina Pankowska for her analysis of the survey data in Kenya and Ethiopia; Digital Divide Data for additional fieldwork in Kenya for the household livelihood study, and Kennedy Songola for fieldwork in dam use monitoring.

The authors also thank the following reviewers for their comments on earlier drafts: Catherine Fitzgibbon, Stephen Ngigi, Paul Harvey, Roger Calow, Stephen McDowell, Steve Wiggins and Guy Jobbins.

This report makes publicly available the findings of research undertaken with financial support from the Australian Aid Programme. The previously unpublished findings resonated with the SPARC research programme, which seeks to learn how best to promote resilience in fragile and conflict-affected places. These are precisely the situations where learning cannot simply be transferred from one place to another without understanding how change happened.

The authors are grateful to the Australian government for the funding of the research and to SPARC for its support in bringing this study to publication. Many thanks to the SPARC communications team, including Julie Grady Thomas and Zoë Windle, as well as Jo Fottrell for copyediting and Lucy Peers for design.

All opinions expressed in the report are those of the authors, and they alone are responsible for any omissions and errors.

About SPARC

Climate change, armed conflict, environmental fragility and weak governance, and the impact these have on natural resource-based livelihoods, are among the key drivers of both crisis and poverty for communities in some of the world's most vulnerable and conflict-affected countries.

SPARC aims to generate evidence and address knowledge gaps to build the resilience of millions of pastoralists, agropastoralists and farmers in these communities in sub-Saharan Africa and the Middle East.

We strive to create impact by using research and evidence to develop knowledge that improves how the FCDO, donors, non-governmental organisations, local and national governments, and civil society can empower these communities in the context of climate change.

How to cite: Levine, S., Ludi, E. and McCord, A. with Agol, D., Amsalu, A., Duvendack, M., Njigua, J. and Tefera, M. (2024) *Do public works programmes create valuable assets for livelihoods and resilience? A retrospective study of the impacts of assets for natural resource management in Ethiopia and Kenya*. Technical report. London: Supporting Pastoralism and Agriculture in Recurrent and Protracted Crises (SPARC). (https://www.sparc-knowledge.org/publications-resources/do-public-works-programmes-create-valuable-assets)

TABLE OF CONTENTS

BOXES AND FIGURES

ACRONYMS

FFA	food for assets
FGD	focus group discussion
NGO	non-governmental organisation
NRM	natural resource management
PRRO	Protracted Relief and Recovery Operation
PSNP	Productive Safety Net Programme
PWP	public works programme
SWC	soil and water conservation
WFP	World Food Programme

EXECUTIVE SUMMARY

Billions of dollars are spent annually on public works programmes (PWPs),[1] a form of social protection where people are helped to meet their immediate needs by a food or cash payment, in return for labour to build 'works' or 'assets'. But PWPs typically spend only half their budget on wages or transfers. The additional costs of building assets and managing labour contributions are justified by twin objectives: immediate needs are supported by the wage transfer, and the assets bring about longer-term benefits.

This attempt to meet short- and medium-term objectives together rests on the assumption that the assets created in PWPs will indeed improve livelihoods or resilience (or enhance well-being in other sectors, e.g., through the construction of schools or health care infrastructure). However, this assumption is rarely tested. Assets for natural resource management (NRM) are a common type of public works, but almost no rigorous studies exist that have assessed the impact of such assets on livelihoods over the medium term. This report explores these issues, finally making available the findings of fieldwork conducted in 2016 testing practical approaches to assessing the impacts of PWP assets in case studies in Ethiopia and Kenya.

The study examines assets for soil and water conservation (SWC) on hillsides in North Wollo in Ethiopia, built through the PWP component of the Productive Safety Net Programme (PSNP); and earth dams to improve water access in Makueni County, Kenya, built through a World Food Programme (WFP) food-for-assets (FFA) programme (WFP, 2010; WFP, 2012). PWP assets had been created at these locations four to five years previously. The study sites were chosen in collaboration with the implementing organisations and were selected from projects regarded as successful.

An analytical framework was created, based on causal chain models that set out how the assets could lead to livelihood improvements. A wide variety of tools and methodologies were adopted using both qualitative and quantitative approaches to test the assumptions behind each causal link. The chain first tested whether the assets were still functioning, whether they had improved natural resources, whether this had led to improved productivity and, finally, what impact this had had on livelihoods and resilience.

In both countries we found that the assets had brought no significant livelihood benefits. In Ethiopia, the SWC measures were largely well constructed and often still functioning, but the absorption of labour was prioritised in programme implementation, rather than the task of SWC. For example, far more structures had been created than could be justified on SWC grounds, and most tree seedlings died because they were planted at far too high a density and at the wrong time of year. The project did bring some environmental benefits on hilltops and some reduction in gully erosion. Claims were often heard that regenerated areas had improved livelihoods by providing fodder and products such as honey, and to have improved agricultural production on the hillsides from the reduction in soil erosion. However, our study found no evidence that these benefits had arisen, nor any plausible mechanism by which this could have occurred.

1 In the humanitarian sector, these programmes are often called cash-for-work, food-for-work or food-/cash-for-assets.

In Kenya, most of the dams were not sited, designed, constructed or maintained well enough to hold water for more than a few weeks after the rainy season, when the villages had enough water from other sources. Here too, the failings stem from the dam construction project having prioritised labour absorption over the quality of the dams.

The case studies were selected because the implementing agencies had presented the assets as successes. However, the application of rigorous research methods – for example, ensuring that testimonies were verified using multiple methods that were not subject to the same potential biases – did not substantiate these claims of success. Some claims were contradicted in blinded group discussions where the participants did not know that the researchers were interested in the livelihoods impacts of the specific PWP assets. Claims of huge increases in crop yields in Ethiopia were contradicted in interviews with local traders and found to be implausible by laboratory soil analysis. In a survey on a representative sample in villages near the dams in Kenya, almost no respondents reported ever using the PWP water sources, and yet 71% of these same respondents still ended by saying how useful the PWP project had been to them in improving their own access to water. This is a rare example of pro-project bias being quantified. The study findings make it clear that the assets' failure to provide the anticipated benefits was attributable to the inherent way that the PWPs were designed and implemented: at every stage from project conception to evaluation, labour absorption and wage payment were prioritised over the quality of the assets. This report details this failure at each stage. Projects were not designed to ensure the production of quality assets, and the impact of assets on livelihoods was not monitored in any study site, suggesting that the livelihoods objective of the PWP was not taken seriously.

Although no universal conclusions can be drawn from these two case studies on the value of assets from PWPs, they do raise serious concerns about the single focus on wage transfer that is common to PWPs globally. Of equal concern is the practice of promoting PWP success stories, even where it should be obvious that reported successes are either implausible or atypical.

The continued popularity of PWPs rests on the optimistic assumption that two birds – employment and asset generation – can indeed be killed with one stone. Reluctance to test this assumption has prevented lessons being drawn to improve the future design and management of PWPs; it risks leading to inappropriate decision-making about when PWPs should be used. Until this is corrected, billions of dollars may continue to be spent on PWPs, but value for money and contribution to resilience and climate change adaptation are likely to be disappointing.

1. INTRODUCTION

1.1 The origins of the study

This research began life as a methodological study to develop approaches to assess the impact of PWPs on people's livelihoods and resilience. Sustained impacts of projects on livelihoods can only properly be understood in the medium term, i.e., a few years after they have closed. To understand the impact of PWPs, this means going back to understand the impact of assets a few years after their construction, rather than seeking only to understand the impacts of the wages paid to those people who contributed labour.

Identifying and understanding the causality of change after several years poses challenges. When projects are initiated, baselines are rarely created: even where they do exist, they are unlikely to have measured what needs measuring, because it is rarely possible to predict exactly how people will choose to use assets. In any case, many other changes are likely to have occurred which make it difficult for research to distinguish the impact of the project from the confounding noise. This is particularly true of approaches based on survey methodologies, because they rely on identifying patterns of correlation rather than trying to understand chains of causality. The correlations that they test rely on a pre-identified theory of change, or set of assumptions, about how assets or any other intervention brings about change.

The challenge is even deeper, because it is not enough to know whether or not an intervention had impact on livelihoods. To understand how to use investments in the best ways, it is also necessary to understand how those impacts were affected by the design of the intervention(s).

The methodological study showed that it was possible to generate rigorous evidence about how assets created by public works (specifically, SWC measures and earth dams) affect people's lives. In all the cases studied, programme impacts were very much less than anticipated: the assets from public works failed to deliver the meaningful contribution to resilience that had been intended. More worryingly, this failure had not been appreciated by those connected to the projects, who had directed us to study the PWPs as examples of success. Our attention was therefore drawn to a deeper analysis of the interventions themselves, in order to understand why the results had been so disappointing.

Our research findings were presented in 2016, but, for various reasons, were not published. The intervening years have shown that the two case studies remain relevant, because little has changed regarding the issues identified. Studies going back to learn about the role of PWP assets on livelihoods and resilience remain rare (see below). Evaluations and other studies continue to focus almost exclusively on the impact of the wages paid ('the transfer'), ignoring the assets which justified the PWP project design. Little, too, has changed in how PWPs are designed and managed.

Issues that concerned us then should continue to be of concern for those promoting, funding, designing, implementing or evaluating PWPs today.

1.2 The rationale of public works programmes (PWPs)

Public works programmes (PWPs) are used to provide support to over 70 million people in low- and middle-income countries around the world every year, including in the aftermath of crises, under a variety of labels such as food- or cash-for-work, employment generation or employment guarantee schemes, etc. PWPs offer a wage (in money or in-kind) in return for labour, which is primarily used to create public goods.

Many kinds of assets can be constructed by PWPs. Projects that create public assets to support NRM (such as SWC or afforestation) are often seen as a vehicle to promote climate change adaptation, resilience and potentially even climate change mitigation (Costella and McCord, 2023; Bird et al., 2015; Béné et al., 2014). Because of the growing interest in investing in climate change adaptation, these assets are therefore of particular interest, including for SPARC.

PWPs are commonly justified with a double-promise: to give direct help through a wage or transfer to enable people to meet immediate needs, while also creating public assets that offer future productive benefits, and thus contribute to improved livelihoods and resilience.

However, the unstated challenge to this double-promise is that PWPs cost significantly more than cash transfers as a means to deliver social protection benefits, due to the administrative, technical and capital inputs required.[2] (The opportunity cost to those working on the PWP often seems to be ignored.) The rationale for this high-cost approach rests on the assumption that the assets created will bring future livelihoods benefits – and without this assumption it is hard to justify the extra costs incurred by governments and donors and by those employed to provide labour.[3]

1.3 Untested assumptions about assets

The assumption that assets will bring future benefits has rarely been assessed. Evaluations of PWPs have focused almost entirely on the immediate impact of the wage transfer and not on livelihoods (see Box 1 for a few exceptions, albeit with methodological concerns). A decade and a half ago, McCord and Farrington (2008: 1) found that 'evidence remains limited on whether the assets created by either short- or long-term PWPs help in the reduction of chronic poverty [...] More and better evidence is urgently needed'. Almost a decade later, a comprehensive review of published and grey literature found that evidence of the impact of PWP assets on livelihoods continued to be neglected (Himmelstine and McCord, 2016), and that reports of assets that did exist tended to lack rigour and thus have limited credibility. Gehrke and Hartwig (2018: 111) find that there is 'not enough evidence on productive effects of public goods generated by PW programs to justify high costs', and Beierl and Grimm (2019: 63) conclude that evidence of the impact of assets is still lacking and that 'the case for PWPs rests mainly on assumed benefits'.

---

2 McCord (2012) finds that the costs of PWPs are roughly double those of a simple monetary transfer. Depending on the type of works and the context, the total transfer value ranges from around 30% to 70% of the cost of the PWP. Social protection programmes without a work conditionality typically pay well over 90% of the total programme cost in transfers.

3 Arguments are also occasionally voiced that it is morally better that people work for a wage rather than being given money for nothing.

BOX 1: EVIDENCE OF THE IMPACTS OF PWP ASSETS ON LIVELIHOODS

A review commissioned for this report in 2023 (Patel, unpublished) to update Himmelstine and McCord (2016) finds only a few new studies on the impact of assets. Two use longitudinal datasets from PSNP in Ethiopia, but they come to different conclusions. Gazeaud and Stéphane (2020) find no evidence from a 2000–2013 dataset to suggest that public works increased agricultural productivity in beneficiary districts, whereas Filipski et al. (2017) find that SWC from PWP had improved grain yields by 2.8%. (There is no discussion of the accuracy or reliability of the survey data from which these conclusions are drawn.)

A handful of studies base their conclusions on asset impacts on the testimonies of 'beneficiaries' (for which we discuss some caveats below). WFP (2016) finds positive impacts in Uganda from a range of water sources, SWC and tree planting through PWP; Thakur (2018) finds a range of impacts from environmental and water conservation assets in India; and Anantha et al. (2021) examine the impacts of rainwater harvesting in India. Steinbach et al. (2020) find that only 11% of farmers reported any increase in crop yields from PWP in India; while Nagaraj et al. (2018) find that desilting a pond (or 'tank') doubled paddy production in India, although the methodology for quantifying the increase is not clear. Of course, different impacts are expected from different kinds of assets and from different locations.

Beyond these few studies – not all of which use a clear or rigorous methodology – no published works have been found. This means that the essential link between public works assets and livelihood outcomes, which is used to justify spending many millions of dollars annually, continues to depend on untested assumptions. It remains difficult to improve the design and implementation of PWPs as long as this almost total evidence gap exists on how public assets contribute to livelihoods, the benefits they create and for whom.

There are several methodological challenges and common methodological shortcomings to filling this evidence gap which this study addressed. These include: the lack of baseline data on livelihoods; a tendency to rely on quantitative approaches using surveys, which exclude context from the analysis; a lack of rigour when relying on focus group discussions (FGDs), where testimonies tend to be reported without attempts at verification or without engaging with the challenge of positive respondent bias (van der Mortel, 2008); and a reliance when using both quantitative and qualitative approaches on programme assumptions for the interpretation of the data collected, rather than *testing* those assumptions.

These shortcomings in conventional appraisal approaches led us to test a different approach to appraising programme impacts based on causal chain analysis, which is explained in the next section.

2. METHODOLOGY

2.1 The research approach

This study sought to overcome two common limitations in evaluation and impact assessment: the challenge of including context in the analysis whilst maintaining rigour; and the general practice of relying on assumptions rather than testing them. These limitations were overcome through the use of a clear analytical framework that allowed for the incorporation of different research methodologies; and, as explained below, through a combination of methodological flexibility with scientific scepticism and 'methodological overlap'.

The dichotomy between quantitative and qualitive research approaches remains strong. The limitations of quantitative survey-based approaches to appraise the impacts of social protection interventions were recognised over a decade ago (Devereux et al., 2013; Stern et al., 2012), but progress in addressing these critiques in relation to PWPs has been slow. This is due, at least in part, to a preference among policy-makers and donors for quantitative analysis of programme impacts.[4] Arguments have long been made for mixed-methods approaches (see, for example, Garbarino and Holland, 2009; Kanbur, 2001), but such approaches often fail to properly integrate quantitative and qualitative methodologies. Rigour is often associated only with statistical analysis of quantitative data collected using conventional survey approaches. Studies may use testimonies from qualitative enquiry to add local colour and anecdote, but they still rely on statistical correlation to prove impact. Meanwhile, analytical frameworks that can properly integrate different methods are rarely used: they can only be adopted successfully if qualitative (i.e., non-statistical) enquiry is as rigorous as quantitative approaches.

In both methodological approaches, conclusions have too often been drawn from pre-identified indicators which are used as evidence of the intended impacts. The identification of such indicators relies, though, on assuming the validity of the theory of change which underpinned the intervention in the first place.

To address these challenges, we adopted a *flexible and exploratory approach* that avoided a pre-planned set of methodological tools. Instead, as described in the sections below, emerging findings opened up new investigatory directions, for which methods had to be devised (qualitative or quantitative), depending on the nature of the evidence needed.

Rigour was built in through an attitude of *scientific scepticism*. Testimony and evidence that appeared to confirm positive assumptions were treated as a possible manifestation of pro-project bias until we could rule out all explanations of change other than the project intervention, i.e., the PWP assets.[5] All testimony and evidence had to show a plausible logic

4 See Cartwright (2007) for a discussion of the 'vanity of rigour'. Rigour in the statistical treatment of data does not lead to rigorous findings if there are doubts about the quality of the data collected (e.g., asking households about their annual income) or in the interpretation of what correlations actually signify. Numbers given to two decimal places somehow continue to be treated with respect and to be equated with objectivity and rigour.

5 See Normand (2008: 42) for an explanation of scientific scepticism, where the author captures our approach well: '… we should accept no things as true unless the evidence available makes the non-existence of the thing more miraculous than its existence. Even extraordinary claims can be true, but the more extraordinary the claim, the more extraordinary the evidence required.' Testimonies that crop yields increased three times following a project intervention are an example of a claim requiring quite extraordinary evidence.

and to have corroborating evidence which came from sources that could not share the same bias (for example, two people offering similar stories in a FGD did not constitute corroboration or triangulation, unless they used very different explanations or brought up different kinds of substantiating evidence). This attitude was intended to avoid pro-project bias both from the largely conscious and deliberate pro-project bias of informants (both beneficiaries and implementors) and the often less conscious bias by which evaluators or researchers favour evidence that confirms what they expect or hope to find.

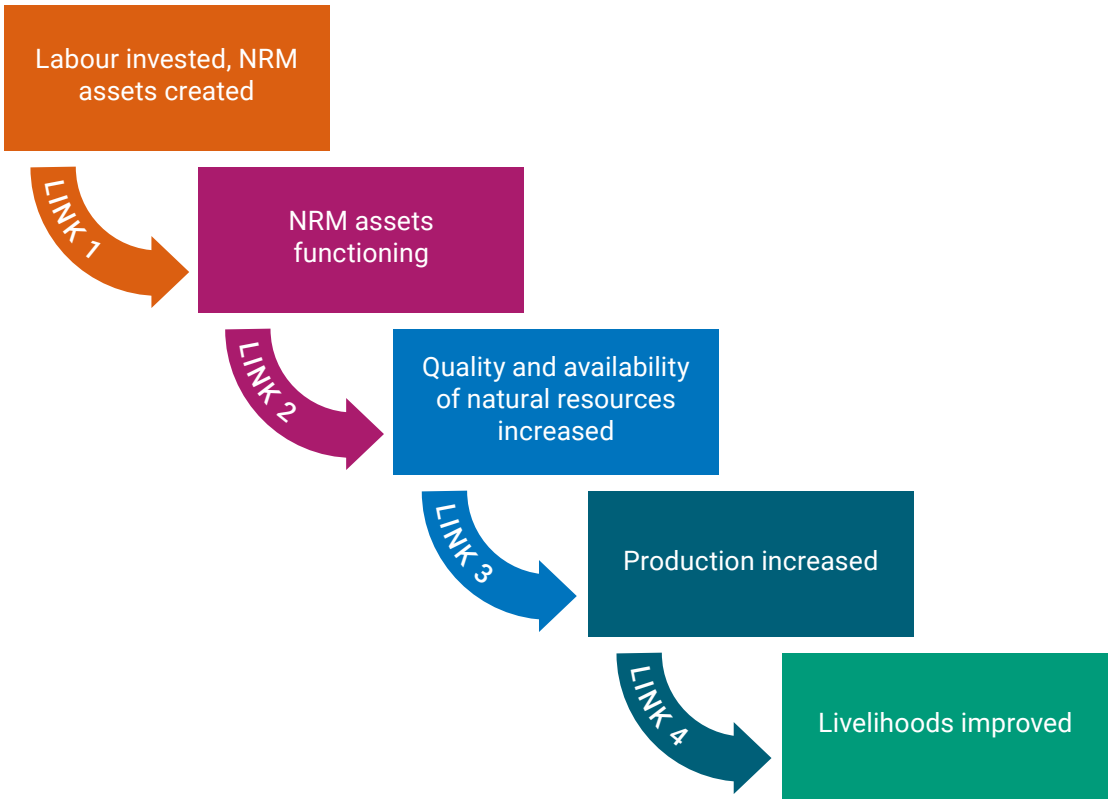
Methodological overlap meant that we looked for at least two different methods, which were not subject to the same bias or error, to gather information for each finding. (This is made clear in the sections below, where individual findings are given alongside an explanation of the different methods used to investigate them.) Evidence ranged from physical and chemical analysis of soils from areas with and without SWC, to published scientific evidence on crop yields in the same agro-ecological area, to counting the number of people who used different water sources on a regular basis over a full season. Where evidence came from interviews, two different approaches were used. In each village, some interviews used an approach adapted from 'goal free evaluation' (Youker, 2013; Scriven, 1991) by adopting a partially blinded approach. Informants were not aware that the object of study was the impact of specific assets.[6] Other interviews asked respondents to follow the causal chain from intervention to improved resilience (see Figure 1).

To enable such a wide variety of tools and methods to be integrated seamlessly, we created an overall analytical framework, derived from theories of change. This approach followed 'theory-based impact assessment' or a 'programme theory evaluation' approach.[7] Theories of change were identified that were either implicit in the programme logic or possible alternative causal chains by which the intervention could have led to positive change. These were tested by first making them explicit in a causal model (see Figure 1 for a generic example), which provided a framework to which a multitude of tools and approaches could contribute. This created a fully integrated mixed-methods and multidisciplinary approach. Various techniques were then used to explore and appraise each causal link, and these are set out in Section 3 and Section 4.

6 In a full goal-free approach, even the interviewers do not know the real object of study.

7 See, for example, Funnell and Rodgers (2011) on programme theory evaluation, Donaldson (2007) and White (2009) on theory-based impact assessment and White and Phillips (2013) on creating an integrated framework for rigorous impact attribution.

FIGURE 1: CAUSAL MODEL THROUGH WHICH NRM ASSETS CAN IMPROVE LIVELIHOODS



Source: Authors.

2.2 Case studies and study sites

The study explored the livelihoods impacts of NRM-related assets that were created around five years previously by PWPs on public land (i.e., public assets) in Ethiopia and Kenya. In both countries, study areas were selected in collaboration with the implementing agencies. Because the study was a methodological exploration and not an evaluation, sites were selected with a deliberate bias, looking only at projects that had been deemed successful.

In Ethiopia, the research studied SWC measures created through the national PSNP in two watersheds in Kalu District, South Wollo Zone, Amhara Regional State. In Kenya, the research assessed earth dams and water pans created through FFA PWPs in northern Makueni County. This work was implemented by international non-governmental organisations (INGOs) in collaboration with the local government under a WFP programme.[8] Details of the sites for both countries can be found in the Appendix.

8 Protracted Relief and Recovery Operation (PRRO) 10666. WFP (2012)

3. FINDINGS FROM THE ETHIOPIA CASE STUDY

3.1 The assets and what was already known about their impact

The SWC measures were of two kinds: the creation of a protected area (an 'exclosure')[9] on hilltops; and gully control and SWC on hillsides on public land. (SWC on private farmland has also been promoted, but it has not been undertaken through PWPs. Such private SWC was not studied.) The creation of the exclosure involved two processes: the use of labour to plant trees and to create physical structures for SWC to control erosion and run-off and to enhance water infiltration; and establishing rules and enforcement mechanisms to limit exploitation to allow for natural regeneration.

The first of these processes was implemented through the PSNP. The exclosures were intended to allow the regeneration of vegetation, which would lead to less erosion from run-off and greater natural productivity of the protected area, including from grass that could be harvested as livestock fodder and the controlled harvesting of other products, including honey (mentioned frequently by almost all key informants at national and district level).

SWC measures included tree planting, infiltration trenches, some limited terracing, building check-dams to control gully erosion and creating micro-basins known as 'eyebrows'.[10] Although the work should have fitted into a watershed management approach, PSNP work was planned and implemented within administrative boundaries.[11]

The official narrative of the success of the PSNP's SWC was confirmed by almost all informants at national and district level, who claimed that SWC had led to reduced erosion, increased yields and greater resilience. Several informants said villages were making considerable profits exporting honey as a result of the environmental recovery in the exclosures. However, our extensive search found no documented evidence of such livelihood impacts as a result of SWC through the PSNP. A few studies report data on positive environmental impacts of exclosures, through increased biodiversity, soil flora, etc., but these all focus only on the conditions *within* the exclosure itself. They do not examine any changes in farm productivity resulting from the creation of the exclosures, and much less on the livelihoods of farmers.[12]

---

9 These are often called enclosures, but the purpose of an enclosure is normally to enclose, i.e., to keep animals *inside*, e.g. to claim exclusive rights to use land. These protected areas were designed to *exclude* animals, therefore they are more properly called exclosures (Yirdaw and Monge, 2017).

10 Technical details of the SWC are not necessary to understand this study.

11 The boundaries of a watershed are set only by topography and how water flows, and they are unrelated to administrative boundaries.

12 Studies have assessed the impact of exclosures: on the regeneration of native vegetation (Damene et al., 2013; Verdoodt et al., 2010 and 2009; Tekele, 2001); on the composition and diversity of plant species (Getachew, 2014; Angassa and Oba, 2010; Aynekulu et al., 2009; Abebe et al., 2006; Mengistu et al., 2005a and 2005b; Asefa et al., 2003); on soil cover, soil erosion, soil fertility and organic matter content (Damene et al., 2013; Yayneshet et al., 2009; Mekuria et al., 2007; Descheemaeker et al., 2006a); on rain water infiltration (Hongoi et al., 1995); on tree cover (Hirvonen et al., 2022); and on the production of fodder and woody biomass (Descheemaeker et al., 2006a and 2006b; Mengistu et al., 2005a).

One published impact study on PSNP treats the impacts of the assets (SuDCA and Soberland, 2015), but it serves to illustrate the weaknesses of approaches to impact assessment rather than to provide evidence in support of them. Here, a reported 20% increase over five years in agricultural yields is based only on reports from government officials; and 40%–60% of this increment in production is reported to have been 'confirmed' in FGD as attributable to PSNP, though without making it clear how this attribution has been made or verified. Given the lack of any formal measurement by farmers of their yields, the variability in growing conditions from year to year, and the absence of any control, these findings are not adequate to draw conclusions about the impact of PSNP assets.

Our findings on the impacts of the SWC assets are set out as follows. The framework of Figure 1 establishes causal links, each of which provides a question, e.g., 'Are the assets still functioning?' (Link 1), and 'Has the quality and/or availability of natural resources increased?' (Link 2), etc. For each question, we set out the findings and briefly outline the methodologies and tools used to confirm or refute each causal connection (see the subsections '*How do we know?*').

3.2 Link 1: Did the PWP create assets that are still functioning?

An underlying problem made it difficult to establish whether the PWP assets were still functioning: the physical location of the PWP implementation had not been documented, therefore it was hard even to locate the assets. Records covered planned, rather than completed, work. Only aggregate figures were recorded on the total numbers of assets created in the district, and no one had documented what had actually been undertaken. This not only made it impossible for monitoring to have taken place, it indicated that monitoring had never been planned. Our research team spent considerable time visiting and interviewing in villages to identify which hillsides had been treated and the range of assets constructed.

As a result of these extensive field visits, we believe it is reasonable to conclude that most of the physical assets (e.g., cut-off drains, terraces, infiltration trenches and check-dams) still existed and were functioning satisfactorily. However, it was impossible to know how many other structures had been created of which no trace remained. Taken individually, their construction was judged to be technically competent. However, the number of structures created was often excessive, so that many served no additional SWC purpose. This indicated that more labour had been spent than had been needed to achieve the objectives of erosion control.

Tree planting had been less successful, with the majority of seedlings having died. Although district officials blamed poor management by the community and insufficient control of grazing by goats in the exclosures, reports by other informants suggested that it was more likely because of weaknesses inherent in the planting process. Seedlings had been planted in the dry season – the appropriate time for PSNP labour absorption targets, but the wrong time of year for tree planting. Furthermore, they were often planted so closely together that they had a low chance of survival.

The over-production of physical structures, mass planting of tree seedlings and inappropriate planting time all suggested that the design and implementation of the PWP had been guided by labour absorption needs at a particular time, rather than by an objective to create assets with lasting benefits.

How do we know?

We searched for local government records of assets created and for monitoring data on their functioning, but records had not been kept.

Our assessment of functioning assets had to be drawn from information from a large enough sample of villages or hillsides. This sample size would be too large for an in-depth study, because of time and resource constraints. We therefore established two different sampling approaches. A relatively rapid review was carried out, whereby we visited 21 hillsides in randomly selected *kebeles*,[13] simply to establish whether or not the assets created in the PWP were still there and still functional. This answered the Link 1 question. Subsequently, only four of these *kebeles* were investigated in depth to answer the questions arising from Links 2–4. This is an example of how our analytical framework allowed us to use different methodologies with different sample sizes for different purposes.

Land cover/land-use mapping was attempted using satellite images, although the resolution of the images was problematic. A Satellite Pour l'Observation de la Terre (SPOT) image from 2007 was obtained, but cloud cover had obscured the locality. An analysis of Google images did not provide data of adequate quality or resolution.

3.3 Link 2: Did the assets improve the natural resources?

Impacts inside exclosures and on farmland need to be considered separately.

In the exclosures:
The PSNP interventions had led to visible environmental improvements and had contributed to reducing soil erosion and run-off. Where enforcement of rules was strong, vegetation had regenerated to some extent and biomass production (grass, shrubs and trees) had increased.

Creation of the exclosures involved two distinct processes: using PWP labour to create physical structures and plant trees; and establishing and enforcing management rules for the exclosures to allow natural regeneration. These processes could be undertaken separately. Some regeneration would have taken place just from the creation of the rules of the exclosure, without requiring the use of labour or PWP. The physical works from PWP labour may have accelerated the process of regeneration, but it is hard to assess their distinct contribution.

13 A *kebele* is the lowest administrative unit in Ethiopia, typically comprising a number of villages.

FIGURE 2: VISUAL COMPARISON OF CONTROL AREA AND PWP AREA

Note: The hillside on the left is still used for grazing animals; the hillside on the right is currently only used to cut grass.

On farmland:

The predominant narrative reported by officials and villagers was that erosion had been reduced, gullies had been rehabilitated and soil fertility in fields had improved. To examine these claims, it was necessary to distinguish between gully erosion (where surface water collects in a stream to dig a channel down a hillside) and sheet erosion (which takes place across a slope). The former causes visible loss of soil and damage down a channel, i.e. over a small area. The latter causes largely invisible damage to significant areas, as the fertile parts of soil are gradually washed away down the slope.

Gully erosion had been controlled by the check-dams. Evidence of previous erosion and of the recovery were visible. Gully erosion had caused significant damage, making areas of land unusable for farming and destroying crops when the erosion occurred during the growing season. Individuals told plausible stories of losses from erosion and of crops at the bottom of the hill being damaged by sedimentation. However, although striking, such erosion had affected only a small number of people and usually only a small proportion of their farmland. Direct and clear evidence supported the testimony that this erosion had reduced as a direct result of the PSNP assets.

We heard some reports of increases in crop yields. Informants attributed this to the SWC interventions having improved soil fertility on farmland by reducing sheet erosion. As noted previously, although these claims have sometimes been presented in studies as conclusive evidence of impact, they needed closer investigation for three reasons. First, rigour demanded that all reported impacts required verification, particularly claims relating to the success of state programmes on what was a highly politicised issue. Second, other interventions had been carried out on farmland at the same time, which could explain increases in yield. These included farmers' own SWC measures on individual farmland, and the promotion of both organic and inorganic fertilisers. Third, reports of success were contradicted by other farmers who argued that neither soil fertility nor production had increased significantly.

Physical and chemical analysis of soils from the hillsides with and without the PWP showed no significant differences for any of the parameters between treated and untreated hillsides. It was not possible to compare changes over time in soil fertility because there were no baselines of soil properties. In the survey, farmers were asked about changes in soil erosion on their fields: in one pairing of a treated and untreated hillside, there was no statistical difference in the answers; in another pairing, more famers said that soil erosion was reducing in the untreated, compared to the treated, hillsides (i.e., things were comparatively worse in PSNP hillsides!).

There was no evidence that water infiltration had improved significantly. There had been no new springs or other improvement in water availability in the five years since asset construction, even though the rains immediately before the field research had been very good.

It is notable that the claims that SWC had improved soil fertility arose in FGDs that were explicitly discussing the impact of the PWP assets, and that these claims were not corroborated in the blinded FGDs. Overall, these claims cannot be accepted as accurate, and we believe there are more plausible explanations for the testimonies:

- Good rains prior to the research visit had resulted in a bumper harvest after three years of poor yields. People may have reported this as a sign of improved soil fertility.
- Close attention to the language used in reporting successes showed that it mirrored the language used in official agricultural extension messages. Public works for SWC have been highly politicised, with significant pressure to have – or to report – success in meeting targets.
- The successful control of gully erosion benefited a small number of farmers directly, resulting in yield increases for them. This may have been conflated with broader improvements in soil fertility.
- The two types of soil erosion have generally been conflated in discussions around PSNP and SWC, including by aid agencies (see below). This has helped agencies give favourable reports.

The clear conclusion is that the PWP had a significant impact on gully erosion on small areas of land that had suffered from highly visible erosion damage, but the works had no significant impact on the majority of farmland.

How do we know?

In the exclosures:
Observation during site visits and visual comparison with the control areas clearly showed a physical difference in vegetation. Figure 2 shows neighbouring hillsides with a clear boundary between a regenerated area (an exclosure) on the right and a degraded area (outside the *kebele* where the PSNP had been implemented) on the left.

Attribution of the impact of PSNP assets was supported by scientific analysis of biophysical changes and was consistent with evidence from other studies. The extent of vegetation regeneration was consistent with the time since PSNP implementation and with studies from other locations (e.g. SuDCA and Soberland, 2015; Mulugeta, 2014; Abebe et al., 2006; Mengistu et al., 2005a; Asefa et al., 2003).

Gully erosion
General testimony in FGDs was complemented by direct reports of the specific experiences of individual farmers on their farmland. These reports were supported by observational evidence and comparison with the control hillsides. The evidence and the testimonies were all consistent with the science of gully erosion. No further methods were needed.

Sheet erosion
There were contradictory claims regarding improved fertility on farmland as a result of reduced sheet erosion, therefore a wide variety of methods had to be used to establish the truth of this question.

Soil samples were collected from hillsides with SWC assets and from neighbouring (control) hillsides where the topography and land-use patterns were similar. These samples were analysed in a laboratory for parameters associated with erosion and soil fertility, including physical structure, nutrient content and percentage of organic matter.

FGD in the villages followed two distinct approaches. Some groups discussed the assets directly and examined their chain of impacts. Other groups were run 'blind', discussing general changes in farming and water, without the participants knowing that the study was interested in SWC assets. (The discussions were held in parallel to avoid 'cross-contamination'.) Pro-project bias or politicisation was more evident in FGDs that focused on PWP assets.

A survey was conducted on a random sample of 1,392 households across the four study sites. The survey included questions on changes in water availability and springs in the area. Survey respondents did not know the purpose of the survey, thus preventing bias.

Rainfall data were collected for the years between the field research and the creation of the assets to examine potential causes of changes in water availability and crop yields (see also Link 3).

The findings of other studies were used to examine the scientific plausibility of claims made regarding improved soil fertility. This included scientific studies conducted nearby at Maybar research station[14] on the extent of soil erosion in farmland with a similar topography (see also Link 3).

14 The Maybar research study is part of the Water and Land Resource Centre of Addis Ababa University. A large number of scientific papers have been published from research at the site.

3.4 Link 3: Did the improvement in natural resources lead to improved productivity?

Logically, it was unlikely that there could have been significant impacts on productivity in the areas where Link 2 was found to have failed (i.e., assets had had no significant impact on the natural resource base). Once a causal chain has broken down, it may not make sense to continue investigating further links. However, we proceeded to examine Link 3 for two reasons. First, this was a form of triangulation or confirmation of the interim conclusion from Link 2: if there was evidence that yields had improved, this might challenge the conclusion that Link 2 had failed on most farmland. (Alternatively, it might suggest that there were alternative causal pathways that had not been tested.) Second, the study was designed to test research approaches. This could be done to some degree even in the absence of impact.

In the exclosures:
There had been a significant increase in some kinds of productivity within the exclosures. By keeping animals out, grasses grew and cut-and-carry of fodder for livestock had been introduced. It proved possible to quantify this productivity to a certain degree, and to ascertain that it was significant. There was consistency and technical plausibility in respondents' answers to questions about areas harvested, fodder yields, the time taken to harvest, the monetary value of fodder sold and the duration it could sustain animals. These reports aligned with published literature from Ethiopia.

As discussed above, these improvements were probably largely attributable to the creation of rules governing the use of the hilltops, rather than to any public works *per se*.

On farmland:
Evidence was visible of crops being planted on land that had been reclaimed after the control of gully erosion, and this was supported by reports from farmers and key informants.

Some claims were made of general increases in crop yields across the hillside as a result of reduced sheet erosion, as discussed above. Some farmers reported yield increases of up to 300%, which was surprising and thus warranted more intensive investigation.

A general increase of 300% in crop production in the *kebele* would be noticeable in trade, e.g., in the movement of food crops into and out of the *kebele*. One would expect to see greater volumes sold, particularly after harvest, and smaller quantities imported into the *kebele* for consumption during the rest of the year. Our approach of methodological adaptation meant that this had to be explored, although trader interviews had not been planned originally. Traders reported that, if anything, the opposite trend was being seen, i.e., towards more food supplies entering the village and less being marketed.

To control for other explanations for changes in yields, we compared reported changes in production between areas where PWP assets had been constructed and control watersheds. No patterns emerged that linked reports of yield changes to PWP assets.

The claim of a causal pathway from SWC assets to hugely improved productivity requires some explanation: what is the physical, chemical or biological mechanism by which yields increase so much?

Soil loss and crop yields under different conservation practices have been monitored over many years at the nearby Maybar research station.[15] Studies there have shown little evidence of the impact of SWC assets on agricultural productivity over the timescale involved in this study. Levels of soil loss that might have been experienced in the case study areas would result in reductions in yields of around 0.25% for every centimetre of soil lost (Ludi, 2004). Uncontrolled erosion would cause yield losses of less than 1% per year. The impact on crop yields of interventions to reverse soil degradation would therefore be very small, even if the interventions succeeded in halting any further erosion.

In fact, studies show that SWC tends to reduce crop yields, because the land taken up and lost from cultivation more than outweighs the positive impact on soil fertility (Adimassu et al., 2017; Adimassu et al., 2012, Herweg and Ludi, 1999).[16] (As discussed above, changes in soil properties that might lead to higher crop yields were not evident in soil analysis.) Putting claims of anything like 300% increases in yield to one side, it has been generally shown that farmers would not be able to detect changes in yields of a magnitude that is plausible against the backdrop of significant annual yield fluctuations caused by so many other factors.

For a combination of these reasons, the reported impacts of PSNP assets on crop productivity can be discounted, apart from in areas of previous gully erosion.

15 Although the Maybar research station is located at a higher altitude than either of the study sites, and therefore is likely to have slightly higher rainfall, it was the closest site from which to obtain credible data and so was used to gauge the feasibility limits of the impact of SWC on natural resources and crop production.

16 Adimassu et al.'s (2017: 87) review of the literature on SWC and crop yields concludes that 'the impact of physical soil and water conservation practices on crop yield was negative mainly due to the reduction of effective cultivable area by soil/stone bunds'.

How do we know?

Productivity of the exclosures could not be measured directly, and so had to be assessed from reports from respondents. To increase the rigour of the evidence, questions were asked in various ways to quantify fodder production, e.g., amounts harvested and periods between harvesting; time taken to cut and trips taken to carry fodder each time; number of animals that could be sustained over what period after each harvest; monetary value of fodder harvested (triangulated by establishing the local price of fodder). Triangulation was facilitated because different *kebeles* had adopted different approaches to exploiting the exclosures (see Link 4), which meant that harvesting patterns varied considerably. Reports were obtained from across the FGDs and from individual interviews. The reports were sufficiently consistent to have credence, and they contained enough detail for us to believe that bias and politicisation had not influenced respondents' opinions.

Literature from the northern highlands of Ethiopia was used to establish typical values for fodder consumption per animal per day, and typical values for fodder production per unit area (Aune et al., 2006).

Observation of crops growing on land visibly reclaimed from gully erosion supported farmers' testimonies and was sufficient to establish a causal link with a high degree of certainty. The SWC structures from the PSNP were clearly identified as the cause of the physical rehabilitation (see Links 1 and 2).

Because of contested reports, crop production increases were assessed again via a wide variety of methods (see also Link 2).

Reports were collected in separate FGDs asking directly about the SWC assets and in 'blind' FGDs asking about changes in farming without mentioning the PWP assets.

A survey on a random sample in the PSNP watersheds and in the control watersheds was used to compare changes in crop production and in household food security (looking particularly at a household's own food production as measured in consumption months).

Local traders who bought and sold agricultural goods and food from and into the *kebeles* were asked about changes in recent years. They were not aware of the particular focus of the study.

Rainfall data were used to identify other causal factors affecting crop yields over the five-year period since asset construction.

Finally, evidence from a nearby crop research station helped establish the scientific parameters of feasible soil changes and how these would translate into changes in crop yields.

3.5 Link 4: Did improved productivity lead to improved livelihoods?

In the exclosures:
Key informants in Addis Ababa frequently offered reports of the benefits of exclosures. These benefits were reported mainly in terms of production from cut-and-carry fodder from exclosures, as well as secondary benefits such as honey production. However, reports of improvements to productivity tended to ignore any assessment of the land's previous economic productivity and its capacity to support the free grazing of livestock.

This study attempted to calculate the economic value of both the previous grazing regime and of fodder production from cut-and-carry. Precise calculations were not possible, but it was still important to obtain a rough estimate that could at least give the scale of magnitude of any benefits. Compared to the numbers of animals that were sustained by free grazing, the net advantage of fodder from the exclosures was roughly enough to sustain eight cattle over a year (shared between around 125 households). The most generous estimate of *gross* economic benefit was just $4.50 per household per year. Significant labour would have been needed to harvest and transport the grasses, however, and once such opportunity costs are taken into account, the net benefit would be zero or negative, despite the benefits claimed.

The distribution of any benefits varied from hillside to hillside. In some locations, all households were allocated a strip that they could harvest (and which they could sell if they did not have livestock themselves). In others, only a small number of households who had prior claims over the hillside shared the grass from an exclosure.

Although informants at national level talked generally of honey production as a common and significant impact of exclosures, no mention was made of honey in any of the case study *kebeles*, or of any other secondary economic products. Indeed, the only documented evidence of honey production that was related to the PSNP came from a 'PSNP Plus' project, which specifically promoted honey production through the provision of inputs and training and the creation of market linkages (Burns and Bogale, 2012).

On farmland:
The lack of improvement in the overall productivity of land beyond gully rehabilitation has already been established under Link 2. Livelihood benefits were limited to a small number of individuals who had fields that had been directly affected by run-off and sedimentation previously.

How do we know?

The economic benefits of exclosures were first sketched out in FGDs and with key informants. Identified benefits were then quantified through detailed interviews in villages that applied different systems to share the benefits. In some villages, fodder was quantified by how much could be sold and at what price; in others, it was quantified by how many livestock could be sustained over what period. The broad similarity in results across these different calculations gave confidence to their reliability.

The economic cost of the labour used to cut and carry fodder was established in FGDs with those who used this practice.

Calculations of the economic value of free grazing was estimated from the numbers of animals grazing, as given in FGDs and from visual estimates of animals grazing on similar hillsides.

The randomised survey sample helped to establish how many people were able to benefit from fodder in each exclosure. Combined with population lists, this allowed an approximate estimation of total harvests.

4. FINDINGS FROM THE KENYA CASE STUDY

4.1 The assets

The Kenya study assessed the livelihood contributions of earth dams and water pans created in 2009–2010 through food-for-work PWPs in northern Makueni County, Kenya. Makueni is situated in a semi-arid area, where water resources are limited and food security is a challenge.

The PWPs were implemented by non-governmental organisations (NGOs) in collaboration with the local government, within the framework of the WFP Protracted Relief and Recovery Operation (PRRO) 10666. The overall programme sought to build 'household and community assets to strengthen resilience against shocks' (WFP, 2012). Water was intended for both human and livestock use, and the anticipated benefits included a reduction in the time taken to collect water; an increase in drinking water for humans and livestock; increasing opportunities for engagement in additional income-generating activities; improved livestock production; and an increase in the value of farm production.

It was appropriate to apply the broad causal chain depicted in Figure 1 to analyse the impact of the assets designed to enhance livelihoods and resilience by improving the quality and/or availability of water resources. Therefore, this section poses the same questions about each causal link.

4.2 Link 1: Did the PWP create assets that are still functioning?

The project had constructed or rehabilitated 27 dams five years before the fieldwork. Of these, 20 dams were randomly sampled for a rapid physical and technical assessment of their design appropriateness, physical integrity and the degree of siltation. This appraisal was made four weeks after the end of the rainy season, by which time only five dams had usable water. The technical appraisal identified serious deficiencies linked to the dams' ability to function as intended. This included shortcomings with regard to structural integrity from design or construction faults in 12 dams; and poor siting for most dams in terms of topography, geological conditions or soil morphology, leading to high seepage and/or poor recharge of water. The design of shallow pans with a large surface area also resulted in high evaporation losses. Thirteen of the 20 dams had high levels of siltation that had further reduced storage volume.

Although water user committees had been established to manage the dams after they were handed over to the communities, none of these committees continued to function after the completion of the works. This is because the committees comprised beneficiaries of the PWP (i.e., people who had contributed labour) rather than dam users. Maintenance had not been carried out. Only three of the 20 dams were protected and fenced.

In conclusion, the assets still existed, but they were not functioning well.

How do we know?

Information on project design, implementation and the current functionality of dams was collected from key informants.

To avoid basing any conclusions on a small number of dams – i.e., a number that could be investigated in depth – a rapid appraisal was conducted of 20 randomly sampled dams of the 27 created. Visits were made four weeks after the end of the short rains. Brief interviews were conducted with local officials; and the physical existence and structural integrity of the dam were visually appraised by technical experts, with functionality noted in terms of water holding. Water availability in the dams was checked again in early March (12 weeks after the short rains) through phone interviews with village elders.

The technical appraisal noted the dam: location, design, siltation, physical damage, contamination (signs of animal droppings and footprints), protection status and vegetation cover in the catchment, as well as water availability. Technical plans and specification documents were reviewed. Overall functionality was appraised against three criteria: design appropriateness, physical integrity and degree of siltation.

The above parameters had not been monitored by the project, and documentation did not exist against which the study findings could be compared.

4.3 Link 2: Did the assets improve the natural resources?

Of 20 dams visited, only five had enough water to be usable just four weeks after the end of the short rains. Eight were dry. One dam had dried within 10 days of the end of the rainy season. Although one informant blamed poor rains for the lack of dam recharge, verification revealed that total rainfall in the season immediately preceding the appraisal had *not* been low (although it had been poor for agriculture because rains were poorly distributed).

A deeper study of five dams was conducted in the following season, after the long rains. Three of the five dams remained functional for only one month into the dry season after the long rains; one was functional for two months; and one dam held water for the length of the dry season. The quality of water in most dams was poor and unsuitable for human consumption, because the dams were not maintained and almost all water was visibly contaminated by livestock.

The impact of any additional water supply for the villages was marginal, because these dams held water for the same period when villagers had several other water sources, including shallow wells and household catchment from roofs. The mean numbers of daily users for each dam ranged from just eight to 18 during the short period when water was still available.

In conclusion, the dams created through the PWP made only a marginal contribution to improving the availability of natural resources (water) for the villages in the catchment areas.

How do we know?

Dam performance was investigated through the rapid assessment of 20 dams after the short rains, and through a more detailed technical assessment of five dams after both the short and long rainy seasons.

Dams were available for use by people from more than one village, and so investigation at village level had to be supplemented with investigation at dam level. Observers were hired to watch and gather data on every user of each of the five dams on one day per week from 07.00 until 19.00, from the end of the long rains until the dams ran dry. This established the duration of water availability within the dam, the duration of use, and the number of people and livestock using them. The observers also asked all users where they came from and what they used the water for.

The explanation for poor dam recharge (low rainfall) was checked – and refuted – by obtaining rainfall records from a nearby weather station.[17]

Information on water availability, water quality and the use of different water sources was gathered in FGDs in the five villages nearest to the dams. At least one blind FGD was conducted in each village, and at least one FGD was conducted in each village where respondents spoke specifically about the water sources created by the PWP. Further information was gathered from key informants.

A quantitative survey was used to gather Information on the different water sources available in villages in each month and which sources households used for different purposes. The survey was administered to 20 randomly selected households in each of the five villages closest to the dams and in five other nearby villages.

4.4 Link 3: Did the improvement in natural resources lead to improved productivity?

Given their very limited contribution to water availability, it was unlikely that the dams could have had a significant impact on livelihoods. However, as with the Ethiopia case study, we continued to look for possible benefits in productivity and livelihoods, both to verify our negative findings about Link 2 and because the research was designed to test approaches to the retrospective study of livelihood change.

The range of livelihood activities that the dams were expected to support was not realised. There was little use of water for human consumption (because of poor water quality). Just one or two people per dam used the water for vegetable gardening and there was almost no use for brickmaking or any other economic purpose beyond watering livestock. Even this use had no impact on livestock production, because the limiting factor for livestock production was fodder, not water. The pans did not provide additional water availability, but instead merely substituted for other available water sources: dam construction did not result in any change in the frequency of watering animals, nor in the time taken to water them.

17 Makindu meteorological station, Makueni County.

The insignificant impact of the dams on productivity was largely because water was available for a short period only. Had water been more readily available, there would have been some other constraining factors to certain economic uses, e.g., vegetable production could only have benefited the few people who owned land near the dams. The dams did not address people's priority needs, which included more water of drinking quality, the supply of water for various uses when other sources were unavailable and other constraints to livestock production.

How do we know?

Information on water use was obtained from FGDs, which included questions about practices related to various economic activities that use water (livestock keeping, horticulture, brickmaking, etc.) and about trends and changes over the previous five years.

Although FGDs offered rough estimations of the prevalence of different activities or phenomena, these cannot be relied upon as being free of bias. The estimations were therefore independently established from the survey detailed in Link 2. The survey was designed and conducted so that respondents did not know the focus of the study, and it was not clear until the last question that there was any particular interest in one specific water source.[18] The survey asked about the sources available by season, and their use by season and by purpose; livestock watering practices, and livestock management practices and constraints; and for comparisons on these parameters with five years previously, (i.e., before the PWP dam construction).

Dam use monitoring, as described in Link 2, provided information on who used the dams created by the PWP and for what purposes.

4.5 Link 4: Did improved productivity lead to improved livelihoods?

Projects had reported on livelihood impacts from the dams (see, for example, WVI, 2014). This included time saved in collecting water for productive agricultural activities, and the use of water from the pans to irrigate maize. But the projects had not monitored impact. Instead, they had engaged in the common practice of collecting 'success stories', which remained unverified.

Our rigorous attempt to find livelihood benefits identified only one or two people who had experienced any meaningful economic benefit from each dam. No corroborating evidence could be found for use of water to irrigate maize, and no time savings were found from using the dams to collect water. (To be strictly accurate, we found one person across the study sites who said they had saved time. However, this individual was also an employee of the implementing agency.) Economic resilience had deteriorated since dam construction: herd sizes had depleted because of distress sales (i.e., reluctant sales to raise money for immediate needs), and food security had declined due to decreasing demand for agricultural labour. Overall, the dams had almost no impact on livelihoods. This was unsurprising, given that the dams had not made a significant contribution to improving access to water or the quality of water available.

18 This last question was added deliberately at the end to quantify pro-project bias (see Box 3).

Because there was no impact to analyse, we instead looked at the *potential* benefit of well-functioning dams. This was to assess what contribution improved water sources could have made to economic resilience.

Analysis showed that the returns from any possible water-dependant economic activity would not have been sufficient for households to escape poverty, let alone to reach any level that could be described as resilient (i.e., sufficient to keep a household out of poverty even in the event of a normal shock). Vegetable gardening could help a household to earn around $250 in a three-month growing season, but only if the household land was immediately adjacent to a water-pan. Brickmaking had a broadly similar income potential, depending on the number of adults in the household with spare labour. Even if a three-month water source were available after both short and long rains, few households would be able to earn as much as $500 over the year from the water resources. Considering only the period of time when at least one adult would be engaged in the activity, this would equate to just $2.80/day, or $0.78 per person per day for a fairly typical household of five.[19] This represented just one-third of the international poverty line.

How do we know?

The insignificant scale of the economic use of the water from the PWP assets was evident from the fact that they offered hardly any additional water. This was confirmed by FGDs, which established the range of economic uses of water, and the survey, which also established the prevalence of different economic activities.

In the absence of actual economic activity as a result of the dams, we investigated potential economic activity by studying the economics of water-related activities carried out in the village by people using other, more reliable, water sources. Very detailed interviews were held with several people about each of four water-related activities that had been identified in the FGDs (vegetable gardening, brickmaking, water vending and livestock watering). Each interview established the scale at which the activity was being carried out by the household, and covered all input and investment costs, time inputs, resultant revenue streams and the time period over which revenue was generated. The interviews also established access constraints to activity (e.g., owning land close to the dam or owning a donkey to transport water) and the opportunity costs of the activity, by asking about the activities that individuals would have carried out had there been no water.

19 Calculations per person use weightings for children following KNBS (2021), i.e. 0.24 for a child under 5 and 0.65 for children 5–14 years. The illustrative household comprised two adults, two children between 5 and 14 years, and one child under 5.

5. IMPLICATIONS FOR PWPs AS A VEHICLE TO BUILD RESILIENCE

The lack of livelihood impacts from the assets was surprising, because we had deliberately chosen two projects that we had been told were successful. Robust investigation clearly established that the assets created had resulted in only negligible impacts on livelihoods – despite this being the rationale and justification for investing significant government and donor funds, and for demanding labour inputs from food-insecure people targeted by the projects.

The projects had both generated 'success stories' around the promotion of sustainable livelihoods benefits from PWP assets, and this had diffused into the collective wisdom of governments and the aid community. However, these stories were found to be highly implausible when investigated properly.

The relevance of this study is not the inadequacy of the assets created by two particular PWP projects. Indeed, it could be argued that little can be learned about the role of PWPs in building resilience today by looking at assets in a limited number of project sites in just two countries a decade ago. The importance of this story lies in two concerns that it raises: that so much is being assumed about the benefits of PWP assets without any confirmation; and that the reasons for the failure of the assets to have any impact is so clearly linked to the fact that they were constructed through PWPs. The particular features in the design and implementation of the PWP that caused the problems in the case studies remain common to much PWP practice globally.

At the very least, this ought to be a major concern among those proposing, funding, designing or implementing PWPs, unless they have taken specific steps to address and mitigate these identified failings. Because so little consideration is still being given to exploring the impacts of PWP assets on livelihoods, however, such remedial action is likely to remain rare. As a result, investment decisions continue to be made in the blind hope that project assumptions will turn out to be true. This shuts off opportunities to learn how best to use PWPs, and in which circumstances.

The study revealed that the PWP assets had not contributed to resilience because of the tension between dual objectives: the short-term objective of addressing consumption needs through a wage transfer, and the longer-term objective of building resilience through the creation of quality assets. This tension is relevant for the majority of PWPs being implemented today. In our case studies, at every stage in the project cycle, the short-term objective either took priority or was the only objective recognised and supported. The assets therefore failed to have any sustained impact on livelihoods because of generic factors in the way that PWPs are conceptualised, designed and executed in the social protection and humanitarian sectors. This starts with the way in which projects are geographically targeted, assets are selected, and projects are designed and implemented. Unless and until these issues are addressed, many PWPs will experience the same failures.

Although the longer-term benefits were central to PWP discourse at the time, both programmes focused on creating a project that could absorb large amounts of labour rather than prioritising the quality and functionality of the assets created. This can be seen most obviously in the implementation of the projects – planting trees at densities almost guaranteed to prevent their survival in an effort to absorb enough labour in Ethiopia, and constructing far more physical structures for gully control than were needed; and, in Kenya, an inappropriate reliance on manual labour to compact dams that were structurally unsound as a result.

The lack of genuine concern for the assets and resilience benefits was evident in both PWPs at all project stages. The **problem analysis** driving programme implementation was based on concerns relating to short-term food insecurity needing a wage intervention. There was no parallel or equal analysis of underlying livelihood issues and which assets could be useful. **Project selection** started by identifying areas that were food-insecure, calculating the transfers that needed to be made and then selecting assets that could absorb the required labour. In Kenya, this resulted in the decision to construct dams because other water projects were not as labour-intensive. **Site selection and asset design**, too, were driven by concerns to create assets in areas where food-insecure people could be hired and where enough labour could be absorbed, rather than siting assets where they were needed and where they would function effectively. In Ethiopia, this led to planting trees when people were hungry rather than when seedlings might survive; and in Kenya, it led to siting dams where water losses were inevitable or recharge would be minimal, because a labour-intensive asset had to be created in a given food-insecure area.

Similarly, **budgeting** was based on the amount of food or cash assistance required. Work was created to match this, with the capital and administration costs added. Budgets did not start from an assessment of the financial requirements to build specific assets to a particular quality threshold. Project documents showed little consideration of their value for money as an investment (see Box 2). Despite increasing demands internationally for development assistance to show evidence of value for money, the governments or donors funding the assets did not demand any reports on their quality, sustained functionality or impact. This clearly signifies that the assets were not considered a key component of the project 'value'.

In project **implementation**, construction was not designed to maximise the quality of the assets or even to use labour most efficiently. Rather, construction was designed to absorb as much labour as possible, even if this compromised asset quality. As discussed, this was seen in the construction of potentially redundant assets and the excessive planting of tree seedlings in Ethiopia, and in poor compaction of dams in Kenya, where construction supervision was by staff competent in managing 'food for work' – that is, measuring work norms and ensuring correct payments – rather than by staff with the required technical competencies for a construction project.

In Kenya, water-user associations were given the responsibility for the **management** of dams after construction, but these associations comprised food-for-work beneficiaries who had built the dams, rather than local dam users. This undermined the incentive to manage the dams, which, combined with the poor performance of the dams, led to the identified failures in maintenance. In Ethiopia, no resources were available to maintain assets, which resulted in some cases of poor maintenance and management.

BOX 2: WOULD THE DAMS IN KENYA HAVE BEEN JUSTIFIED AS A RESILIENCE PROJECT, WITHOUT A PWP WAGE JUSTIFICATION?

According to the design documents for the PWP in Kenya shared with the study, the desilting of Syotovali dam increased its capacity by 4,590 m^3. Although it is known that dams suffer water losses from seepage, evaporation and siltation, we have explored its economic rationale starting with unrealistically optimistic assumptions – that there would be no further siltation, there would be no water losses from evaporation or seepage, that the dam would be fully recharged each season and that all water retained would be used productively.

For simplification, we assume all the water was used for one particular economic activity. Our own activity economy analysis found vegetable production to be the most lucrative use of dam water in Makueni. In the absence of data for evapo-transpiration rates in Makueni, a constant irrigation requirement of 6 mm per day[i] is assumed. On this basis, the additional water capacity would be sufficient to irrigate around 0.6 ha for a four-month growing season.

The maximum production from this, on the impossible assumptions of no water losses, would be around $1,000 per rainy season, or, if the dam were to fill up twice a year, up to $2,000 per year. This production figure *does not* represent net benefit, as it ignores the opportunity cost of eight months of labour (for four people).

According to project documents, the cost of rehabilitating the dam was $33,137. On the assumption that labour had no opportunity cost (i.e., no value to people without the work provided by the water from the dam), the annual return would be 6% of the investment. It is likely that water losses would cut this return in half. Once the opportunity cost of labour is included in the calculation, this figure diminishes further.

Given that resources for development investment are finite, spending should be prioritised on investments with the greatest impact on resilience. This study does not seek to determine the rate of return that is sufficient to justify such an investment, but rather to raise the concern that such calculations were not made in the design phase of this PWP.

i No figures for evaporation exist from Makueni. This estimate is based on Dagg et al. (1970), Woodhead (1968) and Kimani et al. (2015). Irrigation demand depends partly on evaporation rates, and partly on the crop grown and the stage of crop growth.

None of the projects **monitored** asset functionality or use, and no information was available about the assets from the implementing organisations – although this did not reduce their conviction that the assets had been a success. Information was only collected on the quantity of work done and transfers made, and, in Kenya, to verify the completion of the work, not the quality of the asset. Monitoring ceased once the use of labour for asset construction had ended.

Because the PWP project cycle in Kenya focused on short-term food security objectives, it did not accommodate any **evaluation** after asset completion. In Ethiopia, the PSNP is conceived as a medium- or long-term social protection programme, and several large multi-year impact evaluations have been undertaken. However, these evaluations have focused on the impact of

wage transfers, with the impact of assets on livelihoods receiving even less attention than their environmental impact.[20] In some instances, the evidence gap created by the lack of evaluation has been filled with individual 'success stories', chosen to confirm the programme logic but without plausible, substantiating evidence.

The incentive structure in PWP has contributed to a lack of attention to the medium-term benefits from assets and it means significant financial resources are being used inefficiently and with limited scrutiny. Little has been learned over the past decades on how best to select, construct and manage PWP assets. This does not imply that accountability is *inevitably* lacking in PWPs, or that assets built through PWPs can *never* contribute to livelihoods. However, the study shows that a conventional PWP approach introduces particular accountability challenges and increases the risk that investment in assets may bring only limited benefits. These challenges are still not widely recognised, and few attempts have been made to date to mitigate the risks inherent in such PWP programming.

Billions of dollars are invested by governments and donors in PWPs as a vehicle for poverty reduction and resilience-building every year. At the same time, poor and food-insecure people have contributed billions of labour days to construct assets as part of these programmes. All of this continues despite the lack of evidence about the impact of PWP assets on livelihoods. Most worryingly, questions have not been raised widely about the untested assumptions that underpin PWP as an instrument for livelihoods promotion and resilience, and about the challenges and tensions inherent in the dominant approach to PWP implementation.

In the absence of rigorous evaluation of PWP assets, questions remain about what is driving continued investment in PWPs. Unfortunately, it is much harder than anyone would wish to design and implement projects that have sustained success. Relatively frequent failure, or partial failure, should be expected. Too often, though, the aid system incentivises reports of success, and no one in the chain from donor to implementer has any incentive to challenge such stories. This risks wilful blindness and reluctance to look for information that challenges the easy narrative of success – even though this is the very information needed to improve investment decision-making. It is hard to interpret the lack of interest in studying the contribution of PWP assets or the ready acceptance of success stories – despite the ease with which they could be identified as misleading, implausible or highly atypical cases – as anything other than the outcome of this incentive structure.

The ready publication of success stories is highly problematic because it has endorsed and legitimised assumptions about PWP impact that are often unwarranted. This disguises the extent to which evidence is lacking about the medium-term benefits of PWPs, and it limits lesson learning. It encourages continued spending in such investments even though they may be ineffective.

20 Rigour is also a challenge. Asset impact studies (e.g., SuDCA and Soberland, 2015) that assess changes in vegetation cover and in production are based on survey results or they report a constantly upward trajectory in yields as a result of the intervention, based on information from local officials (an example where 'key informants' were perhaps neither key nor informants). As discussed elsewhere, such assessments should be treated with some caution. Additionally, such evaluations do not consider neighbouring (non-PSNP) areas to understand other processes of change taking place.

6. IMPLICATIONS FOR LEARNING AND ACCOUNTABILITY

This study set out to develop and to test methodological approaches for understanding the impacts of one kind of resilience-building project in places where baselines had not been established and where many other factors were driving changes. This report demonstrates that the piloted approach worked: there is a high degree of rigour behind the evidence collected and the way in which it has been used analytically to understand how the PWP – or at least the assets created by the PWP – played out in people's lives after project completion.

BOX 3: QUANTIFYING PRO-PROJECT BIAS

Many people tend to tell you what they think you want to hear. But is it *ever* possible to take what people say at face value? Just how widespread is pro-project bias among reports from the so-called 'beneficiaries' of aid programmes?

We set out to quantify pro-project bias through the randomly sampled survey about water use in Kenya. Respondents did not know that the study had an interest in a particular water source created near their village through the PWP. They were asked many questions about which water sources they used in which seasons, for what purposes and why. A final question asked specifically about the water source created through the PWP, asking respondents to rate how beneficial it had been for them personally. It seems reasonable to conclude that if people had already told us that they never used that particular water source, then the water source was of limited, if any, benefit to them.

Of all randomly sampled respondents, 71% reported that the water source had been beneficial or very beneficial to them. However, none of these respondents had indicated via other questions that they ever used the water source, or that it had been of any other benefit to them. These positive answers seem simply an expression of pro-project bias.

If this case is typical, then almost three-quarters of people who do not see any benefits in a project will still be quick to tell researchers how wonderful it has been.

6.1 Lessons for PWPs

It is difficult to kill two birds with one stone

Recipients of a wage transfer from a PWP contribute their time to construct public assets. Those funding the assets spend about half of a PWP budget on everything needed to make people work (materials, supervision, etc.) and not on the transfer itself. This is justifiable only if the assets created bring sufficient benefits. However, it is not clear that PWPs are designed and managed in ways that ensure that projects result in assets that are of an adequate quality. The immediate wage transfer is prioritised in project documents primarily.

It is far harder than hoped to achieve the twin objectives of meeting immediate needs and building assets. The feasibility and cost-effectiveness of using a PWP approach to construct assets should be assessed before the project design phase, and a clear case should be established for the wisdom of maintaining dual objectives.

The benefits of assets need to be studied

PWP assets are not studied regularly, therefore a lack of realism is maintained regarding the feasibility of achieving twin objectives. It has not been possible to develop an evidence base showing which assets are worth building through PWPs and under which circumstances, because this is never established in individual cases. This situation should not be allowed to continue.

Lesson learning starts with monitoring

Many of the difficulties and costs incurred by this study arose because monitoring and information management systems had not looked at asset functionality and use. Even a minimal investment in documenting the functionality and use of assets created would make future lesson learning much simpler and cheaper. More importantly, this would provide projects with real-time information to enable implementers to understand the impact that an asset is having.

Technical understanding is important

Typically, current practice in PWP evaluation does not demand technical expertise in the appraisal of the assets created. This is an essential part of the multidisciplinary approach needed to assess the full causal chain from PWP activities through to livelihood impacts.

BOX 4: WHY ARE EVALUATIONS NOT CONDUCTED RETROSPECTIVELY?

If investments are intended to have medium-term benefits, why are lessons not looked for in the medium term? It is difficult to justify this failure, although four dimensions are pushing the sector away from learning about how projects play out in real life.

Bureaucratic dimension
Evaluations are tied to finite project budgets, so evaluations are conducted just before a budget closes, i.e., before it is possible to know what impact a project will have on people.

Evaluations mainly have an accountability function, rather than prioritising lesson learning – and this is accountability to donors. Evaluations are thus designed only to look at processes and to check whether outputs were delivered. Assessing how useful the outputs were, or for whom, involves accountability to end-users, which is not prioritised.

Few evaluations look at intervention types, isolating one component of projects and evaluating it across several different projects. Thematic evaluations are hard to organise where evaluations are tied bureaucratically to project budgets.

Incentive dimension

All project actors are incentivised to report 'success stories' without questioning the plausibility or typicality of these stories. There are no incentives to challenge this.

System incentives encourage evaluations that broadly endorse the overall model behind a project. This is possible because product satisfaction by the agency paying for the project is based entirely on reporting.

Evaluators become recognised as specialists in a particular kind of project, which is natural. But this increases the risk that evaluators unconsciously adopt the assumptions made by the models they are evaluating, and it reduces the likelihood that they will look outside the project framework to challenge programming assumptions.

Even where evaluations are independent, powerful agencies can exert pressure on an evaluator to conform to certain expectations and to remove unwelcome criticisms.[21] Evaluators who wish to be contracted regularly have an incentive to avoid being too challenging.

Technical dimension

Most monitoring and evaluation focuses on predefined indicators that presume that project assumptions are correct. It is confirmatory, rather than exploratory (Copestake, 2014). Implicit pro-project bias is thus hard-wired into many methodologies.

Pro-project bias in reporting is also huge (see Box 3). Many studies, not just in PWPs, include impact claims that can only plausibly be attributed to such bias.

Because of the accountability function, evaluations focus on talking to people who have participated in a project. Learning about how project assumptions relate to the real world requires also talking to people who have not participated. However, resources are rarely made available to add in this element.

Cultural dimension

Aid projects inspire defensiveness and territoriality. This discourages open access learning. Most people and agencies are cautious about opening up their project to scrutiny, especially to people from another agency.

Scientists are expected to share their data when sending papers for peer review, so that other scientists can make their own analysis. This does not happen in the aid world, where data are not commonly shared. This makes it very easy to cherry pick data and findings.

21 The delay of several years in being able to publish this research is one example of this.

6.2 Lessons of wide relevance

Retrospective assessment is important

The most important lesson to draw is that projects or investments can and should be studied some time after their completion. This is almost never done. Project evaluations are tied to the budgets of project implementation, and they are conducted just as projects close or before. This may be adequate to assess the quality of implementation but it can tell us little about any benefits of projects. Project assumptions remain untested and are thus repeated, whether they are generally applicable or not. Methodological challenges are not a barrier to learning about what happened to projects after they closed. Conducting evaluations some three to five years after project closure should be standard practice.

Retrospective study is possible, with resourcing

This study shows that it is possible to assess impact retrospectively. It is important and it can be done – but it needs to be resourced.

Rigour, not just theory, can determine research methodologies

The study did not set out with a pre-determined toolkit of methods to assess impact. Rigour was an approach – a state of mind, not a tool. Such rigour was only possible because methodological flexibility allowed the adoption and adaptation of new tools to follow up in different ways on what was being learned. In a sense, the research methodology was demand-led, meaning that the tools used were determined by the need for information, rather than being chosen on theoretical grounds. This freedom, necessary for rigour and learning, is often denied to evaluators who have to prescribe their methodologies in advance. An evaluability assessment is useful, but it does not answer this requirement adequately.

Mixed-methods approaches make sense

In the study, qualitative and quantitative methods performed different roles and they were easily combined in a single analytical framework. In many cases, these methods were used to check each other, since they do not share the same risks of bias. Rigour applied equally to quantitative and qualitative tools. There is no reason not to combine the approaches seamlessly into most studies.

Evaluations must ask the most important questions

The questions that are asked determine the potential usefulness of any evaluation or research study. Impacts on people's lives should be the main focus of most evaluations, and the performance of assets should be a significant question in the domain of PWPs. This is not the case for most evaluations in the aid sector.

Theory-based assessment enables coherent analysis

The analytical framework provided by a theory-based approach was invaluable. It allowed for a wide variety of tools and discrete pieces of information to be fitted into a single coherent analysis. Because all the intermediate links between activities and livelihoods were examined, the 'black box' by which interventions work was opened (Pawson, 2007). Theory-based evaluation is well established, but in practice few evaluations and impact assessments explicitly test each link in the causal chain.

6.3 Recommendations for future practice

1. All organisations funding, supervising or implementing PWPs should satisfy themselves that they have adequate information on the benefits brought by the assets in all programmes that they have completed. If not, they should ensure that systems are put in place rapidly to ensure that current and future PWPs will collect such information.

2. These same organisations should check their design, implementation, and monitoring and evaluation systems to ensure that the failings identified in this study cannot arise in their PWP. Programmes should be designed and implemented in a way that prioritises the assets, and not simply so that they are created to absorb labour and in the vague hope that they will provide benefits.

3. Assumptions about future livelihood benefits should be made explicit when PWPs are designed. Evidence should be sought to justify these assumptions *before* the design phase is completed.

4. The benefits of any assets created through a PWP should be included in monitoring and reporting, with monitoring continuing after the completion of asset construction. If the creation of assets is deemed important enough to demand a labour contribution from those receiving cash or food transfers, and important enough to consume significant budgets, then the assets should be deemed important enough to be given equal priority in evaluations as the process and impacts related to wage transfers.

5. It should be standard practice for impact evaluations to be conducted several years after the completion of development or resilience-building programmes, including PWPs. These evaluations would be for lesson learning, not (upwards) accountability, and those lessons should be shared widely. These studies should be resourced as a common good for the sector.

6. Rigour should be demanded of these retrospective studies. Organisations have to examine the incentive structures that make good and honest learning harder, and they have to be explicit in discussing what can be done to reduce the incentives that discourage such learning.

7. Learning should be seen as a public good for the sector. Collaborative learning efforts are needed, including in setting the terms of reference for individual agency evaluations and in their review.

REFERENCES

Abebe, M.H., Oba, G., Angassa, A. and Weladji, R.B. (2006) 'The role of area enclosures and fallow age in the restoration of plant diversity in northern Ethiopia' *African Journal of Ecology* 44: 507–514.

Adimassu, Z., Mekonnen, K., Yirga, C. and Kessler, A. (2012) 'Effects of soil bunds on runoff, soil and nutrient losses, and crop yield in the Central Highlands of Ethiopia' *Land Degradation and Development* 25(6).

Adimassu, Z., Langan, S., Johnston, R., Mekuria, W. and Amede, T. (2017) 'Impacts of soil and water conservation practices on crop yield, run-off, soil loss and nutrient loss in Ethiopia: review and synthesis' *Environmental Management* 59: 87–101.

Anantha, K.H. et al. (2021) 'Impact of natural resources management interventions on water resources and environmental services in different agroecological regions of India' *Groundwater for Sustainable Development* 13.

Angassa, A. and Oba, G. (2010) 'Effects of grazing pressure, age of enclosures and seasonality on bush cover dynamics and vegetation composition in southern Ethiopia' *Journal of Arid Environments* 74: 111–120.

Asefa, D.T., Weladji, R.B. and Colman, J.E. (2003) 'An assessment of restoration of biodiversity in degraded high mountain grazing lands in northern Ethiopia' *Land Degradation and Development* 14: 25–38.

Aune, J.B., Asrat, R., Teklehaimanot, A.D. and Bune, B.T. (2006) 'Zero tillage or reduced tillage: the key to intensification of the crop-livestock system in Ethiopia.' In: Pender, J., Place, F. and Ehui, S. (eds) *Strategies for sustainable land management in the East African Highlands*. Washington DC: International Food Policy Research Institute.

Aynekulu, E., Denich, F. and Tsegaye, D. (2009) 'Regeneration response of Juniperus procera and Olea europaea subsp cuspidate to exclosure in a dry afromontane forest in northern Ethiopia', *Mountain Research and Development* 29(2): 143–152.

Beierl, S. and Grimm, M. (2019) *Do public works programmes work? A systematic review of the evidence in Africa and the MENA region.* Bonn: GIZ.

Béné, C., Cannon, T., Davies, M., Newsham, A. and Tanner, T. (2014) *Social protection and climate change*. Development Cooperation Working Paper 16. Paris: Organisation for Economic Co-operation and Development.

Bird, N., Manuel, M. and Nakhooda, S. (2015) *Scaling up international support for adaptation: productive safety nets and reimbursable debt service*. ODI Working Paper 427. London: ODI.

Burns, J. and Bogale, S. (2012) *Impact assessment of honey microfinance and livestock value chain interventions: final impact assessment of the PSNP Plus in Sekota*. Washington DC: United States Agency for International Development.

Cartwright, N. (2007) 'Are RCTs the gold standard?' *Biosocieties* 2(1): 11–20.

Copestake, J.G. (2014) 'Credible impact evaluation in complex contexts: confirmatory and exploratory approaches' *Evaluation* 20(4).

Costella, C. and McCord, A. (2023) *Rethinking social protection and climate change: implications of climate change for social protection policy and programming in the Asia-Pacific region.* Canberra, Australia: Department of Foreign Affairs and Trade, Government of Australia.

Dagg, M., Woodhead, T. and Rijks, D.A. (1970) 'Evaporation in East Africa' *International Association of Scientific Hydrology* 15(1): 61–67.

Damene, S., Tamene, L. and Vlek, P.L.G. (2013) 'Performance of exclosure in restoring soil fertility: a case of Gubalafto district in North Wello Zone, northern highlands of Ethiopia' *Catena* 101: 136–142.

Descheemaeker, K., Nyssen, J., Rossi, J., Poesen, J., Haile, M., Moeyersons, J. and Deckers, J. (2006a) 'Sediment deposition and pedogenesis in exclosures in the Tigray highlands, Ethiopia' *Geoderma* 132: 291–314.

Descheemaeker, K., Muys, B., Nyssen, J., Poesen, J., Raes, D., Haile, M. and Deckers, J. (2006b) 'Litter production and organic matter accumulation in exclosures of the Tigray highlands, Ethiopia' *Forest Ecology and Management* 233: 21–35.

Devereux, S., Roelen, K., Béné, C., Chopra, D., Leavy, J. and McGregor, A. (2013) *Evaluating outside the box: an alternative framework for analysing social protection programmes.* IDS Working Paper 431, CSP Working Paper 010. Brighton: Institute of Development Studies at the University of Sussex.
Donaldson, S. (2007) *Program theory-driven evaluation science: strategies and applications.* Thousand Oaks: Sage.
Filipski, M., Taylor, E.J., Abegay, G.A., Taffesse, A.S. and Diao, X. (2017) *General equilibrium impact assessment of the Productive Safety Net Program in Ethiopia.* 3ie Impact Evaluation Report 66. New Delhi: International Initiative for Impact Evaluation.
Funnell, S.C. and Rogers, P.J. (2011) *Purposeful program theory: effective use of theories of change and logic models.* San Francisco, CA: Jossey-Bass/Wiley.
Garbarino, S. and Holland, J. (2009) *Quantitative and qualitative methods in impact evaluation and measuring results.* Discussion Paper. Birmingham: University of Birmingham.
Gazeaud, J. and Stéphane, V. (2020) 'Productive workfare? Evidence from Ethiopia's productive safety net program*' American Journal of Agricultural Economics* 105(1): 265–290.
Gehrke, E. and Hartwig, R. (2018) 'Productive effects of public works programs: What do we know? What should we know?' *World Development* 107: 111–124.
Getachew, M. (2014) 'Vegetation dynamics of area enclosure practices: a case of Gonder Zuria district, Amhara region, Ethiopia' *Journal of Natural Sciences Research* 4(7).
Herweg, K. and Ludi, E. (1999) 'The performance of selected soil and water conservation measures—case studies from Ethiopia and Eritrea' *Catena* 36: 99–114.
Himmelstine, C.L. and McCord, A. (2016) *What is known about the impact of assets for natural resource management created through Public Works Programmes on livelihoods – a literature review*. London: ODI.
Hirvonen, K., Machado, E.A., Simons, A.M. and Taraz, V., (2022) 'More than a safety net: Ethiopia's flagship public works program increases tree cover' *Global Environmental Change* 75.
Hongoi, A., Matsumoto, S., Takahashi, H., Zou, H., Cheng, J., Jia, H. and Zhao, Z. (1995) 'Effect of exclosure and topography on rehabilitation of overgrazed shrub-steppe in the loess plateau of northwest China' *Restoration Ecology* 3(1): 18–25.
Kanbur, R. (2001) 'Q-squared? A commentary on qualitative and quantitative poverty appraisal'. In Kanbur, R. (ed.) *Q-squared: qualitative and quantitative methods of poverty appraisal.* New Delhi: Permanent Black.
Kimani, M.W., Gitau, A.N. and Ndunge, D. (2015) 'Rainwater harvesting technologies in Makueni county', *International Journal of Engineering and Science* 5(2): February.
KNBS – Kenya National Bureau of Statistics (2021) *The Kenya poverty report.* Nairobi: KNBS.
Ludi, E. (2004) *Economic analysis of soil conservation: case studies from the highlands of Amhara region, Ethiopia.* African Studies Series A18. Bern: Centre for Development and Environment.
McCord, A. (2012) *Public works and social protection in Southern Africa: do public works work for the poor?* Helsinki: Juta Press and UNU.
McCord, A. and Farrington, J. (2008) 'Digging holes and filling them in again? How far do public works enhance livelihoods?' *Natural Resource Perspectives* 120. London: ODI.
Mekuria, W., Veldkamp, E., Haile, M., Nyssen, J., Muys, B. and Gebrehiwot, K. (2007) 'Effectiveness of exclosures to restore degraded soils as a result of overgrazing in Tigray, Ethiopia' *Journal of Arid Environments* 69: 270–284.
Mengistu, T., Teketay, D., Hulten, H. and Yemshaw, Y. (2005a) 'The role of enclosures in the recovery of woody vegetation in degraded dryland hillsides of central and northern Ethiopia' *Journal of Arid Environments* 60: 259–281.
Mengistu, T., Teketay, D., Hultén, H. and Yemshaw, Y. (2005b) 'The role of communities in closed area management in Ethiopia' *Mountain Research and Development* 25(1): 44–50.
Mulugeta, G. (2014) 'Vegetation dynamics of area enclosure practices: a case of Gonder Zuria District, Amhara Region, Ethiopia' *Journal of Natural Sciences Research* 4(7).
Nagaraj, N. et al. (2018) 'Economic Impacts of MGNREGA in Dryland Region of India: A Meso and Micro Study in Selected States' in M. Bhattarai et al. (eds), *Employment Guarantee Programme and Dynamics of Rural Transformation in India*. India Studies in Business and Economics.
Normand, M. (2008) 'Science, skepticism, and applied behavior analysis' *Behavior Analysis in Practice* 1(2): 42–49.

Patel, N. (unpublished) *What is known about the impacts on livelihoods of assets for natural resource management created through Public Works Programmes? A review of literature since 2016.*
Pawson, R. (2007) *Causality for beginners.* NCRM EPrints Repository. Southampton: National Centre for Research Methods.
Scriven, M. (1991) 'Pros and cons about goal-free evaluation' *Evaluation Practice* 12(1): 55–62.
Steinbach, D., Soanes, M., Barrett, S., Venkataramani, V. and Cheranda, T.M. (2020) *Deepening knowledge of MGNREGS' contribution to climate resilience: a study of Rajasthan and Uttar Pradesh*. London: IIED.
Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R. and Befani, B. (2012) *Broadening the range of designs and methods for impact evaluations.* DFID Working Paper 38. London: UK Department for International Development.
SuDCA and Soberland Consulting Associates Joint Venture (2015) *Productive Safety Net Programme, phase III 2014: public works impact assessment – final report.* Addis Ababa: Ministry of Agriculture.
Tekele, K. (2001) 'Natural regeneration of degraded hillslopes in southern Wello, Ethiopia: a study based on permanent plots' *Applied Geography* 21: 275–300.
Thakur, S (2018) 'Relevance of MGNREGA Assets' *Journal of Rural Development*, Vol. 37(1): 21–32.
UNDP (2012) Coping with drought and climate change in Ethiopia. Addis Ababa: UNDP.
van de Mortel, T.F. (2008) 'Faking it: social desirability response bias in self-report research' *Australian Journal of Advanced Nursing* 25(4): 40–48.
Verdoodt, A., Mureithi, S.M., Ye, L. and van Ranst, E. (2009) 'Chrono-sequence analysis of two enclosure management strategies in degraded rangeland of semi-arid Kenya' *Agriculture, Ecosystems and Environment* 129: 332–339.
Verdoodt, A., Mureithi, S.M. and van Ranst, E. (2010) 'Impacts of management and enclosure age on recovery of the herbaceous rangeland vegetation in semi-arid Kenya' *Journal of Arid Environments* 74: 1066–1073.
WFP –World Food Programme (2010) *Protracted Relief and Recovery Operation (PRRO) recommendations for an institutional and operational framework for the implementation of PRRO 10666-K': final report.* Nairobi: WFP.
WFP (2012) *PRRO 10666, protecting and rebuilding livelihoods in the arid and semi-arid areas, Makueni county FFA programme.* Final report. Nairobi: WFP.
WFP (2016) Evaluation of the DFID funded "Enhancing Resilience in Karamoja Programme." Kampala: WFP.
White, H. (2009) *Theory-based impact evaluation: principles and practice.* Working Paper 3. New Delhi: International Initiative for Impact Evaluation.
White, H. and Phillips, D. (2012) *Addressing attribution of cause and effect in small impact evaluations: towards an integrated framework* New Delhi: International Initiative for Impact Evaluation.
Woodhead, T. (1968) Studies of Potential Evaporation in Kenya. Nairobi: East African Agriculture and Forestry Research Organisation.
WVI (2014) 'Musisi dam success story'. Nairobi: WVI.
Yayneshet, T., Eik, L.O. and Moe, S.R. (2009) 'The effects of exclosures in restoring degraded semiarid vegetation in communal grazing lands in northern Ethiopia' *Journal of Arid Environments* 73: 542–549.
Yirdaw, E. and Monge, A. (2017) 'Reconsidering what enclosure and exclosure mean in restoration ecology' *Restoration Ecology* 26(1): 45–47.
Youker, B.W. (2013) 'Goal-free evaluation: a potential model for the evaluation of social work programs' *Social Work Research* 37(40): 432–438.

APPENDIX: THE STUDY SITES

Ethiopia

Kalu District (or *woreda*) in South Wollo Zone is highly degraded, drought-prone and chronically food-insecure. At the time of the study in 2016, a wide range of PSNP interventions had been carried out there over many years (UNDP, 2012). Kalu has 30 rural and four urban *kebeles*, which are further subdivided into what are called 'watersheds'. These watersheds are based on administrative units and are not defined entirely by topography. Approximately 200 watersheds exist in the district (although there are discrepancies in the district's documentation regarding the exact number), with the number in each *kebele* ranging from three to six. Two sets of paired watersheds were selected: Guba and Kono as the treated areas, and Menekuse and Barentu as the matching controls without PSNP assets for NRM.

Kono and Barentu are located about 5 km from Habru town at an altitude of approximately 1,500 m above sea level. Their topography is characterised by steep hillsides rising above flat-to-medium steep slopes primarily used for crop cultivation. The predominant activity of inhabitants is mixed agriculture, combined with livestock-rearing. The main crops cultivated include sorghum, teff, wheat, oats and legumes. Cattle, sheep, goats and donkeys are commonly owned.

Guba and Menekuse are located about 12 km from Habru town towards the north-east at an altitude of approximately 2,000 m above sea level. The terrain is characterised by steep hillsides and medium-to-flat arable land. Agriculture is mixed with crop and livestock production. The main crops cultivated are sorghum and teff and some legumes. Again, cattle, sheep, goats and donkeys are commonly owned.

Kenya

Makueni County in the south-east of Kenya has unreliable rainfall and high temperatures, leading to serious water and food security challenges. Maize is the main staple food in the research area, together with cow peas, although both government agencies and NGOs have promoted diversification and the production of drought-resistant crops such as sorghum. The county has two rainy seasons, with the long rains occurring from mid-March to May and the short rains from October to December. Conditions are not fully reliable for rain-fed farming.

The rains in the year preceding dam construction (2009) were particularly bad, with a prolonged drought across the Horn of Africa, leading to crop failure and distress sales of livestock. Population growth in Makueni has resulted in increased pressure on land with limited options for non-agricultural livelihoods and employment. Landholdings are becoming smaller as they are subdivided between family members.

There are a number of seasonal rivers and streams, notably the Thwake and the Kikuo, but the River Athi is the only permanent water source and the major one when the other rivers and streams are dry. Water is scooped from holes dug into the bed of the Thwake River during the dry season, although it is no longer easy to find water near the surface as the level of ground water is dropping.

WFP's project worked at both household level, including technology transfer for SWC and terracing, and on public works of various kinds. This study only looked at dam construction for public asset creation. Some villages already had earth dams to harvest rainwater, and boreholes sponsored by various government agencies and NGOs. Under the project, existing dams were rehabilitated and new dams were constructed in an effort to increase the number of months during which water is available (WFP, 2010).

The project in Makueni County was selected in conjunction with WFP as an example of a successful PWP that had included explicit objectives relating to longer-term livelihoods improvements. The study focused on dams in what was then North Makueni District because the quality of available documentation by the implementing partners (World Vision International) was felt to be better.

Dams were intended to provide water for consumption and for livestock and for other income-generating activities, including horticulture. Livestock are widely owned in Makueni, though ownership is increasingly concentrated in the hands of the better off. The anticipated outcomes of asset creation were stated to be (WFP, 2012):

- improved access to water for human and livestock consumption
- improved pasture for livestock feed
- increased crop production
- reduced environmental degradation.

The study looked in depth at dams near five case study villages: Syotuvali, Kavumbu, Kiangini, Kithuki and Kavingon.

@SPARC_Ideas

sparc-knowledge.org

Cover image: Two problems, one solution – or is it more complex?
© Julie Grady Thomas

Funded by

UKaid
from the British people

This material has been funded by UK aid from the UK government; however the views expressed do not necessarily reflect the UK government's official policies.