
The Department for International Development

MILLENNIUM VILLAGES EVALUATION: MIDTERM SUMMARY REPORT

Date: **November 2016**

Authors: **Edoardo Masset, Dee Jupp, David Korboe, Tony Dogbe, Chris Barnett, Arnab Acharya and Kelsy Nelson**

Submitted by Itad
In association with:



Acknowledgements

This report has been prepared by the team for the impact evaluation of the Millennium Villages Project. The team is composed of staff from Itad, the Institute of Development Studies, the London School of Hygiene and Tropical Medicine, and Participatory Development Associates Ltd (PDA-Ghana). The team is fully independent of the Earth Institute and the Millennium Promise. The authors of this report are Edoardo Masset, Tony Dogbe, Dee Jupp, David Korboe, Chris Barnett, Arnab Acharya and Kelsy Nelson. The report was edited and proofread by Caitlin McCann, Chris Steele and ETTY Payne.

The findings of this report are the full responsibility of the authors, and any views contained in it do not necessarily represent those of the UK Department for International Development or of the people consulted. The team is nonetheless very grateful to all the researchers that have assisted with data collection, the staff at DFID, and everyone else that has provided support, information and comments – including the work of the Earth Institute during the enumeration phase.

Citation

Masset, E., Jupp, D., Korboe, D., Dogbe, T., Barnett, C., Acharya, A. & Nelson, K. 2016. *Millennium Villages Evaluation: Midterm Summary Report*. Itad, Hove.

Disclaimer

This report has been prepared by Itad for the named client, for services specified in the Terms of Reference and contract of engagement. The information contained in this report shall not be disclosed to any other party, or used or disclosed in whole or in part without agreement from Itad. For reports that are formally put into the public domain, any use of the information in this report should include a citation that acknowledges Itad as the author of the report.

This confidentiality clause applies to all pages and information included in this report.

Acronyms and abbreviations

ANCOVA	Analysis of Covariance
CBO	Community-Based Organisation
CF	Faraway Control Village
CHW	Community Health Worker
CLTS	Community led Total Sanitation
CN	Nearby Control Village
CV	Control Village
DACF	District Assemblies Common Fund
DD	Difference-in-Difference
DFID	UK's Department for International Development
DHS	Demographic and Health Survey
G-PASS	Girls–Participatory Approach to Student Success in Ghana
FBO	Farmer Based Organisation
GPE	Global Partnership for Education
GPS	Global Positioning System
HH	Household
IPW	Inverse Probability Weights
JSS	Junior Secondary School
LEAP	Livelihood Empowerment Against Poverty
MDG	Millennium Development Goals
MTDP	Medium-Term Development Plans
MV	Millennium Village
MVP	Millennium Villages Project
NGO	Non-Governmental Organisation
NHIS	National Health Insurance Scheme
PRA	Participatory Rural Appraisal
PTA	Parent Teacher Association
RCA	Reality Check Approach
SADA	Savannah Accelerated Development Authority
SDGs	Sustainable Development Goals
SSS	Senior Secondary School
UN	United Nations
VSLA	Village Savings and Loans Association
WASH	Water Sanitation and Hygiene
WHO	World Health Organization

Contents

ACKNOWLEDGEMENTS.....	2
ACRONYMS AND ABBREVIATIONS.....	3
EXECUTIVE SUMMARY	7
CHAPTER 1. INTRODUCTION.....	9
1.1 THE MILLENNIUM VILLAGES PROJECT	9
1.2 BRIEF OVERVIEW OF THE EVALUATION DESIGN	11
1.3 AIMS AND STRUCTURE OF THE REPORT	12
CHAPTER 2. METHODOLOGY.....	14
2.1 THE EVALUATION DESIGN	14
2.2 METHODOLOGY FOR ESTIMATING DIFFERENCE-IN-DIFFERENCE ANALYSIS.....	16
2.3 DATA COLLECTION AND SAMPLE CHARACTERISTICS.....	19
2.4 INSTITUTIONAL ASSESSMENT	23
2.5 PARTICIPATORY RURAL APPRAISAL.....	23
2.6 REALITY CHECK APPROACH.....	24
CHAPTER 3. MVP IMPLEMENTATION.....	26
3.1 CONTEXT: ACTIVITY PROGRESS REPORTED BY MVP	26
3.2 MILLENNIUM VILLAGE PROJECT PROGRESS IN INFRASTRUCTURE	29
3.3 INVOLVEMENT IN PROJECT ACTIVITIES	33
3.4 PEOPLE’S EXPERIENCE AND VIEW OF INVOLVEMENT IN MVP ACTIVITIES.....	37
CHAPTER 4. ACHIEVEMENT OF THE MDGS	40
4.1 IMPACT OF MVP ON THE MILLENNIUM DEVELOPMENT GOALS.....	40
CHAPTER 5. IMPACT	49
5.1 IMPACT OF MVP ON MONETARY POVERTY.....	49
5.2 IMPACT OF MVP ON INCOME AND FOOD SECURITY	55
5.3. IMPACT OF MVP ON CHILD HEALTH.....	62
5.4 IMPACT OF MVP ON ANTHROPOMETRY.....	64
5.5 IMPACT OF MVP ON ANAEMIA OF UNDER-5 CHILDREN.....	65
5.6 IMPACT OF MVP ON EDUCATION	66
5.7 IMPACT HETEROGENEITY	70
CHAPTER 6. SPILL-OVER EFFECTS	74
6.1 GEOGRAPHIC SPILL-OVER EFFECTS.....	74
6.2 INSTITUTIONAL SPREAD OF MVP PRACTICES.....	76
CHAPTER 7. INSTITUTIONAL SUSTAINABILITY	77
7.1 SUSTAINING SADA-MVP’S MODEL OF COMMUNITY WORKERS AND ALLOWANCES.....	77
7.2 SUSTAINING INTERVENTIONS AFTER THE PROJECT.....	78
7.3 COMMUNITY CAPACITY TO SUSTAIN INTERVENTIONS AND IMPACT	79
7.4 RELATIONSHIPS WITH CITIZENS’ GROUPS, NGOs AND FBOs	79
7.5 LIMITED FISCAL DECENTRALISATION AND INCONSISTENT AND INADEQUATE REVENUE	80
CHAPTER 8. OVERALL FINDINGS AND CONCLUSIONS	82
BIBLIOGRAPHY.....	85

List of tables

TABLE 1. SURVEY INSTRUMENTS USED DURING 2014	15
TABLE 2. COMPLETED HOUSEHOLD INTERVIEWS DURING THE FIRST THREE ROUNDS	20
TABLE 3. REASONS FOR NOT COMPLETING THE INTERVIEWS.....	20
TABLE 4. COMPLETED HOUSEHOLD INTERVIEWS DURING THE FIRST THREE ROUNDS	21
TABLE 5. INDIVIDUALS LISTED IN THE SURVEYS	21
TABLE 6. PERCENT OF POPULATION CHANGES ACROSS SURVEYS.....	22
TABLE 7. ACTIVITIES AND OUTPUTS IN THE AGRICULTURE SECTOR	27
TABLE 8. ACTIVITIES AND OUTPUTS IN THE EDUCATION SECTOR	27
TABLE 9. ACTIVITIES AND OUTPUTS IN THE HEALTH SECTOR.....	28
TABLE 10. ACTIVITIES AND OUTPUTS IN THE INFRASTRUCTURE, WATER AND SANITATION SECTORS	28
TABLE 11. ACTIVITIES AND OUTPUTS IN COMMUNITY ENGAGEMENT AND GOVERNANCE	29
TABLE 12. SUMMARIES OF VILLAGE FACILITIES, 2012	30
TABLE 13. HEALTH RESOURCES IN 2012.....	30
TABLE 14. SUMMARIES OF VILLAGE FACILITIES, 2012 AND 2014.....	31
TABLE 15. WATER SOURCES, 2012 AND 2014	31
TABLE 16. EDUCATIONAL RESOURCES IN 2012 AND 2014.....	32
TABLE 17. HEALTH RESOURCES 2014.....	32
TABLE 18. MEMBERSHIP OF COMMUNITY-BASED GROUPS (MIDTERM).....	34
TABLE 19. INVOLVEMENT IN AGRICULTURE-RELATED ACTIVITIES (MIDTERM)	34
TABLE 20. INVOLVEMENT IN HEALTH-RELATED ACTIVITIES (MIDTERM).....	34
TABLE 21. INVOLVEMENT IN EDUCATION-RELATED ACTIVITIES (MIDTERM).....	35
TABLE 22. F-TESTS OF EQUALITY IN PROJECT INVOLVEMENT PATTERNS (MIDTERM).....	37
TABLE 23. PROJECT IMPACT ON THE MDGs	41
TABLE 24. POVERTY INDICES	49
TABLE 25. WELL-BEING: CHANGES IN RANKINGS SINCE BASELINE	51
TABLE 26. DIFFERENCE-IN-DIFFERENCE IMPACT ON PER-ADULT EQUIVALENT EXPENDITURE	52
TABLE 27. DISTRIBUTIONAL IMPACT ON PER-ADULT EQUIVALENT EXPENDITURE.....	54
TABLE 28. INCOME SHARES BY SOURCE, ALL HOUSEHOLDS	55
TABLE 29. DIFFERENCE-IN-DIFFERENCE IMPACT ON PER CAPITA INCOME (IPW METHOD).....	56
TABLE 30. IMPACT OF MV ON DIFFERENT INCOME SOURCES (LAGGED MODEL).....	57
TABLE 31. DISTRIBUTIONAL IMPACT ON HOUSEHOLD INCOME (AVERAGE DD EFFECTS)	59
TABLE 32. PROJECT IMPACT ON FOOD SECURITY PERCEPTIONS (IPW METHOD).....	60
TABLE 33. PERCEPTIONS OF FOOD SECURITY.....	61
TABLE 34. DD ANALYSIS (FIVE-YEAR INTERVAL)	63
TABLE 35. IMPACT OF THE MVP ON NUTRITIONAL STATUS OF CHILDREN UNDER 5 (IPW METHOD)	64
TABLE 36. IMPACT OF THE MVP ON NUTRITIONAL STATUS OF CHILDREN UNDER-5 (PANEL CHILDREN).....	65
TABLE 37. IMPACT OF THE MVP ON HAEMOGLOBIN CONCENTRATION AND PREVALENCE OF ANAEMIA.....	66
TABLE 38. IMPACT OF THE MVP ON SCHOOL ATTENDANCE.....	67
TABLE 39. IMPACT OF THE MVP ON SCHOOL ATTENDANCE (PANEL CHILDREN)	67
TABLE 40. IMPACT OF THE MVP ON TEST SCORES.....	69
TABLE 41. IMPACT OF THE MVP ON TEST SCORES (PANEL CHILDREN)	70
TABLE 42. IMPACT OF THE MVP BY DISTRICT	71
TABLE 43. IMPACT OF THE MVP BY GENDER (BOYS AND GIRLS)	73
TABLE 44. IMPACT OF MV ON INCOME BY DISTANCE	75
TABLE 45. IMPACT OF MV ON NUTRITIONAL STATUS BY DISTANCE.....	76

List of figures

FIGURE 1. SUMMARY OF DATA COLLECTION BY YEAR.....	16
FIGURE 2. CHANGE IN HOUSEHOLD SIZE BETWEEN BASELINE AND MIDTERM	22
FIGURE 3. INVOLVEMENT IN PROJECT ACTIVITIES BY HOUSEHOLD EXPENDITURE (MIDTERM)	36
FIGURE 4. PROJECT IMPACT WITH BASELINES AT DIFFERENT TIMES IN PROJECT AND CONTROL AREAS	47
FIGURE 5. DENSITIES OF LOG PER-ADULT EQUIVALENT EXPENDITURE IN THE PROJECT AND CONTROL GROUPS	53
FIGURE 6. PER-ADULT EQUIVALENT EXPENDITURE BY PREDICTED EXPENDITURE	54
FIGURE 7. DENSITIES OF HOUSEHOLD INCOME IN THE PROJECT AND CONTROL GROUPS	58
FIGURE 8. PER CAPITA INCOME BY PREDICTED PER CAPITA INCOME.....	59
FIGURE 9. INVOLVEMENT IN PROJECT ACTIVITIES BY HOUSEHOLD EXPENDITURE	62
FIGURE 10. PROJECT AND COMPARISON COMMUNITIES.....	74

Executive summary

- E1. This report presents the midterm findings from an impact evaluation of the Millennium Villages Project (MVP) in northern Ghana. The MVP has been designed to demonstrate how an integrated approach to community-led development can translate the international Millennium Development Goals (MDGs) into results. The project in northern Ghana is one of several instigated over the past 10 years, and is set to reach nearly half a million people across 10 countries in Africa. Central to the MVP approach is the synergistic value of *integrated community-based investments*, focused on scientifically proven interventions, delivered simultaneously rather than as one-off investments. The premise is that a critical platform of basic needs must be reached before economic development can really take off. The project in northern Ghana has been running in three districts since May 2012, investing over £11 million on health, education, agriculture and infrastructure interventions in 35 communities, and reaching around 30,000 people.
- E2. There have been evaluations of the MVP by other researchers. Some have conducted evaluations using data on project and control villages collected by the Earth Institute (Remans et al. 2011, Pronyk et al. 2012), though serious doubts were raised about the validity of these estimates (Bump et al, 2012, and Clemens and Demombynes, 2012). Other studies cleverly exploited data from Demographic Health Surveys to assess the impact of the MVP (Clemens and Demombynes, 2011). Finally, one study used a combination of project data and matched survey data to assess the impact of the MVP in Kenya (Whanjala and Muradian, 2013). Our evaluation differs from previous impact assessments in four ways. First, data are collected from sizeable project and control groups before and after project implementation. Second, data are collected from repeated interviews with the same households and individuals (panel data). Third, data are collected on a wide range of welfare indicators covering all the MDGs. Fourth, the evaluation is complemented and validated by a concurrent qualitative assessment of the MVP's impact. These elements are unique to our study and make our evaluation the most statistically valid, comprehensive and qualitatively informed than what has been conducted to date.
- E3. This is the midterm report based on the third survey round of household data (2012, 2013 and 2014), complemented by additional qualitative studies (reality check approach (RCA), interpretational lens or participatory rural appraisal (PRA) study, institutional analysis). This evaluation uses a mixed methods approach to impact evaluation, founded on a difference-in-difference (DD) design that compares changes in outcomes in the MVP areas before implementation to post-implementation, with changes in the same outcomes for an explicit control group (split into nearby and faraway controls). It is important to note that the findings presented are not those of a full impact assessment, as the project had been running for two years at the time of the 2014 data collection. This is especially so as some of the activities in the first year were preparatory (e.g. consultations with farmers, feasibility studies), and these are unlikely to have contributed much to observable effects.
- E4. Overall, the quantitative analysis finds that the project had moderately positive impacts at the midterm. The project has not yet had a sizable impact on the MDGs but several impacts are visible on other indicators of well-being: (i) The project produced a large increase in per capita income, with an increase in agricultural incomes. This increase in incomes appears to be supported by a large improvement in self-reported food security. Household expenditure increased to a lower extent, and since poverty measures are based on household expenditure, we find little overall impact on poverty. (ii) There was a considerable improvement in the nutritional status of children under five. (iii) The project produced a modest reduction in mortality rates but we highlight the difficulty in calculating these effects after only two years of data. The project did not affect the

prevalence of anaemia. (iv) The project produced only a modest increase in school attendance rates and no improvements on the quality of education as measured by mathematics and English test scores. (v) There are some signs that the project is having a higher impact in the Builsa district compared with the other districts. (vi) We found no difference in impacts by sex.

- E5. For the most part, the PRA study findings align with those of the statistical analysis, with the MVP communities appearing to show some marginal movements in (self-perceived) well-being, and no substantive reduction in poverty. The project seems to be demonstrating initial gains in areas such as school enrolment, teacher attendance and pupil attendance/retention; attended births, emergency care and malarial morbidity; crop yields, livestock health and food security.
- E6. The RCA study, among other things, also highlights how people's aspirations to leave farming and their need for reliable cash incomes calls into question some of the underlying assumptions behind the MVP's means of addressing poverty. People are increasingly concerned about the unpredictability of the climate (particularly, late rains), high costs of inputs, declining soil fertility and risks; they also talk about the increasing need for cash – in contrast to a largely cashless situation observed in 2013. The provision of infrastructure and services is questioned as being insufficient to change behaviour and to address some deeply held cultural and traditional practices, some of which may in fact have a good basis for being preserved. Based on these observations, there seems to be insufficient attention to context in the design of the MVP interventions and the different effects of the same intervention in different situations. The institutional assessment further supports the view that the MVP approach to working with communities has been less empowering and transformative than might have been hoped.
- E7. Our exploratory analysis is also suggestive of potential spill-over effects to neighbouring communities in relation to income indicators but not in relation to nutritional status – based on the statistical analysis. The PRA and RCA studies show that the expected spill-over effects predicted by the MVP's programme theory do not appear to be strong thus far. In several areas, the faraway controls are ahead of the nearby controls. Indeed, the 'control' villages have attracted attention from non-MVP development organisations and political patronage, and, as the RCA study indicates, two 'control' villages seem poised to become quite successful as a result.
- E8. In terms of the institutional assessment, district officials in the three project districts appreciate what has been achieved so far under MVP, but at the same time they are concerned about the sustainability of the various initiatives when the project comes to an end. They are especially concerned about how top-up allowances to community health nurses, midwives and medical assistants, as well as payments to community health/education/livestock workers (who previously were willing to be volunteers) will be sustained. This could have a consequential effect on the future expectations of community workers, the level of community spiritedness and damage to positive traditions of communal labour and volunteering. Furthermore, the district institutions have inadequate and inconsistent revenue and are themselves somewhat dependent on the Savannah Accelerated Development Authority (SADA)-MVP for funds to fuel their vehicles for monitoring and support visits to their field staff.

Chapter 1. Introduction

1. This report presents the midterm findings from an impact evaluation of the Millennium Villages Project (MVP) in northern Ghana, as commissioned by the UK Department for International Development (DFID).¹ This is the first Millennium Village (MV) to be accompanied by an independent impact evaluation. The MVP in northern Ghana runs from May 2012 until December 2016, with interventions targeting a cluster of 35 communities. The MVP has been designed to demonstrate how an integrated approach to community-led development can translate the international Millennium Development Goals (MDGs) into results. It is an approach that has been previously piloted in Kenya and Ethiopia, and in 2006 launched at scale to reach nearly half a million people across 10 countries throughout sub-Saharan Africa.
2. This report is based on the third survey round of household data collection (2012, 2013 and 2014) complemented by additional data and qualitative studies that follow up on the baseline assessment. The report also builds on a series of publicly available reports and datasets, which include:
 - **The Initial Design Document**,² which sets out the detailed evaluation design.
 - **The Baseline Report**,³ which analyses the 2012 household survey, plus the qualitative research (institutional assessment, reality check, PRA assessments).
 - **The 2012 household dataset**,⁴ which is available for re-analysis (by request from the UK Data Archive).
 - **The Year 2 Report**,⁵ which analyses the 2013 household survey – a reduced instrument that focuses on a narrower sub-set of modules (e.g. demographics, consumption, expenditure).
 - **The Analysis Plan**,⁶ which sets out in detail how the evaluation team intends to analyse the main quantitative datasets in the final evaluation report.

1.1 The Millennium Villages Project

3. At the UN Millennium Summit in September 2000, world leaders adopted the Millennium Declaration, committing nations to a new global partnership to reduce extreme poverty and address pressing challenges of hunger, gender inequality, illiteracy and disease. Targets were set for these MDGs to be achieved by 2015. The MVP was created to explicitly demonstrate how the MDGs could be achieved at the local level, using an integrated and scaled-up set of targeted investments based on the recommendations of the United Nations Millennium Project. The interventions within MVP cover food production, nutrition, education, health services, roads, energy, communications, water supply and sanitation, enterprise diversification, environmental management and business development. The initiative makes use of community decision making and uses science-based technologies and techniques, such as agroforestry, insecticide-treated malaria bed nets, antiretroviral drugs, remote sensing, and geographic information systems. The MVP is a 10-year initiative with two five-year phases.

¹ The Terms of Reference for the assignment are included in the Initial Design Document, Appendix A.

² <http://www.ids.ac.uk/publication/an-impact-evaluation-design-for-the-millennium-village-project-in-northern-ghana>

³ <http://www.ids.ac.uk/publication/millennium-villages-impact-evaluation-baseline-summary-report>

⁴ <http://discover.ukdataservice.ac.uk/catalogue?sn=7734>

⁵ <http://www.ids.ac.uk/publication/northern-ghana-millennium-villages-impact-evaluation-preliminary-report-on-the-second-round-of-data>

⁶ <http://www.ids.ac.uk/publication/northern-ghana-millennium-villages-impact-evaluation-analysis-plan>

4. The *first phase* focused on achieving quick wins,⁷ especially in staple crop production and disease control, and on establishing basic systems for integrated rural development that help communities escape the poverty trap and achieve the MDGs. The interventions are delivered at a cost of approximately \$120 per capita per year of which the MVP contributes around half to complement funds from the host government, the local community and other partners. The *quick wins* are complemented by infrastructural improvements (buildings, roads, energy, water and sanitation), which are necessary to strengthen the primary health care and education services, as well as facilitate development of agricultural markets and business development. The revitalisation and strengthening of community institutions (local government, sector-based institutions, etc.) and the implementation of strategies to increase women's participation and leadership are also important facets of the MVP approach. The *second phase* focuses more intensively on commercialising the gains in agriculture and continuing to improve local service delivery systems to support the local scale-up.
5. The MVP was piloted in Kenya (Sauri) and Ethiopia (Koraro) in 2005 and launched at scale in 2006 to reach nearly half a million people across 10 countries. Central to the MVP approach is the value of *integrated community-based investments* in scientifically proven interventions, delivered simultaneously rather than as one-off investments. The premise is that a critical platform of basic needs must be reached before economic development can take off. The interventions are implemented in a multi-layered, multi-sectoral and integrated manner. The reasons underpinning this approach are rooted in the following explanations:
 - To address multiple objectives across a range of areas (health, poverty, disease control, nutrition, etc.) a holistic strategy with a range of tools (community-based clinics, diversified local food production, malaria control, etc.) is needed.
 - These multiple tools are synergistic, with each supporting a main objective but also contributing to progress on several or all of the goals.
6. As part of their second phase of implementation across Africa, the Millennium Promise Alliance (MPA) approached DFID to finance a Millennium Village in rural northern Ghana. The proposal was for the design, implementation and monitoring of a five-year set of integrated interventions to accelerate development in a cluster of communities of up to 30,000 people based on the MV model, with potential to be substantially scaled up. The UK provides £11 million over five years between 2012 and 2016 for the implementation of this new MV site in West Mamprusi, Mamprugu Moaduri and Builsa South districts of northern Ghana.⁸ By supporting a new MV along with this independent and rigorous evaluation, DFID and the semi-autonomous Savannah Accelerated Development Authority (SADA)⁹ aim to provide robust evaluation evidence on the effectiveness of the MV model and integrated rural development approaches as well as how to evaluate them in order to increase the effectiveness of future development policy and interventions. Previous evaluations have been mostly internal, undertaken by the Earth Institute and the MVP implementation team, and therefore not viewed as independent. Furthermore, they have been criticised for not being rigorous due to their reliance on before-and-after datasets¹⁰

⁷ "Quick wins" are defined as scientifically proven interventions that can achieve wide coverage and lead to dramatic gains in short periods of time. The quick win interventions are: (1) distribution of improved seeds and fertiliser; (2) distribution of long-lasting insecticide treated bed nets; (3) basic immunisations; (4) Vitamin A campaigns; and, (5) community wide de-worming to reduce levels of intestinal parasitosis.

⁸ At the start of the MVP in the SADA area at the beginning of 2012 there were two districts (West Mamprusi and Builsa). Over the course of 2012, they were both split into two, making a total of four districts. The MVP is implemented in three of these districts as stated here.

⁹ The SADA Strategy, 'A Sustainable Development Initiative for the Northern Savannah', emphasises mobilising and coordinating increased investment from public and private sources in order to stimulate private sector-led economic growth, along with careful monitoring and evaluation of development interventions in order to maximise impact.

¹⁰ For example, the Millennium Promise (2010) *Harvests of Development in Rural Africa: The Millennium Villages after Three Years*.

(often showing positive trends), without a sufficiently robust control group (to enable the measurement of the net effect in the MV site).¹¹

7. The stated aim of the MVP in northern Ghana is the achievement of the MDGs. The overall goal stated in the project's logical framework is that there will be a regional impact on poverty in the Northern / Upper East regions, with the proportion of the population living below the extreme poverty line reducing from 52.3% (Northern region) and 70% (Upper East region) to 33% overall by the end of 2016. The target for under-5 mortality rates is that they will fall from 124 deaths per 1,000 (Northern) and 98 deaths per 1,000 (Upper East), to 54 by the end of 2016. This is to be achieved through the "*accelerated progress towards the MDGs for up to 30,000 people in the MV site*".¹² Midterm progress against the MDG indicators is presented in Chapter 4, based on local level changes recorded in the project (treatment) and control sites. The implementation of MVP to date is set out in Chapter 3. This covers achievement against each of the projects outputs, infrastructure constructed or rehabilitated (e.g., health facilities, school classrooms, water points, electricity), and involvement of the local people in project activities (e.g. cooperative farmer groups).
8. The northern Ghana MVP does not have a single, overarching Theory of Change (ToC) underpinning the programme logic that outlines how the inputs-to-outputs-to-outcomes achievements ultimately result in MDG-level impacts. This is partly because of the complexity of the programme (multiple interventions designed to lead to multiple outcomes), but also because MVP adjusts its interventions each year and is not a fixed package. At the baseline, the evaluation team attempted to reconstruct the theory of change based on a series of detailed, generic "intervention logics" from the Earth Institute.¹³ While these "intervention logics" provide a basic understanding about the causal chains and assumptions in a select number of MVP sub-components, they do not accurately represent the specific intervention package that has been implemented in northern Ghana (plus they lack sufficient detail about how activities are sequenced and interlinked). Instead, the evaluation has focused on testing the overarching economic theory about the poverty trap and based on theoretical and empirical research that does exist, as outlined in the Initial Design Document. Alongside this, the evaluation team plans to work in two directions for the endline:¹⁴ first, by developing a few micro-level theories of change around interventions that are particularly relevant; and second, by sequencing the qualitative work to understand anomalous quantitative results with the goal of formulating hypotheses that can be tested with the available data.

1.2 Brief overview of the evaluation design

9. The evaluation uses a mixed methods approach to impact evaluation (Masset et al. 2013a; 2013b). At the core of the methodology is a difference-in-difference (DD) design that compares changes in outcomes in the MVP areas before implementation to post-implementation, with changes in the same outcomes for an explicit control group. DD allows the evaluation to isolate the MVP impact on outcomes (including poverty, child development, undernutrition and child mortality) from effects of other variables changing over time.

¹¹ This has been highlighted by several commentators of the MV approach to evaluation, such as Michael Clemens. See for example: "*Impact Evaluation in Aid What for, how rigorous?*", presentation for the Royal African Society and Overseas Development Institute (ODI), Center for Global Development, 3rd July 2012, London, UK.

¹² Based on the outcome objective of the DFID Logframe for the Millennium Village in Northern Ghana.

¹³ See the Initial Design Document, Appendix B. <http://www.ids.ac.uk/publication/an-impact-evaluation-design-for-the-millennium-village-project-in-northern-ghana>

¹⁴ As set out in the Analysis Plan (Masset 2015).

10. In preparing the design of this impact evaluation, a number of alternatives were considered. For instance, a randomised trial was, in principle, possible by randomly allocating the interventions to matched village pairs, but this was considered to be infeasible in this particular case. For a more detailed explanation, see Masset et al. (2013a). In the end, the matching of control villages to project villages (on aggregate characteristics) and further matching of project and control households at the analysis stage (on household characteristics) within a DD approach was considered to be the next best feasible approach after a randomised design.
11. Alongside the quantitative survey data used for the DD analysis, there are a number of supporting qualitative studies that contribute to the evaluation team's understanding about how and why change has occurred – including for the most marginalised in society. In addition, there are four key qualitative modules that complement the DD analysis: (i) a **poverty and vulnerability assessment** using a Participatory Rural Appraisal (PRA) that describes local and multi-dimensional perspectives of wealth and well-being; (ii) an **institutional assessment that** captures empowerment and institutional change, particularly at the community and district administrative levels; (iii) a **reality check approach** (RCA) that uses a light-touch participant observation approach to better understand how the MVP affects the realities of people's lives, as well as to capture any unintended consequences; and (iv) an **interpretational lens approach** that takes some of the preliminary quantitative survey findings and obtains local feedback and interpretation around emerging themes of analysis.
12. As outlined in the Initial Design Document (Masset et al. 2013b), the goal of the evaluation is to assess whether the intervention is meeting the MDG targets; as well as whether it is doing so in a sustainable way by breaking poverty trap¹⁵ constraints. In order to achieve these goals, the Initial Design Document outlined a data collection plan that sets out to gather detailed information on most MDG targets every two years while data on incomes, and expenditure and poverty is collected annually – and this is further elaborated in the Analysis Plan (2014). The annual data on income and expenditure is used to analyse poverty dynamics over the five years of project operations in order to test the project's ability to break poverty traps at community and household levels.

1.3 Aims and structure of the report

13. The overall purpose of this report is to provide a midterm assessment of the MVP; to consider progress towards the MDG targets alongside changes in other outcome variables and spill-over effects; and its likely sustainability. The MVP officially commenced in August 2012, although some of the first year's activities were preparatory (e.g. consultations with farmers) and would not have contributed directly to observable effects. It should therefore be noted that the report does not provide a full impact analysis at this stage (two years into the project at the time of the midterm household survey data collection), but rather highlights those areas where change is beginning to occur (both positive and negative). For this reason, we only focus on some parts of the Analysis Plan, including presenting the impact against the MDG indicators as well as a number of other variables. It does not undertake a full analysis against the poverty trap, which will be assessed in the final report.
14. This report is a summary of more detailed data collection, analysis and findings which are contained in annexes to this report. The most substantive of these annexes are four separate reports covering: (i) statistical analysis of the household datasets, blood tests and anthropometric measurements (Annex A); (ii) poverty and well-being assessment based on a qualitative participatory rural appraisal (Annex B); (iii) an RCA drawing on participant observation in a

¹⁵ For a full discussion of the 'poverty trap', please refer to the Initial Design Document (Masset et al. 2013b).

selection of treatment and control communities (Annex C); and (iv) the institutional assessment, which focuses particularly on the district administrative levels (Annex D).

15. This midterm report is an informed and triangulated summary of the quantitative and qualitative analysis. In this report, Chapter 2 provides an overview of the methodology, including a discussion of the quality of the data and its suitability for a DD analysis. Chapter 3 summarises MVP implementation activities to date and involvement of groups in project activities. Chapter 4 describes the impact MV has had on the MDGs after two years of intervention, without adjusting for differences in baseline characteristics. This is followed by Chapter 5, which presents impact against a number of core variables: monetary poverty, income and food security, child health, anthropometry, anaemia and education. This chapter also includes a discussion of the heterogeneity of programme impacts by sex and administrative districts. Chapter 6 explores the presence of spatial spill-over effects, followed by a chapter on institutional sustainability (Chapter 7). The final chapter provides the overall findings and conclusions.
16. **Note:** As this is an ongoing study at the time of the public release of this report, the names of control villages and the identities of particular focus groups have been replaced with identification references to protect the identity of the control communities and the study population. For individuals and households, these identities are not available and the publicly available dataset has been anonymised with appropriate technical safeguards in order to comply with ethical requirements and those of the Institutional Review Board of Columbia University.

Chapter 2. Methodology

17. This chapter gives a summary of the methodology, with further details provided in the Initial Design Document and the Analysis Plan. Methodological updates for this round of surveys and the qualitative studies are presented in the respective annexes to this report. The key questions for this evaluation were derived from discussions between DFID and other key stakeholders before commissioning the evaluation. The central tenet of the evaluation is to consider impact, both in terms of progress towards the MDG targets as well as a broader set of questions. Indeed, it is perhaps reasonable to expect some progress towards the MDGs given the level of investment in a relatively small area (£11.5 million in 35 communities from DFID alone). As such, the evaluation is tasked with looking further, including: measuring the extent to which these impact/outcome variables can be attributed to the MVP rather than other factors; whether these impacts are sustainable beyond the life of the project; and whether the intervention is cost effective compared to other possible policy options.
18. These concerns underpin the core questions for the evaluation, which can be summarised as:
 1. Does the MVP deliver on promises to **reach the MDGs** within the MV site?
 2. What **externalities or spill-over effects** does the MV generate, and do they significantly add to or detract from the positive impacts that might be achieved within the MV site?
 3. Are the positive impacts of the **MV sustainable after direct implementation** of the MVP has ended?
 4. Is the MV intervention package **cost effective** in the results it achieves, compared with possible alternatives?
 5. Does the MV package **empower disadvantaged or marginalised groups** (e.g. females, the disabled or the elderly)?
 6. Does MV achieve **additional benefits arising from synergies** across implementation of an integrated package of interventions?

2.1 The evaluation design

19. As mentioned earlier, the evaluation design uses a DD approach for the estimation of the project impact. In a DD framework, observations in the project group are compared to observations in a control group. In the simplest set-up, none of the groups are exposed to the treatment in the first period while the project group is exposed to the treatment in the second period but the control group is not. Given a number of assumptions and qualifications discussed below, this approach removes the biases produced by changes over time and by differences between the groups and provides an estimate of project effects.
20. Our sample consists of 35 project villages and 68 control villages. The project villages were selected within the three districts where the project is implemented using a one-to-one matching method based on a set of village level variables from the 2010 census, supplemented by village level observations collected in the field. Each project village was paired to a control village from two strata. One stratum was composed of potential controls in the vicinity of the project and the other stratum was composed of potential controls far from the project. Hence, there are 35 project communities, 34 control communities in the vicinity of the intervention area, and 34 control communities far away from the intervention area but within the district boundaries. The oversampling of the control communities was conducted with the goal of providing an estimation

of project spill-over effects to neighbouring communities and of allowing the use of matching methods at the analysis stage (see Section 6.1, Analysis Plan).

21. The table below lists the data collected during the 2014 (Year 3) survey round in both the MV and CV locations.

Table 1. Survey instruments used during 2014

Survey instrument	Description
Household (full)	Modules designed by the impact evaluation team on income, expenditure, in/out migration and social networks. Other modules are mostly focused on questions regarding achievement of the MDGs (education, malaria, water, sanitation, time use in the home, etc.)
Adult	Developed from internationally accepted standards for demographic and health surveys (DHS) used to calculate child mortality, etc.
Facility	A tool on characteristics, staffing and usage of main health and education facilities (clinics, primary schools and junior secondary schools)
Community	Designed by the impact evaluation team to capture village level data on land area, distance to facilities, economic activities, market prices, shocks and development projects. Prices and shocks are collected in the levels as well as in the trends over the preceding three or five years in order to perform a more rigorous DD analysis of project outcomes
Anthropometry	Heights and weights of all children under 5 are taken
Blood	Haemoglobin of all children under 5 obtained by finger-pricking
Cognitive and learning tests	Observe the MVP's impact on learning outcomes, which are not otherwise captured by the household (HH) questionnaire (i.e. which focus on attendance rates, and highest grade achieved). They include Raven's matrices, backward and forward digit span, short and advanced maths and English language tests
Expectations	These are designed to test people's expectations of survival, income and educational returns – all of which may be impacted by the MVP

22. In addition to the quantitative instruments outlined above, the evaluation design includes a number of qualitative modules. There are four key qualitative methods: (i) a poverty and vulnerability assessment, which was conducted during the baseline to identify local views of wealth, well-being and marginalisation; (ii) an institutional assessment, focusing on the effects of MVP on district government arrangements and longer term sustainability; and (iii) a RCA methodology that draws on participant observation to better understand the realities faced by people living in the area; (iv) an interpretational lens approach using participatory rural appraisal (PRA) tools to gather understanding on emerging findings from the survey data (referred to hereafter as PRA).¹⁶

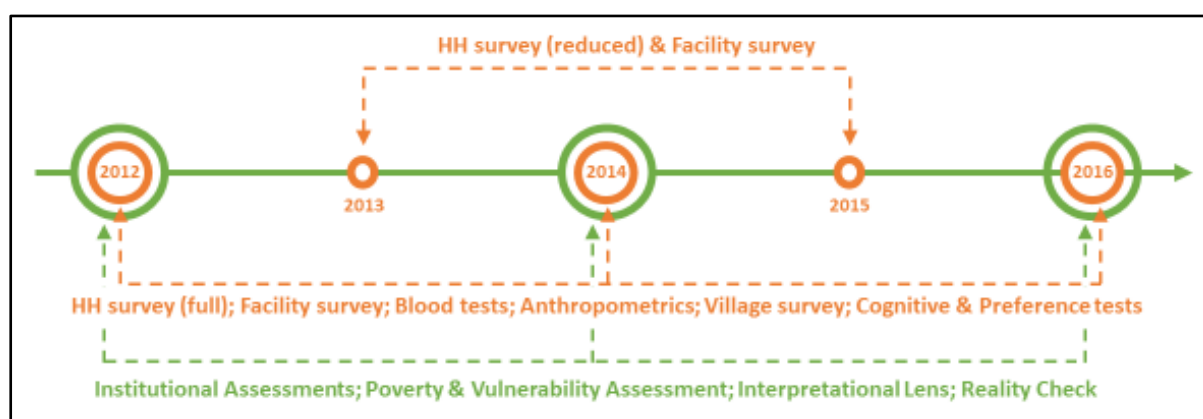
23. The qualitative research elements mainly focus on providing a deeper understanding of 'how' and 'why' change happened including changes in knowledge and behaviour, statistically significant

¹⁶ It should be noted that the interpretational lens in effect repeats the PRA work from the baseline. The main difference is that the interpretational lens does not repeat the well-being categorisation but revisits these same groups with a greater focus on change (and explaining and interpreting change). The mid-term was the first use of this approach, but because of delays in gaining access to the quantitative data, the evaluation team used a shortlist of preliminary issues to guide the qualitative follow-up. The endline plans to make more systematic use of the initial data analysis from either the 2015 or 2016 datasets (depending on what is feasible).

positive and negative differences, what happened where there was no change despite careful and deliberate intervention, spill-over changes in other communities, etc. They also provide local perspectives, gathering evidence that a quantitative dataset cannot be reasonably expected to capture (such as how programme interventions have been linked together or otherwise; how relationships have changed at the household, community, district and regional levels; how institutional arrangements have altered; as well as people's views, perspectives, opinions and reasoning).

24. All the qualitative modules, plus the full household survey and the other quantitative instruments (as set out in Table 1) were conducted during the baseline and midline – and will be repeated for the endline in 2016. During the in-between years (2013 and 2015) only a reduced household instrument was used, mainly to provide an annual dataset for modules on income, consumption and expenditure.

Figure 1. Summary of data collection by year



Source: Itad's elaboration

2.2 Methodology for estimating difference-in-difference analysis

25. Project impact is estimated using a DD analysis: the difference in the change over time in the average outcomes between the project and in the comparator groups. In the simple standard two-period and two-group set-up, the DD effect is:

$$\delta = (\bar{y}_{P,1} - \bar{y}_{P,0}) - (\bar{y}_{C,1} - \bar{y}_{C,0})$$

where δ is the DD effect, y is the average outcome either in the project group (P) or in the comparison group (C) observed in the first period (0) and in the second period (1).

26. We calculate DD effect using regression analysis. We use different regression models depending on whether (i) data are available for panel observations; and (ii) data are available for two or three periods.

The simplest model is the **cross-sectional regression (t=0,1)**:

$$y_i = a + bT_i + cP_i + dP_iT_i + \sum_{j=1}^n g_j X_{ji} + e_i$$

where y is the outcome for the observation i , T is a dummy variable equal to 0 for period 1 and equal to 1 for period 2, P is a dummy variable equal to 1 if the observation is in the project group and equal to 0 if the observation is in the control group, PT is equal to 1 if the observation is both in the project group and observed in the second period. The equation estimates the following: a is the average outcome in the control group in period 1; b is the difference in the outcomes between period 2 and period 1 in the control group (the time trend); c is the difference between

project group and control group in period 1; d is the DD effect of the project. The (X_i) are covariates that improve the balance between the project and comparison group samples, as these were not randomly obtained from an experiment, and increases precision of the estimates by reducing the standard error of the coefficients. One potential problem with the use of covariates in the estimation of project effects is that most covariates are affected by the project or are themselves objectives of the intervention. Think, for example, of a DD regression of height-for-age including changes in total household expenditure. The inclusion of variables affected by the programme will 'absorb' some of the project effects that would otherwise be captured by project dummies. Hence, in order to capture the programme impact with a project dummy interaction, the covariates can only include baseline characteristics or variables that are not affected by the programme Rosenbaum (1984).

27. When panel data are available we use a **fixed effects model** to remove the impact of fixed effects: time-invariant unobservable determinants of the outcomes such as, for example, farmers' motivation or children's abilities. The fixed effect model is:

$$y_{it} = a_i + bT_{it} + dP_{i1}T_{i1} + \sum_{j=1}^n g_j X_{itj} + e_{it}$$

28. The covariates in this case are time-varying variables that are not affected by the project such as, for example, the occurrence of drought or other shocks. As recommended by Angrist and Pischke (2009), we also employ the **lagged outcome model** (Imbens and Wooldridge 2009) also known as the analysis of covariance (ANCOVA) model:

$$y_{i1} = a + by_{i0} + dP_i + \sum_{j=1}^n g_j X_{i0j} + e_i$$

which is simply a regression of the dependent variable in period 2 on the dependent variable in period 1 and a project dummy in addition to the usual baseline covariates.

29. As an aside we note that these models can be expanded to include multiple time periods and for completeness we report below the model specifications employing three periods. For each of the three models above we report the specification estimating the average project effect over the three-year period and the specifications estimating two-year specific project effects.

3-period cross-sectional models (t=0,1,2):

$$y_i = a + b_1T_{i1} + b_2T_{i2} + cP_i + dP_iT_i + \sum_{j=1}^n g_j X_{ij} + e_i$$

$$y_i = a + b_1T_{i1} + b_2T_{i2} + cP_i + dP_iT_{i1} + dP_iT_{i2} + \sum_{j=1}^n g_j X_{ij} + e_i$$

3-period fixed effects models:

$$y_{it} = a_i + b_1T_{i1} + b_2T_{i2} + dP_iT_i + \sum_{j=1}^n g_j X_{itj} + e_{it}$$

$$y_{it} = a_i + b_1T_{i1} + b_2T_{i2} + d_1P_iT_{i1} + d_2P_iT_{i2} + \sum_{j=1}^n g_j X_{itj} + e_{it}$$

3-period lagged models (t=1,2):

$$y_{it} = a + by_{it-1} + dP_iT_{it} + \sum_{j=1}^n g_jX_{ioj} + e_i$$

$$y_{it} = a + by_{it-1} + d_1P_iT_{i1} + d_2P_iT_{i2} + \sum_{j=1}^n g_jX_{ioj} + e_i$$

30. The comparator villages surveyed at the baseline were identified by matching district villages to project villages using a propensity score built using village level characteristics obtained from census data and from field visits. In order to remove remaining baseline differences in characteristics between the project and the control group we further employ matching methods at the household and individual level in the estimation of the project effects. In doing so we follow the methodology for the estimation of treatment effects under un-confoundedness outlined by Imbens and Rubin (2015).
31. Imbens and Rubin recommend the separation of the design stage from the analysis stage in conducting observational studies. The goal of the 'design' stage is to select a propensity score and a sample of observations that maximises the statistical balance of the distribution of the covariates. In the design stage the outcomes are completely ignored in order not to bias the construction of the propensity score. The goal of the analysis stage is to estimate project effects in the selected sample using the propensity score estimated in the design stage. We briefly describe the various steps followed in the design and analysis stage. In order to provide an example, all the results of the procedure adopted to estimate the impact of the project on per-adult equivalent expenditure are reported in Appendix A1.
32. **Design stage:** We estimate the propensity score using a logistic regression model. Imbens and Rubin (2015) propose an algorithm for the estimation of the propensity score which aims at achieving statistical balance of the covariates and does not try to 'explain' participation through a behavioural model. The initial covariates are selected based on substantive knowledge of the existing literature and of the context. Covariates are subdivided into:
- *Basic* covariates that are known to be strong determinants of the outcomes or of participation in the project. These covariates are included in the model regardless of the statistical significance of their correlation with project status.
 - *Additional* covariates that are likely to be correlated with the outcomes or with participation in the project. These covariates are added to the logistic model stepwise based on the statistical significance of their correlation with project status.
 - A group of powers and interactions of all *basic* and *additional* variables identified.
33. We assess the validity of the estimated propensity score by testing the balance of the covariates. We first subdivide the sample based on the propensity score using the algorithm suggested by Imbens and Rubin (2015). We then conduct the three tests of sections 17.3.1, 17.3.2 and 17.3.3 and we plot the figures of section 13.8 (Imbens and Rubin 2015).
34. We assess the degree of overlap in the distribution of the covariates in the project and the comparator groups. We need to avoid making comparisons for observations, either in the project or comparison group, which have very few or no similar observations in the other group. To do this we inspect the distribution of the propensity scores using histograms and calculate the proportion of project and control observations with sufficiently 'good' matches using the method described in section 14.5 of Imbens and Rubin. We then trim the sample to remove the

observations that are outside the region of overlap. The region of overlap is identified using the algorithm described in section 16.4 (Imbens and Rubin 2015).

35. **Analysis stage:** We first repeat some of the operations conducted at the design stage to calculate the propensity score and identify the size of the estimation sample. We run a logit participation model and we calculate the propensity score. We trim the sample to remove observations outside the region of overlap. We re-estimate the propensity score on the trimmed sample.
36. We estimate the majority of project effects using inverse probability weights (IPW) calculated using propensity scores and normalised following the procedure in 17.8.1 of Imbens and Rubin. We employ this method over more accurate estimators such as sub-classification (see below) because: (i) we have to estimate impacts over a large number of outcomes and the application of IPW is straightforward in comparison with other approaches; (ii) in some cases (child mortality rates in particular) the number of observed events is very small and the calculation of effects within small blocks is hardly feasible.
37. In some instances we employ the sub-classification method recommended by Imbens and Rubin. In order to do so, we split the sample into blocks based on the estimated propensity score until the statistical difference in the propensity score within each block is removed following the algorithm of section 13.5. We then calculate project effects as the weighted average of the block level treatment effects, where the weights are calculated using the proportions of control and project observations in each block. The methodology follows sections 17.3 and 17.5 in Imbens and Rubin. Treatment effects within each block are DD regressions.
38. And finally, at the design stage of the evaluation we recognised that the intervention could produce changes in the provision of services by non-governmental organisations (NGOs) and government in both the MV and CV areas, and that this could bias the estimation of project effects. In particular, we identified two potential effects operating in opposite directions: a crowding-out effect and a reorientation of district level expenditure effect. In the first case, the intervention forces NGOs and / or the district / government to move resources away from MV and potentially to CV areas – particularly as MV operates in nearly all assistance sectors (agriculture, health, social mobilisation etc.). If NGOs / district are moving resources to the CV areas from the MV areas, or starting new operations in the selected CV areas, the programme impact is somehow underestimated. In the second case, the government is induced to invest resources from the education, agriculture and health budget in the MV areas because of a cooperation agreement with MVP. This is correctly identified as a positive outcome of the intervention by project managers. The MV acts as a catalyst of public resources for the development of the cluster. However, if the additional public resources invested in MV areas are subtracted from resources normally invested in the selected CV areas, the project impact is overestimated.
39. Thus far, our qualitative work and quantitative information collected on NGOs and district investments suggests that both effects are occurring to some extent. However, it is difficult to estimate the size of these effects and anticipate the final direction of the bias on the estimation of project effects as it could go either way.

2.3 Data collection and sample characteristics

40. This section describes the characteristics of the sample and its suitability for assessing project impact. The validity of a DD approach rests on the assumption that project and comparison groups are similar. Differential trends in the outcomes and covariate shocks were discussed in the previous second round analysis reports and were not found to be major threats. Here, we focus

on changes in the composition of the project and comparison groups produced by attrition, migration and measurement error.

41. The baseline survey targeted a sample of 755 households in the MV villages and 1,496 households in the CV villages. However, the baseline sample comprises only 711 MV households and 1,461 CV households because not all the target households were found at the time of the baseline interviews. The second round of interviews and the midterm surveys targeted the same 755 MV households and 1,496 CV households originally selected at the baseline. The largest number of household interviews was conducted in the second round, followed by the third round and the baseline round. No systematic differences emerged in the completion of survey interviews between project and comparison areas in any of the three rounds (Table 2). If anything, more households were interviewed in CV areas than MV areas at the baseline, and the difference decreases over the three rounds (97% CV against 94% MV at baseline; 99% CV versus 98% MV in the second round; and 97% in both CV and MV at the midterm).

Table 2. Completed household interviews during the first three rounds

Sample	Target	2012	2013	2014
MV interviews	755	711	743	735
%		94	98	97
CV interviews	1,496	1,461	1,487	1,456
%		97	99	97
ALL interviews	2,251	2,172	2,230	2,191
%		96	99	97

42. These very high completion rates are the result of efforts to locate households in repeated visits and of a very limited number of cases of household relocation or dissolution. The numbers of the latter cases are reported in Table 3 and are negligible in all survey rounds.

Table 3. Reasons for not completing the interviews

Reason	2012	2013	2014
No. competent household member at home	21	1	8
Entire household absent	22		11
Interview postponed	10		
Interview refused	1		
Partly completed			
Dwelling vacant or destroyed		4	2
Dwelling not found	19	9	13
Household has relocated		6	15
Household dissolved or deceased		1	6
Other	6		4
ALL	79	21	59

43. Because not all target households are interviewed every year and because those that are not interviewed differ from year to year, the number of full panel households interviewed every year decreases over time. The decrement is, however, very small (Table 4). The baseline survey interviewed 97% of target households, the second round retained 96% of the panel target households, while the midterm survey retained 94%. The absolute numbers of households not re-interviewed are so small (14 households in MV areas and 51 households in the CV areas were lost at midterm in comparison to the baseline) that a comparative analysis of the characteristics of

attriters and non-attriters in MV and CV areas is hardly feasible. In addition, these are not proper 'attriters' but simply households that could not be interviewed in all three survey rounds. The number of households permanently leaving the sample at some point is even smaller. Finally, attrition rates are very similar across the project and comparison groups and differential attrition is not a major threat.

Table 4. Completed household interviews during the first three rounds

Sample	Target	2012	2013	2014
MV panel interviews	755	711	707	697
%		94.2%	93.6%	92.3%
CV panel interviews	1,496	1,461	1,454	1,424
%		97.7%	97.2%	95.2%
ALL panel interviews	2,251	2,172	2,161	2,121
%		96.5%	96%	94.2%

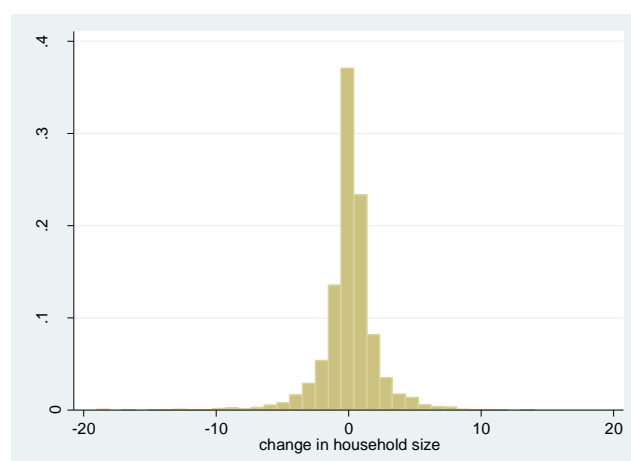
44. Although the rate of household attrition is very small, attrition of individuals is quite large (Table 5). Only 90% of the individuals originally enumerated at the baseline were again enumerated after two years. This, however, is only partly the result of migration, and is largely produced by natural processes (deaths and births) and changes in household composition.

Table 5. Individuals listed in the surveys

Sample	2012	2013	2014
MV individuals	5,231	5,576	5,854
<i>MV panel</i>		4,930	4,654
Fraction of baseline sample		94.2%	89%
CV individuals	10,337	10,649	11,023
<i>CV panel</i>		9,869	9,378
Fraction of baseline sample		95.5%	91%

45. Average household size is stable across survey periods (7.1 at baseline, 7.0 at second round and 7.2 at midterm) but there is considerable change in household composition, as shown in Figure 2. We calculated average household size of panel households at baseline and at midterm and the changes in size between the two periods. Less than 40% of households preserve the same sample size across a two-year period and there are some dramatic changes by more than ten household members. These changes reflect a high level of individual mobility across time, though are comparable with changes observed in very different socio-economic contexts.¹⁷

¹⁷ See for example Halliday for examples of observed changes in households size in El Salvador and in the United States HALLIDAY, T. J. 2005. Mismeasured Household Size and its Implication for the Identification of Economies of Scale. *Oxford Bulletin of Economics and Statistics*, 72, 246-262.

Figure 2. Change in household size between baseline and midterm

Source: Itad's elaboration

46. There are only two ways for people to become a household member: being born or moving into it. Similarly, there are only two ways to leave a household: death and moving out of it. Household size at any time is therefore the result of changes in the natural population growth (births minus deaths) and changes in net migration (in-migrants minus out-migrants). To these changes we need to add reporting errors. The fluidity of household membership in the area, the large number of household members and circumstances of enumeration result in large errors. To simplify, we define as type I error all enumerated individuals that should not have been enumerated, and we define as type II error all enumerated individuals in the current round that were erroneously not enumerated in previous rounds.
47. The decomposition of population change in natural growth, migration and errors is shown in Table 6 for MV and CV areas separately using only those sample households (2,212) that were interviewed for three consecutive rounds. Rates of change are calculated over the number of household members in the previous round. The natural population change is positive and is very similar in MV and CV areas. The fraction of people moving in and out of households is considerable but the two terms tend to cancel out and the changes are very similar in MV and CV areas. Enumeration errors also cancel out at the second round but not at the midterm when type I error is larger (thus leading to an increase in population as more people are found at the midterm that had not been previously, and erroneously, enumerated). Again the patterns of enumeration errors are very similar in MV and CV areas.

Table 6. Percent of population changes across surveys

	All		MV		CV	
	2nd round	midterm	2nd round	midterm	2nd round	midterm
Overall change	-1.3	3.5	-1.3	2.1	-1.3	3.7
Natural change (births minus deaths)	0.8	1.5	0.9	1.5	0.7	1.5
Migration: people moving in	1.7	4.0	2.3	4.3	1.5	3.9
Migration: people moving out	-3.2	-4.0	-4.2	-4.5	-2.7	-3.8
Type I error	4.1	3.6	5.0	2.4	3.6	3.6
Type II error	-4.3	-1.5	-5.1	-1.9	-3.9	-2.1
Residual unexplained difference	0.4	-0.1	-0.2	0.4	-0.5	0.6

48. Overall this data suggests that, despite great challenges, the survey teams were able to enumerate households with great accuracy and that no large differences emerge between MV and CV areas

in population changes driven by natural change, movements in and out of households and reporting errors.

2.4 Institutional assessment

49. The institutional assessment was undertaken by the midterm institutional team comprising four researchers. The assessment took place in May 2015 in the three districts. Six focus group meetings (two in each of the three districts) were held. One was held with the district administrative team, and the other with the heads or representatives of the district technical departments and agencies.
50. On a few occasions the study team used simple scoring to provide an idea of participant emphasis or weighting. Using a scale of one (low) to ten (high), the focus group discussion participants were asked to discuss and give a score to an issue or projects. However, the main point of using the scoring methodology was to open debate and explore the reasons and explanations people gave for their scores rather than the scores themselves.
51. Analysis of the focus group discussions with the district administrative team in each district revealed very similar issues. The same was true for focus groups with the district technical departments. For this reason, instead of producing separate reports for each district, as was done at the baseline, there is only one report for the midterm institutional study, incorporating the perspectives of the district administrators and those of the technical personnel. Where deemed necessary the views of the SADA-MVP team have also been incorporated. The report has focused on what the district officials and the SADA-MVP team had to say.

2.5 Participatory Rural Appraisal

52. This focus group study was undertaken in February and March 2015 with a team of 17 researchers, including five team leaders. They were supported by a research trainer/coordinator and a lead writer. A total of 20 field sites (grouped into seven comparator clusters) were visited by the research team. Each comparator cluster comprises three communities – one Millennium Village (MV) in which the MVP is intervening, a *nearby* control village (CN) and another *faraway* control village (CF).¹⁸
53. The categorisation of villagers into focus groups was informed by a prior well-being ranking exercise at the baseline (and updated at the midterm) in which community representatives compiled locally relevant descriptors of well-being for their community and matched those descriptors with between four and six well-being strata. In addition to the four recurring categories of rich, moderately rich, poor and very poor households,¹⁹ a small minority of communities also identified either a very rich or destitute stratum. In those few cases where there were more than four strata, participants were asked to merge some strata (e.g. rich and very rich, or very poor and destitute) to facilitate comparison across the communities.
54. In order to enhance the tracking of qualitative changes and promote local learning from the evaluation process, the composition of the focus groups interviewed at the baseline was maintained for the midterm assessment. While there were cases in which some participants had either migrated out or died since the baseline, the core composition remained largely intact across the communities.

¹⁸ One of the seven sets is short by a faraway control because the site visited at the baseline was found to have been included in error.

¹⁹ The words 'rich' and 'poor' in this context do not connote the same meaning as used in market-led economies but are rather comparative terms. It is a relative scale of wealth and well-being.

55. The study built on the areas of conversation outlined for the earlier baseline. These were refined and augmented by specific areas of interest flagged by the quantitative research team for deeper enquiry, thereby helping to ensure a link between the survey data analysis and the qualitative work. Additional key informant interviews were conducted with children attending school and those not attending school, teachers, some lead farmers/opinion leaders such as chiefs, district assembly representatives and individuals who the focus group discussions identified as having unique or informative experiences.
56. Throughout the fieldwork, each team tried to make time to reflect on the information they were receiving, identify apparent conflicts and take steps to clarify those findings that were either incomplete or that required further investigation. At the end of the fieldwork, a three-day debriefing workshop was held between the research assistants and team leaders, the research coordinator and the lead writer to distil the provisional findings and identify areas requiring further clarification – either from the communities, from specific institutions mentioned or from the SADA-MVP office in Bolgatanga. A half-day meeting was held with the SADA-MVP team in Bolgatanga to that end.

2.6 Reality check approach

57. The RCA extends the tradition of listening studies (see Anderson, Brown and Jean 2012)²⁰ and beneficiary assessment approaches (see Salmen 1998²¹ and Shutt and Ruedin 2013)²² by combining elements of these approaches with researchers actually living with people whose views are being sought, usually those who are directly experiencing poverty. The RCA can be likened to a ‘light-touch’ participant observation. It is similar in that it requires participation in everyday life within people’s own environments but differs by being comparatively quick and placing more emphasis on informal, relaxed and insightful conversations than on observing behaviour and the complexities of relationships.
58. The six RCA study villages are the same as those used in the baseline study (April 2013). The decision to select four treatments and two controls was based on consultation with the MVP M&E expert after the pilot in 2013. The four MVP are not among the PRA sites, but used the same ratio of Mamprusi to Builsa communities as the programme. Controls are not directly matched but represent one far and one near control. The same households selected in April 2013 were revisited. These households had been selected together with the community to fulfil the following criteria: (i) Poorer households; (ii) Households with different generations living in the house including, where possible, school-age children; (iii) Households in each village which were at least ten minutes’ walk from each other; (iv) Households located at the centre of the village as well as the periphery; and (v) Households with a number of close neighbours (to enable interaction with them as well). Four households selected in 2013 did not fulfil the criteria of having multiple generations in the household and so this time while researchers stayed with the original households, they interacted more with neighbouring households with larger and multi-generational families. One household was dropped as there had only been a single person living there and there were limited opportunities to engage with her and her neighbours. In addition to the 19 households where the RCA had intense interactions, the team had extensive conversations with more than 560 other people in the communities.
59. The RCA study was conducted in two parts with two teams in the field from 30 April to 5 May, and from 28 May to 2 June 2015. Two researchers who were involved in the 2013 fieldwork had to be

²⁰ Anderson, M.B., D. Brown and I. Jean. 2012. *Time to Listen; Hearing People on the Receiving End of International Aid*. Cambridge MA: CDA.

²¹ Salmen, L.F. 1998. Toward a listening bank: review of best practices and efficacy of beneficiary assessments. *Social Development Papers* 23. Washington, DC: World Bank.

²² Shutt, C. and L. Ruedin (SDC). 2013. *How to Note Beneficiary Assessment*. Berne, Swiss Agency for Development Cooperation.

substituted since they were unavailable. This was facilitated by former researchers providing thorough briefings to those who would take over their host households. Refresher training was provided to all researchers, which gave an opportunity for the team to induct four new researchers.

60. The focus of this second RCA study was on 'change' and what people felt had changed for better or worse or had stayed the same since our previous visit in 2013. Using photographs taken in 2013 as prompts, picking up on conversations recalled from the baseline and sometimes our own observations of change (e.g. the installation of electricity poles, construction of new roads) we were able to have open and frank conversations about how people felt about these changes, what were good and less good changes and the perceived drivers of those changes both in their family and community spheres.

Chapter 3. MVP implementation

61. This chapter explores the implementation of MVP activities from 2012 to 2014 as recorded by the MVP itself, complemented by the impact evaluation's own analysis. The first part of the chapter (Section 3.1) is a summary of the project's self-reporting of progress (extracted from SADA-MVP reports), which is followed by the evaluation team's initial analysis of progress in terms of infrastructure (Section 3.2). We then proceed to summarise in section 3.3 'participation rates', which is the survey respondent's views of 'being part of', 'included in' or 'having access to' a particular activity. The final part (Section 3.4) summarises some of the findings derived from the evaluation's qualitative elements to provide perspectives on the experience of these activities from community members themselves (beneficiaries and non-beneficiaries).

3.1 Context: activity progress reported by MVP

62. Since 2012, the MVP has implemented a number of coordinated interventions across the health, education, agriculture, infrastructure, energy, water and sanitation sectors. These have been supported by community engagement efforts encouraging citizens to participate in the MVP's activities in order to build community capacity and local institutions as essential features of a sustainability strategy. This section provides a brief overview of MVP implementation, based on self-reported data from various project documents, including the 2013²³ and 2014²⁴ Annual Reports and the 2015²⁵ mid-year report. The following paragraphs are structured around MVP's logframe,²⁶ and set out as a summary overview of the project's own reporting on key achievements. In this initial section, we do not attempt to triangulate this information with the evaluation's own datasets, and do not seek to make any judgements about activities, achievements, performance or non-performance. The purpose is rather to provide background information on the MVP activities to date.

Output 1: Improved food and nutrition security and development of agricultural value chain

63. Under this output, the MVP developed an agricultural strategy prioritising the improvement of extension work and agricultural services delivery, increasing agricultural productivity, supporting cooperative and new business development, linking producers to markets and promoting access to financial services. Activities in the period have concentrated on distributing agricultural inputs (primarily fertiliser, seeds, saplings, livestock and tractor services), improving crop irrigation to optimise growing seasons (especially rice), providing training and extension services, promoting links and access to markets, building grain stores, developing value chains, preventing bushfires and ensuring access to rotating credit or group savings through farmer-based organisations and village savings and loans associations (VSLAs). Specific activities and outputs over the past three years are documented in Table 7.

²³ 2013 Annual Report on the Millennium Village Project in Northern Ghana, SADA, December 2013.

²⁴ 2014 Annual Report on the Millennium Village Project in Northern Ghana, SADA, December 2014.

²⁵ 2015 Mid-Year Report on the Millennium Village Project in Northern Ghana, SADA, July 2015.

²⁶ Millennium Villages Evaluation Logframe, DFID, January 2014 (N.B. the Logframe was updated in April 2016).

Table 7. Activities and outputs in the agriculture sector

<i>Sector description</i>	<i>Key achievements</i>	
	2013	2014
<p><i>Agriculture, cooperatives and business development, including:</i></p> <ul style="list-style-type: none"> • <i>Surveys</i> • <i>Creating farmer-based organisations</i> • <i>Distributing seedlings</i> • <i>Trainings</i> • <i>Cultivation</i> • <i>Recruitment</i> • <i>Technical assistance</i> 	<ul style="list-style-type: none"> • Survey and feasibility studies of irrigation schemes • 192 farmer-based organisations established • 9,000 mango and 3,000 acacia seedlings distributed • 2,935 farmers trained 	<ul style="list-style-type: none"> • 271 hectares cultivated for rice paddy production • 23 Agricultural extension area training communities on agronomic practices • 166 lead farmers recruited • 3 grain warehouses constructed • 173 vulnerable community members identified • 8 VSLA saved \$6,000

Source: SADA-MVP 2013 & 2014 Annual Reports and 2015 SADA-MVP Mid-Year Report.

Output 2: Enhanced access to quality primary education

64. The goals of the MVP under this output are twofold: first, to achieve universal primary education; and second, to increase the overall quality of education. The priority areas under this output are education quality, primary school enrolment, participation in secondary education, improving gender parity and engaging communities in education. To date, activities have included constructing or rehabilitating classrooms, toilets, water pumps, recreation facilities and teachers' quarters, teacher training and placement, recruiting and building the capacity of community education workers to motivate the community on the importance of enrolling children (particularly girls) in school, awarding scholarships to girls pursuing secondary education and procuring teaching and learning supplies. Specific activities and achievements over the past three years are documented in Table 8.

Table 8. Activities and outputs in the education sector

<i>Sector description</i>	<i>Key achievements</i>	
	2013	2014
<p><i>Education, including:</i></p> <ul style="list-style-type: none"> • <i>Construction</i> • <i>Curriculum development</i> • <i>Training</i> • <i>Technical assistance</i> 	<ul style="list-style-type: none"> • 4 teachers' quarters constructed • 16 classrooms constructed • 18 playgrounds installed • Curriculum manuals delivered to 18 schools • Radio programmes & school dramas hosted • 60 students sent to a girls' camp 	<ul style="list-style-type: none"> • 26 classrooms constructed; 27 rehabilitated • 3 teachers' quarters built; 1 reconstructed • Pupil to classroom ratio is 43:1 • 50 Community Education Workers appointed to primary schools • 160 teachers trained

Source: SADA-MVP 2013 & 2014 Annual Reports and SADA-MVP Mid-Year Report.

Output 3: Enhanced access to health care services by skilled personnel and trained community health worker (CHW) providers

65. The main aim of Output 3 is to reduce common, treatable causes for morbidity and mortality. By increasing access to clinics and referral hospitals, the MVP hopes to reduce death rates and

incidence of HIV, malaria, TB and other tropical diseases, reduce mortality among children under five years of age and improve nutrition and maternal health. Since 2012, the MVP has worked to achieve these outcomes by constructing, refurbishing and equipping village health facilities, hiring skilled health professionals, using rapid diagnostic tests to diagnose malaria, educating communities about signs of diseases, improving health information systems, recruiting CHWs, monitoring children's health, educating expectant and new mothers about neonatal care, and training midwives. Specific activities and achievements over the past three years are documented in Table 9.

Table 9. Activities and outputs in the health sector

<i>Sector description</i>	<i>Key achievements</i>	
	2013	2014
<p><i>Health, including:</i></p> <ul style="list-style-type: none"> • <i>Recruitment</i> • <i>Training</i> • <i>Rehabilitation and reconstruction</i> • <i>Construction</i> • <i>Purchases</i> • <i>Health information systems</i> • <i>Technical assistance</i> 	<ul style="list-style-type: none"> • 63 health professionals recruited • 39 CHWs trained • 2 health centres reconstructed • 1 maternity ward and 1 staff quarter constructed • 2 new ambulances deployed • CommCare²⁷ deployed 	<ul style="list-style-type: none"> • 97 health professionals recruited • 53 CHWs recruited • 2 health centres rehabilitated • 1 maternity ward constructed • 8 students received training • 100% of population has access to a health facility with one skilled staff within 10km

Source: SADA-MVP 2013 & 2014 Annual Reports and 2015 SADA-MVP Mid-Year Report.

Output 4: Improved access to electricity, roads and transportation, communications and improved water and sanitation facilities

66. Infrastructure is a pivotal area for the MVP as all other development activities hinge on having functional transportation systems, communication and energy. In collaboration with the government, the MVP has prioritised roads, transport, electricity, energy, increased access to internet and mobile phone services and water sanitation and hygiene (WASH) facilities. The main activities have included improving roads and transport services, expanding access to electricity (mains and solar), increasing access to improved water and sanitation, monitoring water points, implementing community-led total sanitation (CLTS) and installing communications infrastructure. Specific activities and achievements over the past three years are documented in Table 10.

Table 10. Activities and outputs in the infrastructure, water and sanitation sectors

<i>Sector description</i>	<i>Key achievements</i>	
	2013	2014
<p><i>Infrastructure, water and sanitation, including:</i></p> <ul style="list-style-type: none"> • <i>Lobbying the government of Ghana</i> • <i>Rehabilitation</i> • <i>Construction and installation</i> • <i>Community-led total sanitation (CLTS)</i> • <i>Technical assistance</i> 	<ul style="list-style-type: none"> • 10 roads being rehabilitated • 29 boreholes installed • 49 water points rehabilitated • CLTS rolled out • 7 Information communication technology (ICT) centres installed 	<ul style="list-style-type: none"> • 11 roads rehabilitated • 464 HHs with grid electricity • 49 water points rehabilitated • 28 boreholes under construction • CLTS triggered in 35 communities • 191 latrines constructed

²⁷ Mobile platform for building 'apps' for data collection.

Source: SADA-MVP 2013 & 2014 Annual Reports and 2015 Mid-Year Report SADA-MVP Mid-Year Report.

Output 5: Strengthened local institutions and community capacity to secure sustainability of MVP

67. Under this output, the MVP aims to empower people to advance the achievement of the MDGs in their communities. Increased participation in public decision making and enhanced local governance, with inclusion of vulnerable and minority groups is regarded as key to ensure that long-term project outcomes can be sustained. These aims have been addressed through capacity building activities (particularly community-based organisations – CBOs), constructing community centres and facilitating workshops with the local government. Specific activities and achievements over the past three years are documented in Table 11.

Table 11. Activities and outputs in community engagement and governance

<i>Sector description</i>	<i>Key achievements</i>	
	2013	2014
<i>Community engagement and governance</i> <ul style="list-style-type: none"> • <i>Training</i> • <i>Distributing information</i> • <i>Capacity building</i> • <i>Establishing CBOs</i> 	<ul style="list-style-type: none"> • 140 people attended a forum on MDGs • 62 people attended capacity building workshops • 22 CBOs established to support the MDGs 	<ul style="list-style-type: none"> • 68 CBOs established • 8 media houses told about the MVP

Source: SADA-MVP 2013 & 2014 Annual Reports and 2015 Mid-Year Report SADA-MVP Mid-Year Report.

3.2 MVP infrastructural progress

68. Among the central features that the MVP seeks to change are physical connectivity to the sites, access to healthcare, education and water. This section provides an account of infrastructure in the MVs and the two CVs (near and far) over two years (the baseline in 2012 and the midterm in 2014). We use the Earth Institute's data to undertake the assessment.²⁸ The content of the data differs across the two years and, thus, temporal comparisons are incomplete. In some instances data on functioning and non-functioning facilities were reported. When total numbers are reported we do not make the distinction. For each year we provide an overview, and then offer a more detailed account of water, education and health.

A summary of the baseline

69. Table 12 lists the differences observed at the start of the project in these factors across the MVs and the near and far control sites. At the start of the project none of the MVs were linked by paved roads; and a few of the villages in all three types of sites had no connection by road (paved, gravel or dirt). In 2012, there were no higher secondary schools and the MVs had dramatically fewer health facilities than were present in CVs. As far as we are able to ascertain from the dataset, neither the MVs nor CVs were connected to electricity gridlines.

²⁸ Earth Institute. 2012–14. Facility Surveys.

Table 12. Summaries of village facilities, 2012

Facility/infrastructure	MV	CV near	CV far
Transport Infrastructure to the settlement (Villages connected by road)	Gravel: 19 Dirt: 18	Gravel: 7 Dirt: 9 Paved + dirt: 2	Paved + Gravel: 1 Gravel: 13 Dirt: 11
Number of water sources	198	238	374
Number of primary schools	16	17	17
Number of junior secondary schools	10	6	9
Number of health facilities (not hospitals and those identified as other health clinics)	1	6	8
Hospitals	0	0	2

Source: Earth Institute Surveys, 2012

70. In general the CVs had better schooling facilities at the baseline. The facility surveys collected information on schools such as enrolment rates, number of classrooms in good physical condition (as reported in the survey) and the number of qualified teachers (as deemed by the school system). A few summary measures can be presented to assess the adequacy of educational resources that were available (see Table 16 on page 31). In 2012, the 'student to teacher' and 'student to functioning classroom' ratios were poor in the MVs compared with CVs.
71. There were also striking differences in the health sector in 2012 across the MVs and CVs, although both areas seemed inadequately served (Table 13). In the CVs there were two hospitals among other health posts. The combined formal health facilities (health posts, mobile and stationary clinics, hospitals and dispensaries) in the CVs totalled 16 whereas there was only one health post in the MV areas. There were no doctors working in the MVP areas in 2012. A number of questions were asked regarding services at the health facilities about transport for referral, malaria ACT²⁹ treatment, family planning and vaccination, etc. More questions were asked in 2012 than in 2014 and the 2014 questions were a sub-set of the 2012 surveys. Given that the designation of health facilities can be complicated, these questions are important indicators as to the kind of medical care residents in MV and CV areas can expect. In 2012, the only health facility in the MV area reported that it could provide more than half the activities or services out of the total number of services recorded by the survey. In CV areas, 14 of the facilities could provide more than half of these services.

Table 13. Health resources in 2012

Categories (number within the category)	MV	CV near	CV far
Health posts	1 ³⁰	1	2
Health clinics	0	4	5
Mobile clinic	0	1	1
Dispensary	0	0	0
Hospital	0	0	2
Other type of facilities	2	1	3
Doctors	0	3	4
Nurses	2	0	2
Medical officers	1	5	13
Medical assistants	1	4	8
Vaccine cold chains, operative	1	1	8
Total beds	3	10	273
Centres reporting number of services provided as asked (>13)	1	0	7

Source: Earth Institute. 2012. Health Survey

²⁹ Artemisinin-based Combination Therapy recommended for treatment of *P.falciparum* malaria³⁰ The health facility reported at baseline was funded by SADA.

Infrastructure changes from 2012 to 2014

72. The midterm dataset is more limited than the baseline dataset. For example, there were no data on road access by communities, which were reported in 2012. Another disparity is that no hospitals were reported in the CVs in 2014, which requires further enquiry since two hospitals were reported in 2012. It is possible that these two hospitals were closed or missed from the facility data collection after the districts were re-zoned. Since this data consistency issue was only discovered recently, the evaluation team is investigating whether the hospitals can be located and, if not, why they were shut down or downgraded. Table 14 summarises the 2012 and 2014 facility datasets.

Table 14. Summaries of village facilities, 2012 and 2014

Item	2012			2014		
	MV	CV near	CV far	MV	CV near	CV far
Water sources	198	238	374	526	430	443
Primary school	16	17	17	20	14	16
Junior secondary school	10	6	9	12	7	9
Health centres	1	6	8	13	5	5
Hospitals	0	0	2	1	0	0

Source: Earth Institute Survey, 2014.

73. Overall, the total number of water sources in the MVs and CVs has increased, which appears to contradict the household survey findings where a negative impact on the proportion of households able to access safe drinking water was reported (see Section 4.1 for details). A larger portion of water sources were functional in all areas in 2014 compared to 2012. Whereas 32.3% of water sources were non-functional in 2012 in the MVs, in 2014 it was reduced to less than 4%. The functionality of water sources has also increased in the CVs when compared to 2012 although at a lower rate than in the MVs (Table 15). Even considering that CVs consist of twice as many villages, the MVs had poor access to water in 2012. The facility survey shows that the situation has improved since both sites have an increased number of boreholes and increasing percentage of functioning waterholes. However, these findings were contradicted in the PRA study which found in early 2015 that significant numbers of MVP-provided boreholes had still not been fitted with pumps and thus were not functional. Water is the biggest problem in one of the RCA MV locations and has worsened since 2013. For instance, despite receiving one new and one rehabilitated borehole, it takes up to 30 minutes to lift water and then people have to wait hours in between. As a result many people have started to revert to using the dam or river water.

Table 15. Water sources, 2012 and 2014

Item	2012		2014	
	MV	CV total	MV	CV total
All water sources	198	612	526	873
Boreholes	70	198	134	246
Functioning water source	134	502	506	785
Percentage functioning	67.7%	82.0%	96.2%	89.9%

Source: Earth Institute Surveys, 2012 and 2014.

74. In terms of road improvements the Earth Institute reported that by June 2014, 60.5km of new or improved roads were completed in the MVs though we do not have data about improvements by community.³¹ The RCA study found that new or upgraded roads were provided in four of the six study villages and in some places this led to an increase in numbers and use of motorised tricycles that are used to transport both goods and people. People complained that some of the roads have

³¹ Annual Report on the Millennium Village Project in Northern Ghana, SADA, December 2014.

already deteriorated due to use by overlaid trucks and tractors, often used by outsiders rather than benefiting the villagers. They also blamed poor workmanship which they said was 'typical of government employed contractors'. New roads were also apparent in CV areas.

75. Table 16 compares educational infrastructure in 2012 and 2014. An important measure of educational provision are the number of pupils per teacher and pupils per classroom, which are expressed as ratios. The education literature emphasises both measures as components that increase children's educational performance. The improvement in the MVs is striking, but needs to be considered with caution since the large gains in student/classroom and student/qualified teacher ratios could be because there are fewer pupils. We only consider classrooms that were assessed to be good (designated by the question asked) as opposed to considering the total number of classrooms. There appears to be some reduction in the number of teachers in the CVs. The 'pupil to qualified teacher ratio' has improved in the MVs while this ratio has worsened in the CVs by a small margin. As both areas saw considerable decline in enrolment numbers this is worth further enquiry at endline. Furthermore, it is not yet clear why there is a decline in the number of qualified teachers in CVs even if it is not by a significant amount.³²

Table 16. Educational resources in 2012 and 2014

Item	2012		2014	
	MV	CV total	MV	CV total
Number of students	6,756	9,676	5,751	8,301
Number of classes in good condition	103	207	192	243
Student/classroom in good condition ratio	66	47	30	34
Total number of qualified teachers	61	140	90	114
Student/qualified teacher ratio	110	70	64	73

Source: Earth Institute Surveys, 2012 and 2014.

76. It is not easy to analyse how medical facilities have changed between 2012 and 2014 given the lack of comparable data. The 2014 MVP report³³ states that 85 health professionals are currently working in the MVP clusters or villages, although the MVs comprise of a smaller area than both the CV areas combined. Two health centres were rehabilitated and reconstructed and three health posts were constructed. The PRA study suggests that this may be due to MVP's deliberate investment in providing on-site accommodation for such staff. The RCA also noted that the policy aimed at bringing midwives out of retirement has had a significant effect, and many of the midwives are more willing to live in the village if accommodation is provided. A few of the MVs (two) and CVs (three) have health facilities on electricity gridlines, making it possible for them to be equipped with refrigerators for storing vaccines and other basic supplies. Maternity wards have been constructed in some posts. Two fully equipped ambulances were deployed.

Table 17. Health resources 2014

Item	MV	CV near	CV far
Physicians	1	0	0
Nurses	30	4	14
Medical assistants	5	0	3
Connection to electricity grid	2	2	1
Number of services provided greater than half the number for which survey was conducted	7	2	2

Source: Earth Institute Surveys, 2012 and 2014.

³² The reason for this was not immediately apparent from the midterm studies and requires further investigation.

³³ Annual Report on the Millennium Village Project in Northern Ghana, SADA, December 2014.

77. The RCA study found that despite the provision of new infrastructure there was much evidence of poor workmanship and deterioration even in brand new buildings, such as classrooms and health facilities. New play equipment for schools was found to be in particularly poor condition and some had already been discarded. There was no knowledge, at the community level, of how routine maintenance should be carried out and whose responsibility it was. The district officials also expressed concern that some of the roads and culverts that were constructed or recently rehabilitated were already beginning to deteriorate, which would not have been the case if they had been tarred.³⁴
78. District officials acknowledged during the institutional assessment that non-MV communities benefit from the new infrastructure, and most of them keenly feel the imbalances created by the MVP with its multiple advantages awarded to a minority of communities. Also, despite much improvement in infrastructure in the MVP communities there is no sense at this stage that a critical mass of the human, financial and material resources has been mobilised within the communities and the districts, and thus the likelihood is that the work will continue to depend on a continual injection of funds from external sources. However, it is worth stating that since the baseline study in 2012, the role of the District Works Department in the MVP is said to have increased: they are now involved in the design of infrastructure projects as well as in their monitoring and evaluation, tasks that are often undertaken by external contractors.
79. The facility surveys show that the MVs seem to have much greater improvements in nearly all indicators for which there is comparable data between 2012 and 2014. The improvements show that as of 2014 people in MVs are likely to have better schools and greater access to healthcare and water than people in CVs, which the opposite was found in 2012. It is concerning that health infrastructure in the CVs has deteriorated considerably within the two-year period, which this may need further exploratory work at endline. Education results are slightly mixed: MV area indicators are now better than the CVs, albeit not dramatically better. Although one of the indicators for the CVs has worsened slightly, the educational indicators measured in 2014 are similar in both sites. Overall, the infrastructure improvements in the MVs are significant. To the extent that comparisons can be made, the facility survey shows that MVs have achieved greater improvements than the CVs and that health improvements are more noteworthy than in other sectors.

3.3 Involvement in project activities

80. This section summarises household involvement in MVP activities, based on the household and facility surveys. It focuses on membership and access to services and facilities rather than the nature of their participation, which is covered in Section 3.4.
81. Tables 18 to 21 show the 'numbers of people involved' in activities promoted by the MVP in social mobilisation, agriculture, health and education compared with control villages. Overall the data suggests that membership of groups and access to activities promoted by the MVP is similar to control areas where the same activities are promoted by the government, local members of parliament (MPs) or by other projects and NGOs.
82. MVP does not have a significant impact over CV on social mobilisation as determined by membership of community-based groups (Table 18), with the exception of farmer-based organisations (FBOs), cooperatives and WASH committees. There are more women's groups and parent teacher associations (PTAs) in MV areas than in CV areas, but the differences are neither large nor statistically significant. Membership of a wide range of other groups promoted by the

³⁴ This statement may not necessarily mean poor quality work on the roads but an appeal for the laterite roads to be tarred.

MVP (including water and sanitation groups, MDG school groups, mother-to-mother support groups, daddies clubs, VSLAs and PTA/school management committees) is very small in both MV and CV areas, and no differences are discernible.

Table 18. Membership of community-based groups (midterm)

	MV	CV	P-value	Observations ³⁵
Cooperative	21.1***	5.2	0.000	2,191
Farmer based organisation	23.9***	5.0	0.001	2,191
Farmer field school	0.8	0.4	0.398	2,191
Women's group	19.2	15.1	0.146	2,191
Parent teacher association	41.0	35.0	0.218	2,191
WASH committees	0.9**	0.1	0.011	2,191
MDG school club	-	-	-	2,191
Water and sanitation development board	1.0	0.2	0.116	2,191
Mother-to-mother support group	0.5	0.1	0.238	2,191
Daddies club	0.0	0.1	0.323	2,191
Village savings and loan association	3.0	2.5	0.725	2,191
School management committee	0.8	0.7	0.781	2,191

*Difference statistically significant at 10%, ** significant at 5%, *** significant at 1%.

83. There is a large and statistically significant impact on involvement in agricultural extension, fertiliser use and receipt of loans in MVP areas (Table 19).

Table 19. Involvement in agriculture-related activities (midterm)

	MV	CV	P-value	Observations
Adults received agricultural training	47.8***	16.4	0.000	5,078
Any household member received a loan	11.8***	1.2	0.000	2,191
Used any fertiliser	55.2***	36.8	0.000	2,158

*Difference statistically significant at 10%, ** significant at 5%, *** significant at 1%.

84. There is a large and statistically significant impact on National Health Insurance Scheme (NHIS) membership and home visits by health workers who provided condoms, advice on breastfeeding, child feeding and use of bed nets and basic child nutrition monitoring (Table 20). Plus, a larger fraction of children are given food supplements in MV areas. However, the distribution of bed nets, clinic usage, and provision of deworming and vitamin A tablets are not greater in MV areas compared with CV areas.

Table 20. Involvement in health-related activities (midterm)

	MV	CV	P-value	Observations
Membership of NHIS	87.0***	55.5	0.000	15,958
Someone distributed bed nets	32.4	36.6	0.452	2,191
Visit by a CHW	65.3***	20.5	0.000	2,191
CHW provided condoms	21.6***	3.4	0.000	2,191
CHW measured children's arms	48.3***	13.9	0.000	2,191
CHW advised on breastfeeding	50.9***	14.6	0.000	2,191
CHW advised on child feeding	55.5***	16.6	0.000	2,191
CHW advised on use of bed nets	54.3***	17.4	0.000	2,191
Visited a health facility	75.0	71.4	0.303	2,191

³⁵ Number of respondents (sample size).

Children given deworming	40.3	35.2	0.124	2,191
Children given vitamin A	40.4	38.9	0.704	2,191
Children given food supplements	9.3***	1.8	0.000	2,191
Children given sanitary pads	2.0**	1.1	0.006	2,191

85. Significantly more children in MV areas receive school feeding than in control areas but there are no significant differences in access to other education benefits such as stationery, books and bursaries (Table 21).

Table 21. Involvement in education-related activities (midterm)

	MV	CV	P-value	Observations
Child had a school meal on previous day	43.7**	27.3	0.027	2,191
Children received a bursary	0.8	0.3	0.104	2,191
Children received stationery, uniform, etc.	24.6	20.6	0.356	2,191

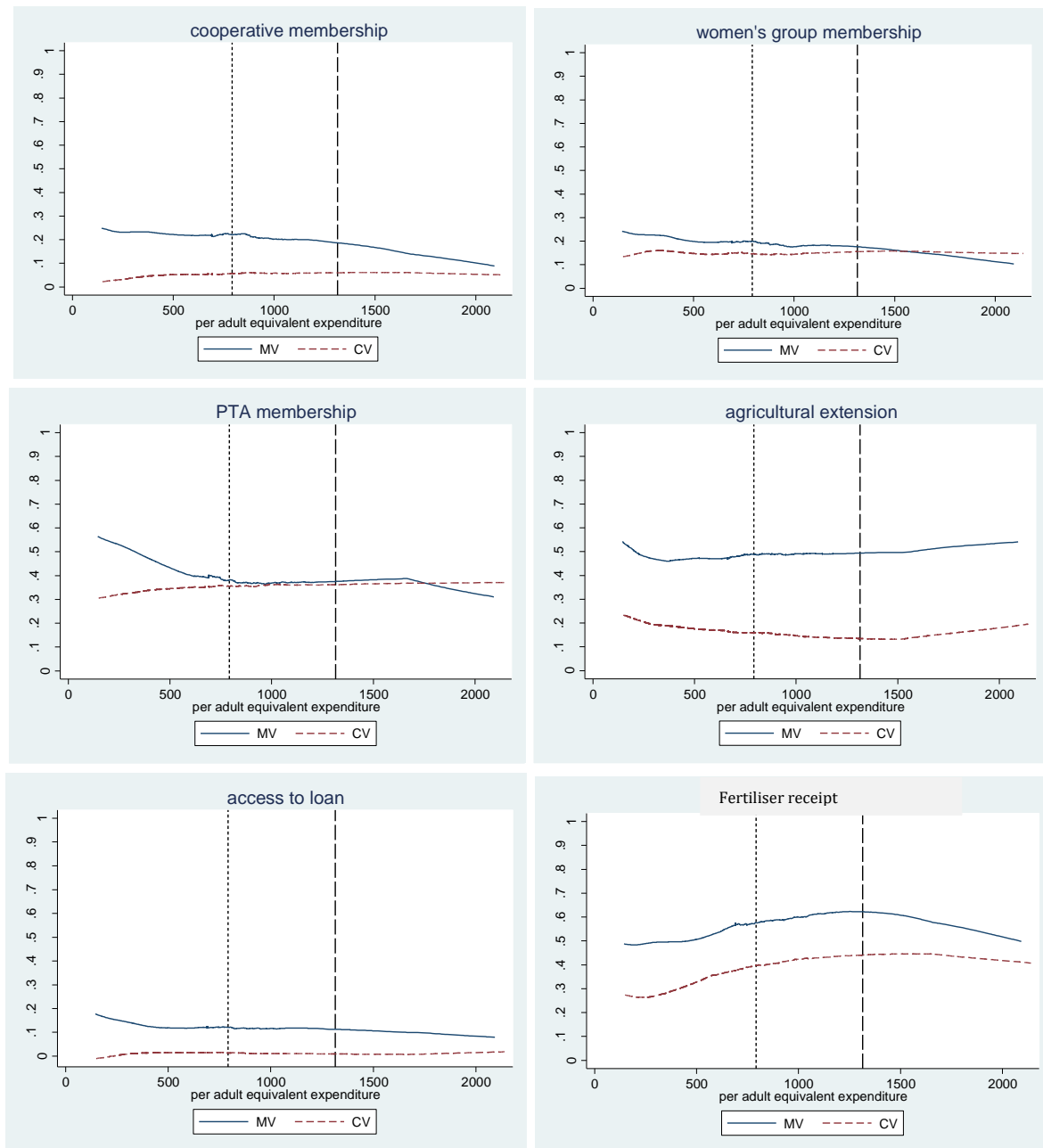
*Difference statistically significant at 10%, ** significant at 5%, *** significant at 1%

86. Next we investigate targeting of the MVP interventions to establish how benefits are distributed across different economic groups, specifically to assess the ability of the intervention to reach the poorest sectors of the population. We do so by plotting the numbers involved in activities in MV and CV areas against baseline per capita expenditure. Figure 3 provides charts of these plots for nine of these activities and Table 22 presents the results of a series of statistical tests that assess the statistical significance of the visual differences observed in these charts. The main findings are as follows:

- The charts show no selective targeting by MVP of households by expenditure levels as the lines are flat over the whole expenditure distribution in most cases. This is confirmed by the null results of Test 2 in Table 22, which shows no pattern in MV access by five expenditure quintiles. The only exceptions are involvement in agricultural extension, which appears to increase slightly with income, and receipt of fertiliser, which slightly decreases with income.
- There are no obvious patterns by per capita expenditure in involvement in activities in the control group either (Test 1 in Table 22).
- As a result of the two points above, there are no differences in patterns between MV and CV with the exception of agricultural extension (less pro-poor in MV compared with CV areas) and fertiliser receipt (more pro-poor in MV compared to CV areas – upper quintiles of the distribution benefit less in MV) (Test 4 in Table 22).

87. Overall there is little indication that the intervention targets poorer people. Rather the programme appears to be directed to the whole population, which in turn has equal access to the interventions.

Figure 3. Involvement in project activities by household expenditure (midterm)



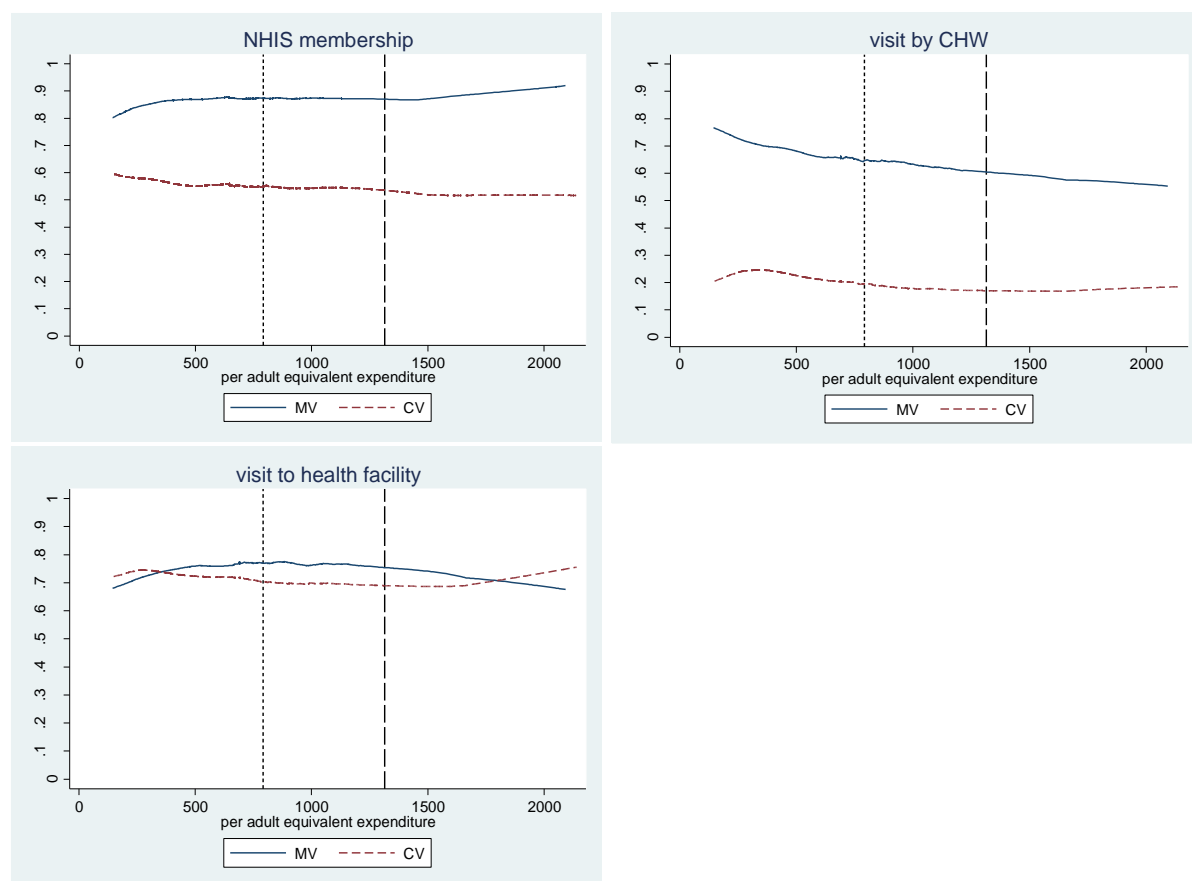


Table 22. F-tests of equality in project involvement patterns (midterm)

	Test 1	Test 2	Test 3	Test 4
Membership of cooperatives	0.47	0.39	20.93***	0.34
Membership of women's groups	0.50	0.87	1.19	0.43
Membership of PTA	0.09	0.91	2.21	1.23
Agricultural extension	1.45	3.05**	36.17***	2.80**
Loans	1.23	0.52	9.16***	1.25
Fertiliser use	1.47	1.66	3.11**	2.15**
Membership of NHIS	1.29	0.32	10.26***	0.51
Visits by CHW	0.84	0.39	20.93***	0.34
Visits to clinics	0.59	0.27	1.26	1.05

Notes: Tests are conducted by running regressions of participation dummies on quintiles of baseline per-adult equivalent expenditure of MV and CV households. Test 1 evaluates the presence of a pattern in CV areas (equality of all CV slope coefficients). Test 2 evaluates the presence of a pattern in MV areas (equality of all MV slope coefficients). Test 3 evaluates the difference between MV and CV areas at each quintile (all MV slope coefficients are zero). Test 4 evaluates the difference in patterns between MV and CV areas (joint difference between MV and CV slope coefficient). *Difference statistically significant at 10%, ** significant at 5%, *** significant at 1%.

3.4 People's experience and view of involvement in MVP activities

88. Drawing extensively on the qualitative evidence from the PRA and RCA studies, this section augments the survey data with insights from local people's perceptions of the project. These insights are important in terms of understanding the quality and perceived value of their involvement and the extent to which local institutions have been strengthened and communities have the will and capacities to sustain gains after the project has ended.

89. The analysis of the survey data shows that MVP has not had a significant impact on social mobilisation in terms of membership of groups, with the exception of cooperative group membership, WASH Committees and farmer-based organisations. These groups have been specifically formed as conduits for agricultural extension, fertiliser use, collaborative farming and access to loans in MVP areas. Both the PRA and RCA studies found that although local people may be members of a farming group, the initial enthusiasm around the provision of agricultural loans by MVP has waned over the project period as the pattern of erratic rainfall persists, undermining crop yields and fuelling a degree of risk aversion among potential beneficiaries. In most MVs, farmers who received the first cycle of credit have been unable to access further credit as a result of defaulting on their loan repayments – influenced, in large part, by rain failure over the last two farming years. It can be seen, therefore, that membership of a group does not necessarily equate to effective participation in its activities or in its expected benefits. The RCA, which concentrates on poorer farmers, found that they had either been refused further loans or had opted out of groups because of indebtedness.
90. Analysis of the survey data also shows a large and statistically significant impact on NHIS membership, which is unsurprising since the project funds NHIS membership, at least in the first years. The qualitative evidence suggests that there may be a tailing-off of membership, and that it is even declining – which raises concerns about the long-term sustainability of the benefits. Both the PRA and RCA studies for example, indicate that since 2013–14 the government has increasingly been requiring subscribers to acquire biometric cards when renewing expired NHIS cards, and there seems to be a falloff in uptake because of the financial and time costs associated with doing this renewal. Furthermore, the PRA team also noted that NHIS renewals are already declining in the MVs following SADA-MVP’s decision to close that line of support.
91. In terms of CHWs, the quantitative survey data reports a large and statistically significant impact on home visits by health workers and their provision of condoms, advice on breastfeeding, child feeding, use of bed nets and basic child nutrition monitoring. The CHWs appear to have been a particularly active part of MVP’s implementation, and although CHWs are providing increased advice in MV areas, there are suggestions that the advice is not being followed in many instances. The RCA study shows that women we met in 2013 in both MV and CV locations who did not know about exclusive breastfeeding now do, and that they got this information from the radio, school and other non-governmental organisation (NGO) programmes. Women would tend to affirm to nurses (and presumably enumerators) that they are doing this, but our observations suggest that other liquids (and some solid foods) are still being introduced early and well before the six-month recommendation. They often affirmed ‘we lie to the nurses because otherwise they tell us off’. The RCA study also saw little change in child feeding practices.
92. Regarding the utilisation of bed nets, both the PRA and RCA studies found that many households either avoid using the nets altogether or else only use them late at night when the ambient temperature has dropped, by which time they may already have suffered bites and, indeed, several participants in the PRA focus groups reported staying up late, thereby exposing themselves to mosquito bites before getting into bed. As in 2013, the RCA study found that none of the households the researchers stayed with actually used bed nets, although they all had them. In virtually every focus group, participants also complained of either itching/burning sensations or sweating when they sleep under the nets, sometimes causing a rash. Such adverse experiences discourage people from using the nets routinely. Some observed a rapid deterioration in the efficacy of the insecticide used on the nets after washing. At one MV site, it was revealed that residents involved in dry season farming by the river sometimes sleep on their farms to keep animals away and to guard their crops against wildfires, leaving bed nets for those in the household who stay behind in the compounds. However, people noted that there were fewer

mosquitos than at the same time in 2013, which was attributed to the long dry season and various large-scale mosquito spraying programmes (that were not implemented by the MVP) in 2014–15.

93. The RCA study also found few people who could tell the team that deworming or vitamin tablets were distributed, although certainly there was a general increase in the use of local health clinics mostly, people said, because they were more accessible and more likely to be staffed.
94. According to the statistical analysis, children in MVs are more likely to receive school feeding but there are no signs of greater access to other education benefits such as stationery, books and bursaries. The PRA-based interviews and RCA study suggest that this is likely because Camfed, Girls–Participatory Approach to Student Success in Ghana (G-PASS)³⁶ and the Global Partnership for Education (GPE) are also rolling out similar educational input initiatives in several control communities and MV areas. However, their interventions do not usually include MVP-type investments in teacher recruitment, handwashing facilities and play equipment. Nevertheless, people in the RCA study felt there was much duplication of incentive programmes for education in MV and CV areas with the same children targeted by different programmes while others, who were also considered in need, missed out. This seemed to be largely because of schools providing organisations with the same lists of eligible children. These programmes that provide uniforms, shoes and school supplies as well as school feeding programmes were not thought to increase enrolment or retain children in school as the basis for school going was more to do with whether the child was regarded as the ‘school type’ as well as older children’s own decisions to leave school to earn their own cash incomes and be accepted as ‘adults’. Corporal punishment and teacher absenteeism are still cited among the main reasons for children to be out of school.
95. In other words, in education too we find that the input gap is closer than one might have expected. The spread of educational inputs in non-MV communities is probably contributing to closing the educational outcomes gap, too. Despite whatever contributions the inputs are making, the fact that learning outcomes remain unimpressive in MVs (but also CVs) are probably attributable, in part, to low levels of teacher competence. Regardless of the superior numbers and more consistent supervision in MV schools, learning will not improve considerably if teachers (both regular teachers and volunteer, CEWs, etc.) lack the skill needed to motivate children to learn and impart knowledge to them. This issue will be explored further during the endline analysis.
96. And finally, the statistical analysis shows that there is no indication that the intervention targets the poorer people, but rather that the programme appears to be directed to the whole population. While this seems to hold true in general, the focus groups interviewed in the PRA and people in the RCA studies noted some de facto (if unintended) exclusion of the poorest households from the MVP credit scheme. This is because applicants must first raise their own resources to finance the ploughing of their farmlands as a precondition for accessing the MVP credit package, a requirement that poorer farmers are less able to fulfil. The RCA came across a few beneficiaries of the special targeting of vulnerable groups who had been given agricultural inputs or livestock. As many of these were elderly, some indicated that they passed these on to others because they could not manage the extra burden, but others were pleased with the inputs.

³⁶ G-Pass is funded by UKAid.

Chapter 4. Achievement of the MDGs

97. This chapter considers the midterm findings through the lens of the MDG indicators. It does not consider qualitative changes nor critique the indicators in any great depth, as this is covered in the next chapter on specific impact topics (Chapter 5). The chapter first presents a sequenced MDG table, which has been simplified – with the full statistical analysis instead presented in Annex A. Some explanations and limitations of each indicator are discussed, but primarily this chapter is about presenting any attributable midterm impact against the MDG indicators largely at face value.

4.1 Impact of MVP on the Millennium Development Goals

98. In this section we illustrate the impact of the project after two years of activities on the UN Millennium Development Goals. Countries have adopted a new and more comprehensive set of development goals known as the Sustainable Development Goals (SDGs). However, we investigate the impact of the intervention on the MDGs rather than on the SDGs for a number of reasons: (i) the achievement of the MDGs is the original project goal and the project activities were chosen and designed to this aim. It seems natural that the impact of an intervention should be first measured against the ultimate goals it set out to achieve;³⁷ (ii) the quantitative survey instruments were largely designed to assess changes in MDGs rather than SDGs or other indicators; (iii) the quantitative impact of the project on the MDGs provides a first approximation of overall project effectiveness; (iv) finally there is substantial overlap between the two sets of indicators as many MDGs were included again into the SDGs.

99. The impacts of the intervention on the MDGs are shown in Table 23. Impacts are expressed as difference-in-differences (DD): before-and-after changes in the outcomes in the MV areas minus the same changes in the comparator group. They represent the impact of the intervention after removing changes produced by other factors unrelated to the project such as government spending, historical trends or natural disasters. All outcomes are reported as ratios, shares, proportions or rates. P-values assessing the statistical significance of the observed effects are reported in parentheses under each DD effect. An additional column with ‘smiley’ and ‘un-smiley’ faces gives a quick qualitative overview of changes that are found to be statistically significant rather than being the outcome of chance. Smiley and un-smiley faces represent statistically significant positive and negative project impacts respectively. When testing multiple outcomes at the same time the probability of finding effects by chance when no real effects occurred increases. Therefore an additional column gives a more conservative overview of the statistically significant effects following best statistical practices.

100. The indicators were developed following the UN instructions on the measurement of MDGs.³⁸ Because of the characteristics of our survey instruments, in some cases our indicators differ slightly from the official UN definitions but great care was taken in reproducing the official methodology. Also, the interpretation of the MDG indicators is not always obvious. Therefore we added a column summarising how the indicator was obtained and what it represents. We first discuss the size of the impact of each family of indicators and their statistical significance using standard approaches, and we then move to discuss what impacts hold after adjusting for multiple testing.

³⁷ See <http://www.unmillenniumproject.org>

³⁸ See <http://mdgs.un.org/unsd/mi/wiki/MainPage.ashx>

Table 23. Project impact on the MDGs

MDG	Difference-in-difference	DD impact (at 10%)	Adj. impact (at 10%)	Indicator
Goal 1 Eradicate extreme poverty and hunger				
Proportion of population below \$1 (PPP) per day	-7.4 (0.132)			The proportion of the population below the international <i>poverty line</i> of \$1.25 a day at purchasing power parity, thus adjusting for cheaper cost of living in Ghana
Proportion of population below the national poverty line	2.3 (0.521)			The proportion population living below the official Ghanaian <i>national poverty line</i> allowing the purchase of a minimum basket of food and non-food items
Poverty gap ratio	-3.9 (0.213)			The mean shortfall of population from the national <i>poverty line</i> . It measures the depth of poverty by calculating how far the poor are from the poverty line
Share of poorest quintile in national consumption	1.7 (0.447) ^a			The share of total expenditure in the study area that goes to the poorest 20% of the population. This is a measure of inequality in the population
Employment to population ratio	5.5* (0.088)	☺		The percentage of individuals older than 15 who did any work, paid or unpaid, over the previous year not including domestic work
Proportion of employed people living below \$1 (PPP) per day	-7.5* (0.095)	☺		The percentage of the employed (as defined above) who are poor. This indicator was developed to provide a measure of the lack of decent work in a country
Proportion of own account and contributing family workers in total employment	4.9** (0.004)	☺	☺	The proportion of the employed population (as defined above), engaged in farming, animal husbandry, fishery or any other self-employment without being remunerated
Percentage of underweight children under 5	0.8 (0.803)			The percentage of children aged 0–59 months, whose weight is below an international benchmark. This a short term indicator of undernourishment

MDG	Difference-in-difference	DD impact (at 10%)	Adj. impact (at 10%)	Indicator
Proportion of population below minimum level of dietary energy consumption	-7.8 (0.159)			The proportion of individuals below the Ghanaian official food poverty line which allows the purchase of a minimum basket of food items
Goal 2 Achieve universal primary education				
Net enrolment ratio in primary education	3.3 (0.282)			The proportion of children of official primary school age (6–11) who are reported having attended primary school at any time during the previous year
Proportion of pupils starting grade 1 who reach last grade of primary	-3.2 (0.304)			The proportion of children aged 11–14 who completed primary among those who ever attended primary school
Literacy rate of 15–24 year olds, women and men	-0.6 (0.674)			The proportion of adults aged 15–24 who were able to read correctly two English sentences ('The child is playing with the ball'; 'Farming is hard work') and to do some arithmetic (9+4 and 4x5)
Goal 3 Promote gender equality and empower women				
Ratio of girls to boys in primary education	-0.09 (0.292) ^a			The ratio of the net attendance rate in primary school of boys and girls aged 6–11. A ratio below one implies fewer girls are attending primary than boys
Ratio of girls to boys in secondary education	0.15 (0.860) ^a			The ratio of the net attendance rate in junior secondary school of boys and girls aged 12–14. A ratio below one implies fewer girls are attending primary than boys
Ratio of girls to boys in tertiary education	-0.63 (0.761) ^a			The ratio of the net attendance rate in senior secondary school of boys and girls aged 15–18. A ratio below one implies fewer girls are attending primary than boys
Share of women in wage employment in the non-agricultural sector	-24.2** (0.020)	⊖		The proportion of women above 15 in overall employment in the non-agricultural sector. The indicator measures to what extent women have equal access to jobs outside agriculture
Goal 4 Reduce child mortality				

MDG	Difference-in-difference	DD impact (at 10%)	Adj. impact (at 10%)	Indicator
Under-5 mortality rate	-1.12 (0.514) ^a			The child's probability of dying before 5 years of age, calculated per thousand of population over the 5 years preceding the interview using the DHS method
Infant mortality rate	-0.09 (0.953) ^a			The child's probability of dying before 12 months of age, calculated per thousand of population over the 5 years preceding the interview using the DHS method
Proportion of 1-year-old children immunised against measles	0.6 (0.890)			The proportion of children aged 0 or 1 whose vaccination card reports a measles vaccination or whose mother recalls the child being given an injection in the upper arm to prevent measles
Goal 5 Improve maternal health				
Proportion of births attended by skilled health personnel	16.7** (0.002)	☺	☺	The proportion of deliveries assisted either by doctor, clinical officer, nurse, midwife, or CHW for all children of age 0–2 at the time of the interview
Contraceptive prevalence rate	4.1** (0.030)	☺		The proportion of women aged 15–49 using any contraceptive method at the time of the interview (sterilisation, pill, IUD, injections, implants, condoms, rhythm, abstinence and withdrawal)
Antenatal care coverage	-4.2 (0.127)			The percentage of women aged 15–49 who received at least one antenatal visit (doctor, clinical officer, nurse, midwife, CHW) for children who are aged 0–2 years at the time of the interview
Goal 6 Combat HIV/AIDS, malaria and other diseases				
Proportion of population aged 15–24 with comprehensive correct knowledge about HIV	0.1 (0.953)			The proportion of population aged 15–49 who answered correctly 8 (yes/no) questions about obvious causes of HIV infection transmission
Proportion of children under 5 sleeping under insecticide-treated bed nets	33.0*** (0.000)	☺	☺	The proportion of children aged 0–59 months who slept under an <i>insecticide-treated mosquito net</i> the night before the interview

MDG	Difference-in-difference	DD impact (at 10%)	Adj. impact (at 10%)	Indicator
Goal 7 Ensure environmental sustainability				
Proportion of the population using an improved drinking water source	-7.6** (0.042)	☹		The percentage of households whose main source of drinking water is: piped into dwelling, yard or plot; public tap; tube well and borehole; protected dug well; protected spring; bottles; sachet water
Proportion of the population using an improved sanitation facility	0.7 (0.754)			The percentage of households that normally uses toilets: flush to piped sewer system; flush to septic tank; flush to pit (latrine); ventilated improved pit latrine; pit latrine with slab
Goal 8 Develop a global partnership for development				
Fixed telephone subscriptions for 100 inhabitants	0.0 (0.997)			The percentage of households reporting having a landline in the home
Mobile cellular subscriptions for 100 inhabitants	-7.5 (0.201)			The percentage of adults aged 15–49 reporting a personal use of a mobile phone during some or all the year before the interview

All figures are percentage changes. For example, -7.4 means that poverty decreased by -7.4 absolute per cent points in MV areas compared to CV areas. P-values in parentheses. ^a P-values calculated using 1,000 bootstrap sample replications at the cluster level. In all other cases P-values are obtained from cross-sectional regressions of the outcome (0/1) on project dummies and using cluster standard errors. *Difference statistically significant at 10%, ** significant at 5%, *** significant at 1%.

101. MDG 1 is *Eradicate extreme poverty and hunger*. We found little impact on overall poverty although a larger impact is visible on extreme poverty (proportion of population below \$1 (PPP) per day and proportion of population below a minimum level of food consumption). The negative impact on the poverty gap ratio and the positive impact on the share of expenditure going to the bottom quintile of the population seem to confirm that the intervention is helping the extremely poor. None of these impacts are, however, statistically significant.

102. The project has not reduced the proportion of underweight children. However, it has had some impact on employment patterns. The proportion of individuals performing any work has increased as well as the proportion of own account workers in total employment (the two ratios represent almost the same phenomenon because the predominant activity in the area is subsistence agriculture) and poverty among the employed has decreased. Poverty rates are, however, sensitive to the poverty line used. In particular, most households are poor based on the national poverty line and changes in expenditure have little impact on poverty rates based on the official poverty line. On the other hand, the poverty headcount changes to a larger extent when using the \$1 PPP poverty line, which is also used in the MDGs that measure poverty among the employed. The decrease in the proportion of the poor among the unemployed is therefore not in contradiction with the lack of change in poverty as measured using the official poverty line. Rather, it is related to changes in poverty using the \$1 PPP poverty line. Indeed, changes in overall poverty based on \$1 PPP and changes in poverty among the unemployed based on the same

criteria appear to be very similar. The impacts observed on employment indicators are statistically significant.

103. The second MDG is *Achieve universal primary education*. We found no impacts in relation to this goal. Project effects on net attendance of primary school, on the proportion of children completing primary school and on literacy rates of adults aged 15–24 are very small and never statistically significant.
104. The third MDG is *Promote gender equality and empower women*. In the study area more girls than boys are attending school at all education levels (primary, junior and secondary high schools). The project has had no impact on this ratio. There is a considerable, and statistically significant, reduction in the share of women in wage employment in the non-agricultural sector in the MV area. However, the proportion of the population employed in the non-agricultural sector is very small so that the absolute impact on employment of this change is very small. This reduction in wage employment could be the result of a reallocation of women's labour across activities. As agricultural work becomes more productive because of the MV intervention, women are increasingly moving away from non-agricultural activities and investing more time in farming. A confirmation of an increase in the family time spent on farming is the 5% increase in the proportion of the population of either sex employed in agriculture (Proportion of own account and contributing family workers in total employment, Table 23, Goal 1). An analysis disaggregated by sex shows that women's employment in agriculture increased by 6 percentage points so that women are increasingly engaged in the economic activity favoured by the intervention.
105. There is a reduction in both infant and under-5 mortality by about 1 percentage point. These impacts are calculated over a five-year interval, which heavily underestimates the actual impact because the programme has been running for only two years. The differences are not statistically significant, partly because of the relatively small sample size employed. We found no impact on the percentage of children immunised against measles.
106. Goal number 5 is *Improve maternal health*. The project has a large and statistically significant impact on the number of births attended by a skilled professional, on the use of birth control methods and on the proportion of children reported to be sleeping under insecticide-treated mosquito bed nets. We found no impact on antenatal care and on HIV knowledge.
107. Goal 6 is *Combat HIV/AIDS, malaria and other diseases*. The quantitative analysis found that the proportion of children under-5 reporting to be sleeping under insecticide-treated bed nets increased in the MVs compared to the CVs, while the RCA found otherwise. Part of an explanation could be the seasonal effect on bed nets. Baseline data were collected during dry season in the MVs thus showing a large increase, while baseline data were collected during rainy season in the CVs thus showing little change.
108. There are a number of observational explanations highlighted in the RCA study. First, families know that when they are asked whether they sleep under bed nets the expected answer is "yes" since they have been given them to sleep under and in the MV they would have been directed to. Second, it may be seasonal since it was extremely hot at the time of the RCA and many families did not sleep inside their houses. When sleeping outside people are not inclined to string up bed nets and it is also too hot to do so (RCA researchers confirm this). Third, a further explanation related to seasons is that mosquitos may have been a nuisance (people tell RCA researchers they use bed nets because mosquitos are a nuisance so the bed nets are helpful in reducing noise and itching rather than protecting against malaria) during the survey period. This coincided with the massive mosquito spraying programme this year. However, conversations during the RCA suggest

that families rarely use bed nets during any season and, as was found in 2013, some of the distributed bed nets were never unwrapped. New mothers are provided with bed nets as part of the antenatal check-up process and, as with other instructions from nurses, always indicate that they use them for fear of being reprimanded by nurses.³⁹ This is confirmed by the survey which shows much higher levels of advice being given about use in MV but this not translating into use.

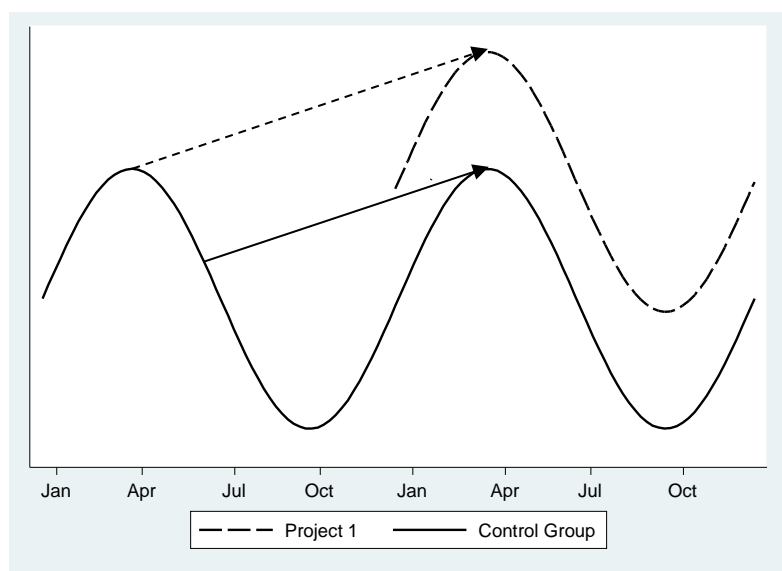
109. MDG 7 is *Ensure environmental sustainability*. We found no impact on the proportion of households using an improved toilet facility, while we found a negative impact on the proportion of households having access to safe drinking water. The latter effect is large and statistically significant, implying that access to safe water sources has worsened in MVP areas relative to the CV areas.⁴⁰
110. MDG 8 is *Develop a global partnership for development*. The project has not increased the use of mobile phones. No household owns a landline phone in the area and no impact is observed on use of landlines.
111. To summarise, the data suggests that there was a drop in extreme poverty and a visible increase in employment in the MV areas. The data are also suggestive of a reduction in child mortality, although the observed effect is not statistically significant, and there is a visible increase in births attended by skilled personnel, and in the use of contraceptive methods and mosquito bed nets. We found no impact of the intervention on: undernutrition, school attendance, girls' school attendance, measles vaccination rates, antenatal care, use of safe toilets and use of mobile phones. Finally, the data seems to suggest that access to safe drinking water has worsened as a result of the intervention.
112. This analysis is potentially affected by what statisticians call the 'multiple comparisons problem'. Since many hypotheses are tested at the same time, there is a high probability that some of the statistically significant effects found are the result of chance rather than real. Statistical tests are usually conducted by setting the probability of wrongly rejecting the null hypothesis of no impact at 5%. An implication of this choice is that there is a 5% probability of finding an impact when there are none at each test. As we increase the number of tests, the probability of finding at least one statistically significant impact when in fact there are none grows. For example, with 28 tests, as in Table 23, the probability of finding at least one statistically significant result when the project has no impact whatsoever is 76%. Since the probability of finding a statistically significant effect increases with the number of tests, researchers may be led to do additional tests, with more variables splitting the sample into different groups until a statistically significant result emerges by pure chance. This heuristic process of obtaining results is commonly referred to as "fishing", "data mining" or "torturing the data until they confess". In order to avoid this problem, we applied two common statistical procedures: a Bonferroni Adjustment and the False Discovery Ratio. The results of the latter adjustment are reported in the fourth column of Table 23. The results of adopting this approach are represented by the smiley faces in column 4. Only three project impacts are found to be statistically significant after adopting more conservative approaches: (i) the increase of own account and contributing family workers in total employment; (ii) the increase in the proportion of births attended by skilled health personnel; and, (iii) the increase in the proportion of children sleeping under insecticide-treated bed nets.
113. Overall, the table suggests that the impact of the intervention after two years is limited. This analysis is, however, affected by a number of limitations. These include:

³⁹ This is similar to mothers also claiming to nurses that they exclusively breast feed or do not give babies water.

⁴⁰ Note: From the PRA and RCA studies, it seems that the functioning/maintenance of water sources is the key problem. Plus, it has been exceptionally dry, although the intervention should have taken this into consideration.

- **Selection bias.** Impacts are estimated as if the observations had been obtained from a randomised experiment. These differences are not adjusted for baseline differences in characteristics between project and comparator areas. After adjusting for these differences the programme effects may also differ.
- **Seasonality.** While midterm data were collected simultaneously, baseline data were collected at different times of the year. In particular, data in MV villages were collected between 16 April and 14 June 2012 (height of dry season) and CV data were collected between 1 August and 18 September 2012 (height of rainy season). Many outcome variables are correlated with rainfall including: amount of mosquitos, incidence of diarrhoea, food availability, and expenditure patterns. The result is that estimated project effects can be biased. The figure below illustrates this. In this example, the midline data were collected in April for both the project and control areas, whereas, the baseline data were collected in the control areas two months later than in the project areas. Given the seasonal pattern of the outcome variable, a difference-in-difference analysis will find no impact, when in fact the project has a visible impact. Using the same figure it could be shown how the evaluation could find a positive or negative impact, when in fact there is no impact.

Figure 4. Project impact with baselines at different times in project and control areas



- Appendix H to the Baseline report⁴¹ contains an extensive discussion of these issues. In order to assess the risk of seasonal bias in the estimation of project effects we looked at seasonal patterns of anaemia, malaria, anthropometry, income and expenditure using secondary data and by reviewing the existing literature. We found little support for seasonality of anthropometric data. Height can hardly be sensitive to rainfall, while the impact of rain on disease does not seem to be sufficient to compromise a child's weight. On the other hand, the absorption of iron is lower in the rainy season, while mosquitos are more common and bed nets are more commonly used. Finally, income and expenditure patterns are greatly affected by seasonal patterns, but income and expenditure data were collected on an annual basis and are not affected by any seasonal bias. To make this more tangible, consider at baseline that data were collected in MV areas during the dry season, when mosquitos are rare and bed nets are not used, while data were collected in CV areas during the rainy season when mosquitos are common and so is the use of

⁴¹ See: <http://www.ids.ac.uk/publication/millennium-villages-impact-evaluation-baseline-summary-report>

bed nets. Now consider that midterm data were collected in the rainy season in both the MV and CV areas. In these conditions, a comparative analysis of changes in the two areas will show a positive impact of the project on use of bed nets even if the project had no impact at all.

- **Sample sizes.** Some of the MDG indicators are calculated using small groups because of age restrictions in defining the relevant samples. Small sample sizes result in large standard errors and a reduction in the likelihood of finding statistically significant small effects. Other indicators are based on large samples but because the results are small, the standard errors are nevertheless large: this is particularly the case of child mortality rates. Absence of statistical significance for these variables may reflect a lack of impact as well as a small sample size.
- **Inadequacy of MDG indicators.** The official MDG indicators do not describe accurately socio-economic conditions in the area and are unable to fully detect the impact of the project on living standards. For example, employment-related indicators have little relevance in a place where most individuals do some work at any time. On the other hand, relevant indicators of well-being such as prevalence of anaemia and malaria or cognitive skills and maths test scores are missing. In addition, MDG indicators based on prevalence rates pay little attention to changes in the distribution of the outcomes so that, for example, a reduction in the percentage of extremely malnourished children cannot be observed.

114. The analysis conducted in the following sections of this report will address these limitations. In particular, comparisons will be made for non-MDG relevant welfare indicators and all comparisons will be performed using matching methods thereby adjusting for baseline difference in characteristics between the MV and CV groups.

Chapter 5. Impact

115. This chapter further explores the impact of MVP by looking beyond the MDG indicators, using a diverse set of variables to assess change. The first part looks at monetary poverty and the extent to which MVP is breaking the poverty trap. The chapter then goes on to consider the impact of MVP on income and food security, child health, anthropometry, anaemia and education. The chapter ends by considering the heterogeneity of impact.

5.1 Impact of the MVP on monetary poverty

116. We calculated monetary poverty rates using the methodology and the poverty lines used by the Ghana Statistical Service (with the full procedure set out in Annex A). Monetary measures of poverty are based on consumption–expenditure data at household level which includes purchases as well as own-produced items and in-kind gifts from other households. Consumption is an overall measure of well-being as it includes the consumption of food, education items, health, non-food items and an evaluation of the value of the use of durable goods and housing. Households that fail to consume a minimum basket of essential goods are classified as poor. However, in the qualitative strands which complemented the survey, households employed more diverse, locally relevant descriptors of poverty/well-being in analysing changes in their conditions.

117. Poverty rates for the first three survey rounds are reported in Table 24 for MV and CV areas, respectively. We show two sets of poverty rates: (1) overall poverty, calculated in relation to the ability to consume a minimum basket of food and non-food items (the ‘general poverty line’); (2) food poverty, calculated in relation to the ability to consume a minimum basket of food items (the ‘food poverty line’). The first poverty rate is an indicator of overall poverty while the second is an indicator of extreme poverty. All poverty indices are very similar in MV and CV areas at the baseline. Overall poverty has not decreased at the midterm in MV areas, although food poverty shows a noticeable⁴² decrease of more than 10 percentage points.

Table 24. Poverty indices

	Overall poverty		Extreme poverty	
	MV	CV	MV	CV
Baseline	0.863	0.874	0.652	0.647
2nd round	0.915	0.911	0.686	0.672
Midterm	0.883	0.871	0.548	0.620

Local perceptions of well-being

118. In terms of people’s own perceptions of their poverty status, the majority of households typically retained their ‘ranks’ (e.g. poor, very poor, rich, very rich), although the PRA exercise also identified some households who had either moved up or down the well-being ladder. This is based on local people’s own (subjective and composite) well-being criteria. The most dominant criterion by which participants distinguish well-being from ill-being continues to be the degree to which annual harvests provide an assurance of year-round food security. Other significant markers of well-being – in their subjective assessment – include the range and/or quantity/quality of tangible assets (such as housing, motorcycles, livestock, bullock ploughs and farmlands); personal health and educational attainment; ability to recover financially after experiencing adversities (such as death in the family, rain failure or flooding); the degree to which they have control over their

⁴² This section is not discussing the DD effects but simply the rates set out in Table 24. As such, a 10 percentage points decrease in just two years is quite large in poverty analysis terms, and hence regarded as ‘noticeable’. The DD effect is, however, not statistically significant for poverty (as set out in the previous table in Chapter 3).

destinies (e.g. ability to hire labour in lieu of using own children to perform productive tasks); and dignity (often reflected in the ability to engage effectively in community life). Based on the communities' criteria for describing well-being, the study found that households headed by women, disabled persons and chronically ill parents were often assessed by local informants as experiencing multiple deprivations. As a result, such households were often classified among the poorest. However, when one compares the current situation with the baseline, an increasing number of these households have experienced marginal improvements in their conditions. The improvements are accounted for mainly by the fact that the Livelihood Empowerment Against Poverty (LEAP) programme has been extended to more communities in the study area. Further, the quality of targeting is also reported to have improved, with the screening process becoming much more inclusive and fair. For now, however, many of the LEAP stimulated improvements have not been large enough to prompt a re-categorisation in the well-being rankings.

119. Table 25 attempts to illustrate how each MV compares with its nearby and faraway controls, based on this well-being ranking activity. We found that improvements in well-being were not substantially greater in MVs than in the faraway control villages. The improvements were also stronger in the faraway control villages than in the nearby controls, raising questions about the MVP's propagative effects on nearby communities.
120. Regarding changes in well-being outcomes, we also identify from the focus group discussions two main explanations for the closeness of the gap between MVs and CVs. First, the Presbyterian Agricultural Station (PAS) is delivering agricultural support – similar to what MVP facilitates for the MVs – in several control communities. This support typically includes improved seed, subsidised traction services and agronomic know-how. Second, many of the immediate gains in the non-MV communities are driven by perverse livelihood strategies – e.g. indiscriminate logging of trees for charcoal production, over-extraction of *chewing sticks*⁴³ from the grassland's shrubs, illegal gold mining and uncontrolled extraction of sand from the river banks and river beds – which are not sustainable over the long term. In many instances, the focus groups reported the involvement of children in such activities.
121. The RCA study found that nearly equal proportions of study households (both MV and CV) felt that they were better off or worse off than two years ago. Those feeling 'better off' attribute this mostly to having more income-earning members in the family (especially ones who could migrate for work) and good Bambara bean harvests. These families had acquired some new assets. Those feeling 'worse off' attribute this to poor harvests, consequent debt and more dependents (non-working people) in the household. Households are fluid in terms of membership, with migration for work having a significant effect, including leaving young children with grandparents. Migration for work seems to have increased according to our conversations, particularly with the increasing unpredictability of farming and late rains. In another RCA study conducted in central Ghana we met many migrants from this region living in large slums who explained to us that they were sending money back to support their families, especially because of the late rains.⁴⁴ The increasing need for cash has resulted in all the study families, except those that are fully dependent on others, to have diversified their income-earning activities and all now have between one and five additional cash earning strategies beyond farming (which is mostly undertaken to feed the family and rarely cash earning).

⁴³ Herbal toothbrushes sold on the local market.

⁴⁴ 2015, Adolescents' Views on Sexual and Reproductive Health in Ghana's Brong Ahafo Region, <http://www.reality-check-approach.com/ghana.html>

Table 25. Well-being: changes in rankings since baseline

Comparator cluster number ⁴⁵	Millennium village (treatment)		Control village (nearby)		Control village (faraway)	
	Number of HH who have changed rankings ⁴⁶	Exits from poverty (versus falls into poverty) ⁴⁷	Number of HH who have changed rankings	Exits from poverty (versus falls into poverty)	Number of HH who have changed rankings	Exits from poverty (versus falls into poverty)
1	6 improvements (6 declines)	6 exits (6 falls)	3 improvements (1 decline)	2 exits (0 falls)	7 improvements (1 decline)	6 exits (1 fall)
2	12 improvements (0 declines)	6 exits (0 falls)	1 improvement (4 declines)	0 exits (0 falls)	4 improvements (1 decline)	4 exits (1 fall)
3	5 improvements (7 declines)	5 exits (5 falls)	9 improvements (3 declines)	6 exits (0 falls)	1 improvement (2 declines)	1 exit (0 falls)
4	32 improvements (0 declines)	20 exits (0 falls)	1 improvement (0 declines)	0 exits (0 falls)	46 improvements (3 declines)	13 exits (0 falls)
5	16 improvements (12 declines)	0 exits (0 falls)	13 improvements (53 declines)	0 exits (0 falls)	No matched faraway control village	
6	17 improvements (6 declines)	9 exits (1 fall)	8 improvements (27 declines)	6 exits (7 falls)	8 improvements (16 declines)	3 exits (7 falls)
7	12 improvements (10 declines)	5 exits (2 falls)	11 improvements (13 declines)		27 improvements (6 declines)	0 exits (3 falls)

* For each row (representing a comparator cluster), the cell(s) in bold text indicate which communities have been most effective at making well-being gains since the baseline.

Consumption and expenditure distribution

122. The assessment of project impact based on changes in poverty rates from the statistical analysis is blind to changes occurring in the distribution of outcomes at different levels of the consumption distribution. The project may improve living standard of the poor, or of some sections of the poor population, without affecting overall poverty rates. In order to remove this problem we assess the impact of the intervention first by looking at changes in average per-adult consumption rather than looking at changes in poverty, and second by looking at changes in inequality and at the impact for households at different expenditure levels.

123. Table 26 shows the average treatment effect of the project on per-adult equivalent expenditure using three different model specifications and adjusting for baseline differences in characteristics between MV and CV households. The intervention has a very limited impact on per-adult equivalent expenditure. The effect size is an average 3% increase in real terms per year, which disaggregates into a 0% impact in the first year and an impact of about 6.5% in the second. These

⁴⁵ The clusters are matched MVP communities with nearby and faraway control villages. The matching was undertaken as part of the baseline sampling strategy. See: Initial Design Document, 2013.

⁴⁶ The well-being rankings used are rich, moderately rich, poor and very poor. Using a participatory exercise, respondents define each according to a community classification. This indicator shows the number of households that have moved from one ranking to another.

⁴⁷ The term 'exit' is used as a shorthand, but poverty can be cyclical and so this may be a temporary exit. This indicator is based on people's perceptions, and it captures those that move ranking in (and out of) poverty. This is defined as those that move between the rankings of the 'richest' (i.e. rich or moderately rich) and the 'poorest' (i.e. poor or very poor).

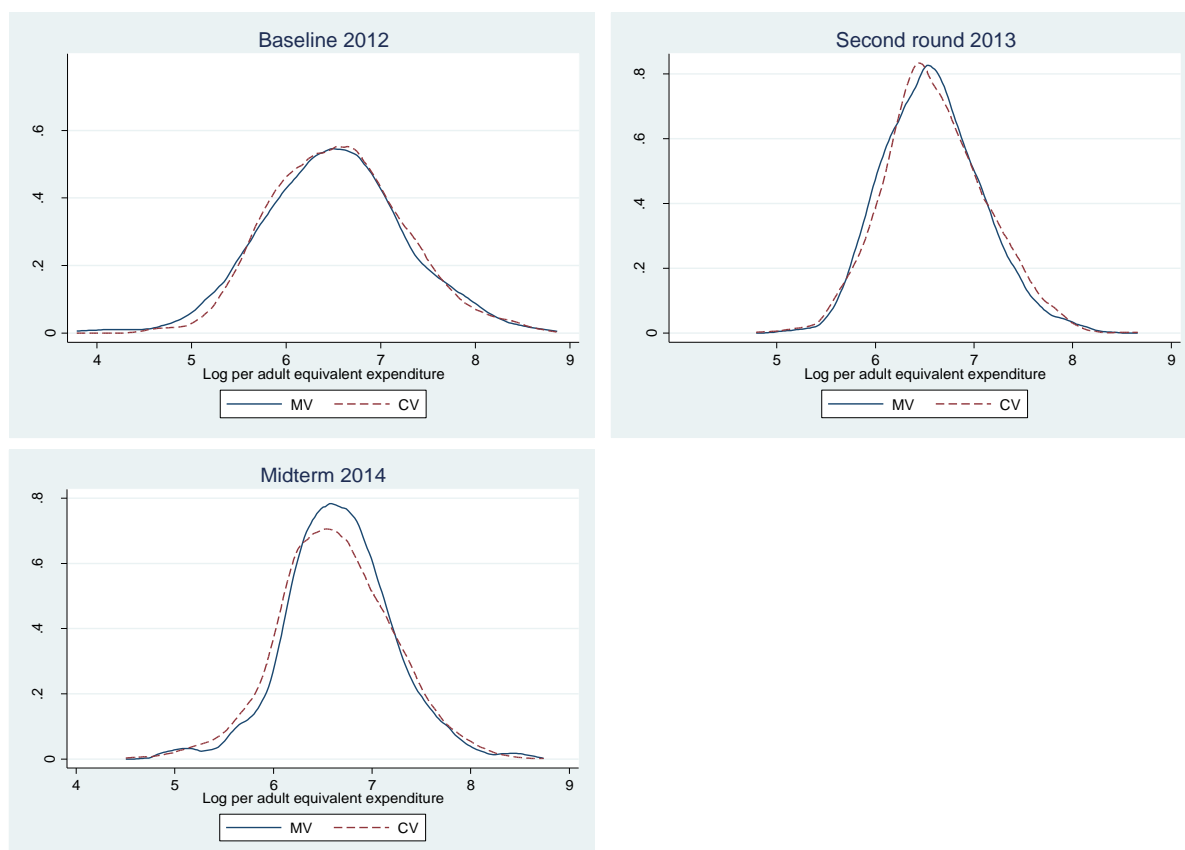
effects are never statistically significant, although it must be noted that our sample was not designed to detect statistical significance of extremely small effects such as those reported in the table.

Table 26. Difference-in-difference impact on per-adult equivalent expenditure

	Cross-section	Fixed effects	Lagged model
Average DD effect	0.031 (0.618)	0.029 (0.644)	0.024 (0.569)
DD effect second year	-0.005 (0.938)	-0.007 (0.922)	-0.009 (0.829)
DD effect third year	0.069 (0.324)	0.066 (0.347)	0.057 (0.337)
Sample size	5942	5942	3942

124. The distributions of per-adult equivalent expenditure over the three rounds in MV and CV areas are plotted in Figure 5. The distributions are nearly identical at baseline and for the second round, but there is a shift to the right of the expenditure distribution in the MV areas at the midterm. Since the increase in expenditure mostly occurs below the poverty line this does not translate into a reduction in the poverty headcount.⁴⁸ The RCA study found many families had increased debt compared to 2012. This means that while material living standards may have improved, people often admitted that they felt scared and stressed about their futures. Some had discovered that they could no longer access credit from official sources or family and neighbours and some had had assets repossessed for non-repayment of loans. Others spoke about not being able to meet social obligations because they are in debt. Some people also divulged to the team that since they are ashamed of their debt they do not tell anyone about it (presumably including enumerators). Further conversations revealed that increased indebtedness raises concerns about how they can access NHIS, support school costs, pay for toilets and electricity connections and buy agricultural inputs, which are all activities supported by MVP.

⁴⁸ Note that poverty analysis based on expenditure is concerned with the material living standards, with higher expenditure regarded as an indication of an improved condition. The analysis is neutral to the source (whether income, debt, donation, lotteries, inheritance, etc.). Whether expenditure is funded by debt or loans is considered irrelevant here.

Figure 5. Densities of log per-adult equivalent expenditure in the project and control groups

125. We then assess the impact of the intervention for households at different levels of the expenditure distribution. These effects are reported in Table 27 at selected percentiles from lower (meaning poorer) to higher (meaning richer), as we move from columns left to right. The impact of the intervention decreases with household per-adult equivalent expenditure. Though none of the coefficients reported in the table are statistically significant, the MV increases the expenditures of household in the bottom 50% of the expenditure distribution and does not affect households above this threshold. Note that in Table 27, as well as in Figure 5, we are simply comparing the distribution of outcomes in the MV and CV areas. These estimates do not provide information on how people at different levels of the distribution are affected by the project. Different households are affected by the project to a different extent and some households might even be negatively affected. A change in the distribution of expenditure is consistent with an extremely large number of individual changes in every direction and since it is impossible to observe the same person 'with' and 'without' the project, these effects cannot be observed directly (Heckman et al. 1997). In other words, these estimates allow us to conclude whether the distribution of expenditure has become more equitable because of the project, but does not allow us to conclude whether, for example, the intervention benefited some particularly poor segments of the population. The same applies to the distributional effects discussed in relation to Figure 6 below and to the similar discussion of distributional effects on incomes discussed in relation to Figures 7 and 8 and Table 27 below.

126. The RCA study purposively includes poorer households and found very little or no accumulation of assets (in several cases fewer assets) and basically no changes in food consumption patterns. The only household that seemed to have improved its overall economic situation was from a CV village and they attributed this change to their successful Bambara bean harvests in 2013 and 2014. Other families' asset accumulation (solar panels, radios, mobile phones and in one case a

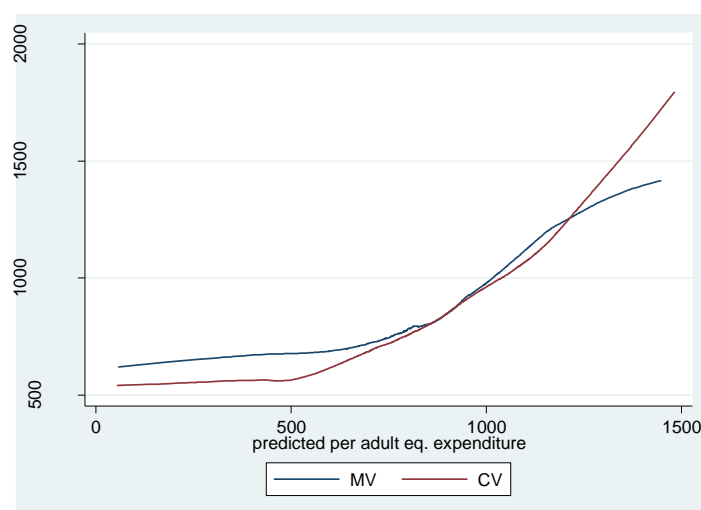
TV) was mostly due to gifts from migrant worker relatives. The research also revealed that people have very little knowledge of the monetary value of gifts they receive. Both the RCA and PRA studies further confirm what was beginning to be apparent in 2013: that people spend large sums of money on funerals which are delayed, sometimes many years, until they can be afforded. This would suggest that even if there are increases in income, these will not be spent on food consumption for the family or on increasing household assets.

Table 27. Distributional impact on per-adult equivalent expenditure

	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
Average DD effect	0.038 (0.493)	0.008 (0.864)	0.008 (0.877)	-0.001 (0.987)	-0.003 (0.966)
DD effect second year	-0.009 (0.861)	-0.028 (0.623)	-0.010 (0.843)	-0.029 (0.594)	-0.014 (0.818)
DD effect third year	0.137 (0.179)	0.068 (0.416)	0.088 (0.266)	0.015 (0.830)	-0.007 (0.916)

127. We conducted some additional analysis of the impact of the intervention across the expenditure distribution and we obtained similar results. We employed non-parametric regression as in Barrera-Osorio et al. (2011). We plot per-adult equivalent expenditure at the midterm for the project and control groups against the predicted values obtained from a regression of midterm per capita expenditure on poverty determinants using control observations only (Figure 6). As in Table 27, the impact of the project on expenditure is largest at the lower end of the expenditure distribution and nil for richer households.

Figure 6. Per-adult equivalent expenditure by predicted expenditure



Source: Itad's elaboration

128. In summary, we find that the project had a very small impact on consumption in the first two years and that much of the impact appears to have occurred in the second year of operation. We also find that the project affected households in different ways. In particular, the impact was larger for extremely poor households (notwithstanding RCA and PRA observations around debt and little accumulation of assets) and nil for households classified as non-poor. Note, however, that none of the observed differences are statistically significant.

5.2 Impact of the MVP on income and food security

129. The quantitative analysis of the expenditure data above showed that the project had a modest impact on consumption and that this impact was larger for the poorer sectors of the population. In this section we investigate the size of the impact of MV on households' incomes and we try to disaggregate this impact by income sources. The economy in the study area is dominated by mostly self-sufficient agricultural and livestock production. Agricultural production and livestock rearing together make up nearly 80% of household income (see Table 28). Some 7–15% of income is obtained from small business activities such as petty trade, microenterprises and community services. Less than 5% of household income is generated by wage employment in the non-agricultural sector and the fraction of income obtained from transfers and remittances is negligible (less than 1%). There is a noticeable change in the composition of income sources over time. The share of agricultural income has decreased at the same time as livestock and business income have increased. It is difficult to interpret these patterns. They might be the result of agricultural shocks related to rainfall patterns, the impact of MV on different economic activities or an improvement in enumeration. The focus group discussions from the PRA study confirm that, overall, rain failure and unpredictable rain has undermined agricultural production since 2012, even in MVs – although some communities did experience net gains in the form of relatively larger harvest volumes.

Table 28. Income shares by source, all households

Income source	Baseline	1st round	Midterm
Agriculture	56.0	45.1	42.1
Livestock	28.4	30.1	32.2
Business	7.1	16.5	16.9
Employment	4.5	4.3	5.5
Transfers	0.8	0.8	0.8

Note: Some incomes are negatives because costs exceed revenues and as a result shares do not add to 100%.

130. We estimate the impact of MV on household income using the same approach employed in the estimation of the impact on household expenditure. Since a large fraction of incomes is negative we are unable to use logarithms. We have therefore decided to standardise the income figures by the baseline common standard deviation of income. The latter was approximately 1.5 times average income at baseline so that changes in standard deviations can be translated to changes with respect to the mean with little effort. As before, we use three different DD regression models: simple cross-section estimates, fixed effects and lagged dependent variable models. The estimation employs the same IPW used in the expenditure models, and the DD regressions include determinants of income that may differ across the MV and CV groups.

131. The results are surprising (Table 29). Even the most conservative estimates of the lagged model suggest an increase in household income by 0.2 standard deviations corresponding to an increase of about 30% per year and all the effects are highly statistically significant. These results contrast with the small effects observed on household per-adult equivalent expenditure. At least three possible explanations are in order for the simultaneous large impact on income and the small impact on consumption.

- The project has a larger impact on agricultural profits than on non-farm activities such as petty trading and incomes from employment. If non-agricultural incomes are unreported, the impact of MV on total household income might appear larger than it actually is. Suppose, for example, that the project increases agricultural incomes by 50% but has no impact on non-agricultural income. Now suppose that agricultural income is 50% of total income but because of under-reporting of non-agricultural income, the reported agricultural share of income is 80%. While the actual increase in total income is 25%, the increase in income based on the under-reported figures is 40%. The fact that reported

income is roughly 50% of reported expenditure suggests that under-reporting is common and the comparative attention given by the questionnaires to agricultural income versus non-agricultural incomes, suggest that it is the latter which is under-reported.

- Households might be saving, rather than spending, their income gains. For example, households may increase purchases of durable goods, productive assets or animal stocks, which we do not include in the expenditure figures. Data on savings from bank accounts and net credit positions are collected by the questionnaire. However, either households save and borrow very little or they rely on other means of savings rather than cash, such as animals and durable goods. Any analysis of saving patterns will have to be based on the latter figures rather than cash holding or debit positions.
- The increase in incomes could also be the result of a positive covariate shock in the MV areas. The MV villages are clustered in a small geographic area and might have been positively affected by rainfall. However, note that this still does not explain the divergence between changes in incomes and changes in consumption, unless it is assumed that farmers do not increase consumption during good agricultural years, which is unlikely. It is worth noting the same positive effect is observed both in the second and in the third year of data collection (Table 29), which makes the covariate shock explanation less likely. Additionally, data on covariate shocks such as droughts, floods and excessive rains reported by farmers were not found to be very different in MV and CV areas throughout the years.

132. The PRA study found that most households in the MVs are using less of their food stocks on funerals due to a direct campaign by the project. In the Builsa district there is an area-wide ban imposed by the paramount chief on the sale and use of liquor at social gatherings. These two measures have left households with more food and capital, which is likely to benefit the richer ones. The RCA households indicated that they had no savings and many indicated greater debt than observed previously at the baseline. Nevertheless cash earning is essential to their livelihoods, and all RCA households had added some form of petty trading to achieve increased demands for cash. Even after extensive conversations, however, these households could not explain what incomes they were earning from these trades. Our conversations also revealed that people are not very good at monetising their consumption, which needs to be taken into account when interpreting the statistical results.

Table 29. Difference-in-difference impact on per capita income (IPW method)

	Cross-section	Fixed effects	Lagged model
Average DD effect	0.398*** (0.000)	0.464*** (0.000)	0.210** (0.013)
DD effect second year	0.391*** (0.000)	0.414*** (0.000)	0.209* (0.054)
DD effect third year	0.405*** (0.000)	0.423*** (0.000)	0.212** (0.042)
Sample size	5,942	5,941	3,941

P-values in parentheses based on cluster adjusted standard errors *** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

133. We further investigate the impact of the intervention on household income by disaggregating changes by income source. Changes in agricultural profits would suggest that the project is effective in increasing agricultural productivity (such as through training, better agricultural practices, tractor services, fertiliser, improved seeds, etc.). Changes in livestock income may be (in

part) due to asset transfers from MVP.⁴⁹ An increase in business income in the MVs would suggest an increase in general economic activity possibly stimulated by the injection of liquidity and facilitated by easier access to markets through improvement in infrastructure. A change in income from employment would suggest that the project is increasing incomes of those directly employed by the intervention or indirectly through the stimulus to economic activity. Finally, an increase in transfers would be hard to interpret from the project perspective as the project is not encouraging migration and does not provide direct monetary transfers.⁵⁰ To simplify the presentation of the results we only show the more conservative estimates obtained using the lagged dependent variable model (Table 30).

Table 30. Impact of the MVP on different income sources (lagged model)

	Average DD effect	DD effect second year	DD effect third year
Agricultural income	0.264** (0.041)	0.261** (0.038)	0.266 (0.101)
Livestock income	0.207** (0.015)	0.044 (0.687)	0.241** (0.006)
Business income	0.244** (0.016)	0.327** (0.046)	0.161 (0.163)
Employment income	-0.087* (0.078)	-0.046 (0.196)	-0.127 (0.141)
Transfers income	0.114*** (0.000)	0.066 (0.245)	0.163** (0.021)

P-values in parentheses based on cluster adjusted standard errors *** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

134. The MVP appears to change all income sources. The largest increases are observed in business income and agricultural income. Since the latter represents nearly 50% of overall income, it largely explains the overall income change observed from the survey data. There is also a noticeable increase in livestock income and an increase in income from transfers. Interestingly there is a small but statistically significant reduction in income from employment. These results would suggest that the project has so far been successful in increasing agricultural productivity and stimulating overall economic activity in the area. While the PRA did not assess incomes per se, its findings on overall well-being indicate that MVs are somewhat ahead of the controls at the midterm point, though not by a particularly large margin. Despite poor rains, the PRA finds that MVs have generally experienced relatively higher crop yields as a result of the diversified support from the MVP. The RCA found that this is mostly attributed to timely tractor services for land preparation which are provided at a standard rate and are offered in MV areas more than CV areas.

135. While migration has slowed overall in the MVs, the survey data shows that migrant remittances/incomes seem to have been a more important source of agricultural capital in the MVs than in the CVs. This may be because of the opportunities to utilise the agronomic know-how acquired through the MVP. The RCA, on the other hand, noted an increase in migration for work especially because of the late onset rains. Many told the team that this was their strategy to be able to pay for agricultural inputs, but others left the area to ease the financial burden at home.

136. Some microenterprise start-ups (e.g. grind-mills and cold drink sales) have also been reported in communities that have been connected to the national electricity grid since the baseline, though the numbers are small. The RCA found that people felt that improved road connectivity had had

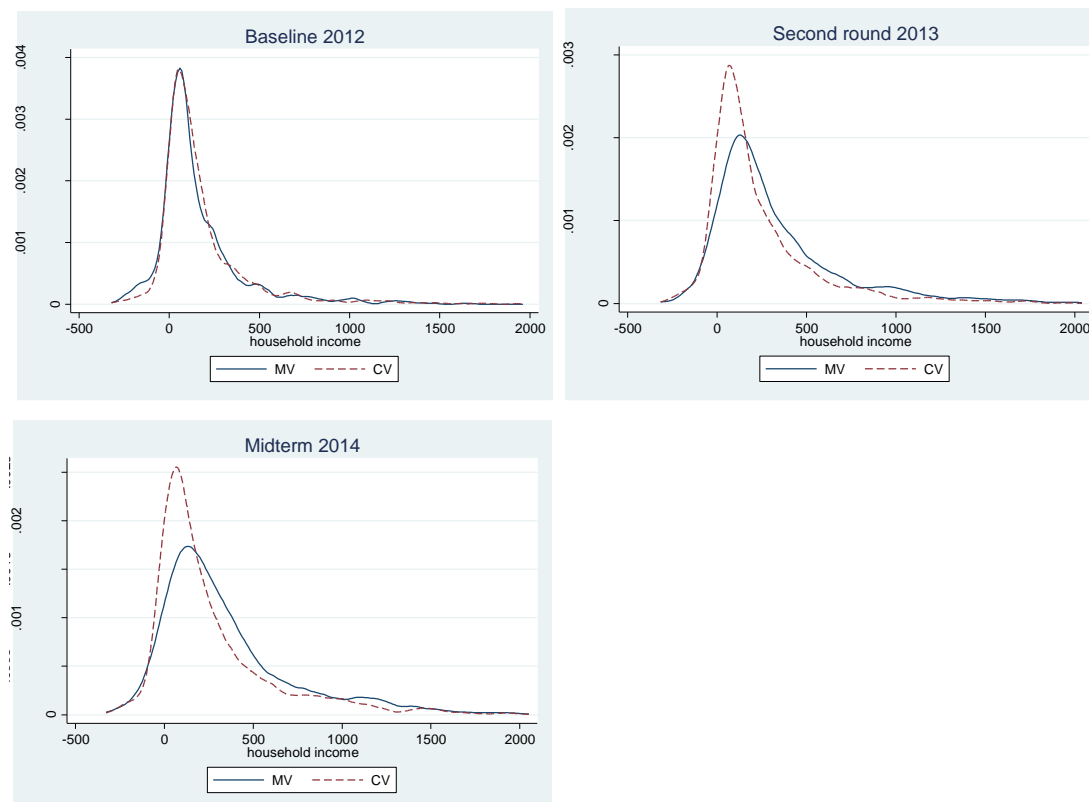
⁴⁹ According to the 2014 SADA-MVP Annual Report, each household was provided with small livestock (goats and chickens) plus training on how to generate income from their production.

⁵⁰ Although LEAP is operating in the area, and some in the RCA study claim to receive payments.

a positive effect on their willingness to start other small enterprises, for example, in transport services, food preparation and petty market trade to service more 'passing trade' and because it was now potentially profitable to do so. People talked about the importance of generating cash this way whereas they hardly ever talked about any increased agricultural income. The range of civil works activities (e.g. construction of culverts and other road improvements as well as health and school infrastructure), the recruitment of community health/education workers (into positions which did not exist prior to the MVP), the expansion of LEAP and new opportunities to provide portage and other low level services at some of the increasingly vibrant markets have all contributed incremental supplements to the incomes of some households. The RCA noted that jobs directly generated by the MVP such as in security, fire protection, waste management, and CHWs were often commandeered by a small connected group in the village rather than being more equitably shared. The significant reduction in income from employment picked up by the quantitative analysis does not necessarily mean that more people are engaged in agriculture. The RCA study found that more people were undertaking short migration trips for work than before and that the money earned from this was also spent away and contributed to reducing the burden on the home-based family. During these trips agricultural inputs are often purchased and the family at home cannot accurately put a monetary value on these. This may result in under-reported incomes from employment.

137. We then investigated whether the intervention had a different impact at different levels of the income distribution. The Figure 7 charts show the densities of household incomes in the project and comparison areas for the three survey rounds. The shapes of the distributions are remarkably similar in the project and control groups at the baseline. There is a clear shift to the right of the MV income distribution at the second round which becomes even more accentuated at the midterm survey. The impact of the intervention appears to be more pronounced in the middle and upper end of the income distribution.

Figure 7. Densities of household income in the project and control groups



Note: the horizontal axis reports per capita household income in Ghanaian Cedis.

138. These impressions are confirmed by quantile regression analysis which shows impacts that are higher at higher expenditure quantiles (Table 31). The results are in sharp contrast with those observed in relation to per capita expenditure where impact of the intervention appeared to be larger for the poorer sectors of the population. This discrepancy in the project effects on income and expenditure could occur because households are saving income gains, and the discrepancy in the distribution of effects on income and expenditure could be the result of the better-off households saving gains while the worse-off are spending. This is presently unclear and, as mentioned above, the RCA households are not saving. They shared a need, should they have surplus cash, to meet social and financial obligations including cash contributions to weddings and funerals and paying off accumulated debts, some of which may have existed for many years.

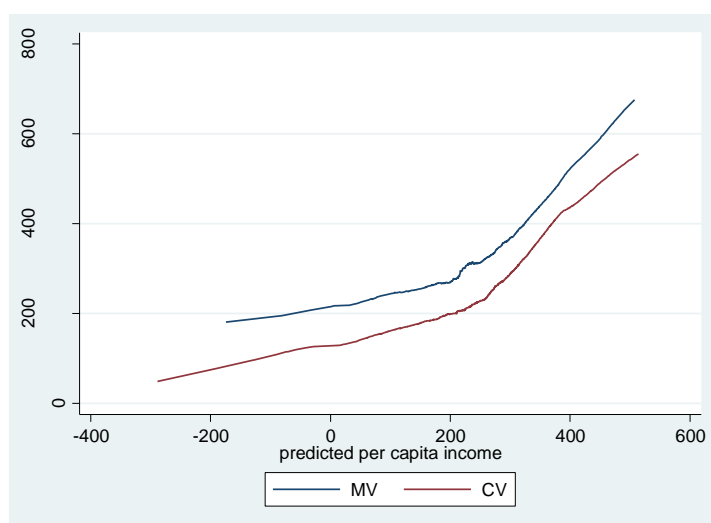
Table 31. Distributional impact on household income (average DD effects)

	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
Average DD effect	0.083** (0.010)	0.157*** (0.000)	0.224*** (0.000)	0.291** (0.002)	0.397** (0.015)
DD effect second year	0.087** (0.037)	0.125*** (0.000)	0.185** (0.001)	0.277** (0.035)	0.373* (0.051)
DD effect third year	0.068 (0.136)	0.176*** (0.000)	0.263*** (0.000)	0.548** (0.027)	0.430 (0.104)

DD coefficients of lagged models. P-values in parentheses are corrected by clustering at the village level using bootstrap sample replications. *** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

139. We repeat the distributional analysis conducted above using data on per capita income (Figure 8). Unlike the results in Table 31, Figure 8 shows that per capita incomes increase equally across the income distribution rather than benefiting the upper half. The most interesting conclusion from the comparative analysis of the distributional project effects on expenditure and income, however, are the implicit implications for household savings. The two figures are consistent with an increase in income that is spent by the poorest half of the population and saved/invested by the richest half. As could be expected in a deprived area, the relatively 'rich' save more, while the poor spend any income increase.

Figure 8. Per capita income by predicted per capita income



140. We also investigated the impact of the intervention on food security. The household surveys collect information on perceptions of food security via two questions. The first question elicits a

binary response: “In the past 12 months, were there months in which you did not have enough food to meet your family’s needs?” The second question obtains a continuous response: “How many days in the last 30 days did you not have enough food to meet your family’s needs?” Note that these questions were asked only at the baseline and midterm rounds and not during the ‘in between’ rounds so that DD effects are reported with respect to the change between the baseline and the midterm surveys. The effects are, as usual, calculated in three different ways and adjusted by baseline characteristics using the inverse probability method. For simplicity, we employ the same propensity score already employed when calculating the impact of the project on income and expenditure as the determinants of food security are similar to the determinants of income and expenditure.

Table 32. Project impact on food security perceptions (IPW method)

	Cross-section	Fixed effects	Lagged model
Not enough food in the last 12 months	-0.33*** (0.000)	-0.33*** (0.000)	-0.32*** (0.000)
Days without enough food in the last month	-1.9 (0.213)	-1.8 (0.211)	-2.5** (0.005)
Sample size	3,960	3,916	1,958

DD coefficients of lagged models. P-values in parentheses are corrected by clustering at the village level using bootstrap sample replications. *** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

141. The project appears to produce a large percentage reduction (30%) in the fraction of the population reporting not having enough food to eat over the last 12 months (Table 32). There is also a reduction in the number of reported days during which the household did not have enough to eat over the last 30 days.⁵¹ The impact for this second measure of food security is somewhat smaller and not robust across different specifications. Overall, these data show that perceptions of food security have improved in the MV areas. The findings of the PRA study in this regard are rather more mixed (see Table 33). For instance, food security has indeed risen overall in a small majority of MVs (four of the seven MVs sampled for the PRA), with participants reporting fewer months of food scarcity – in spite of recurrent disappointments about poor rains. This is supported somewhat by recurrent reports of a decline among MV households in the traditional practice of sending young girls off to serve in better-endowed urban households, often prompted by food insecurity. Overall too, women in the MVs have achieved more progress in enlarging their access to farmland, which helps with food security as women are traditionally responsible for providing the vegetables and seasonings required in the kitchen. Somewhat surprisingly, however, while the analysis shows MV households as being markedly more food secure than their counterparts in the nearby controls, the faraway control sites appear to match the MVs quite respectably on this indicator. Further analysis reveals that much of the improvement in food security in faraway control villages has been on the back of economic activities with a detrimental impact on the environment – such as indiscriminate harvesting of wood to produce charcoal, unbridled mining of sand and illegal mining. These activities, often involving children, then enable households to purchase the food they need. The RCA households were not eating any better than before and some were facing food insecurity in that their production had declined, they had more mouths to feed or they had had to sell stocks to service previous debts. Similar to the PRA study, the RCA faraway CVs were also found to have slightly better or similar food stocks compared with the MV. Unlike the PRA study, however, the RCA found that in MV people were migrating (particularly in areas with good road accessibility) for informal work in markets, the food industry, porter jobs as well as some illegal operations partly to reduce the burden of food provisioning at home. So, without these members of the family those remaining could eat better.

⁵¹ Although this is a very season-dependent question, as even a shift of a month may result in very different responses depending on how many mouths a household is feeding at any one time.

Table 33. Perceptions of food security

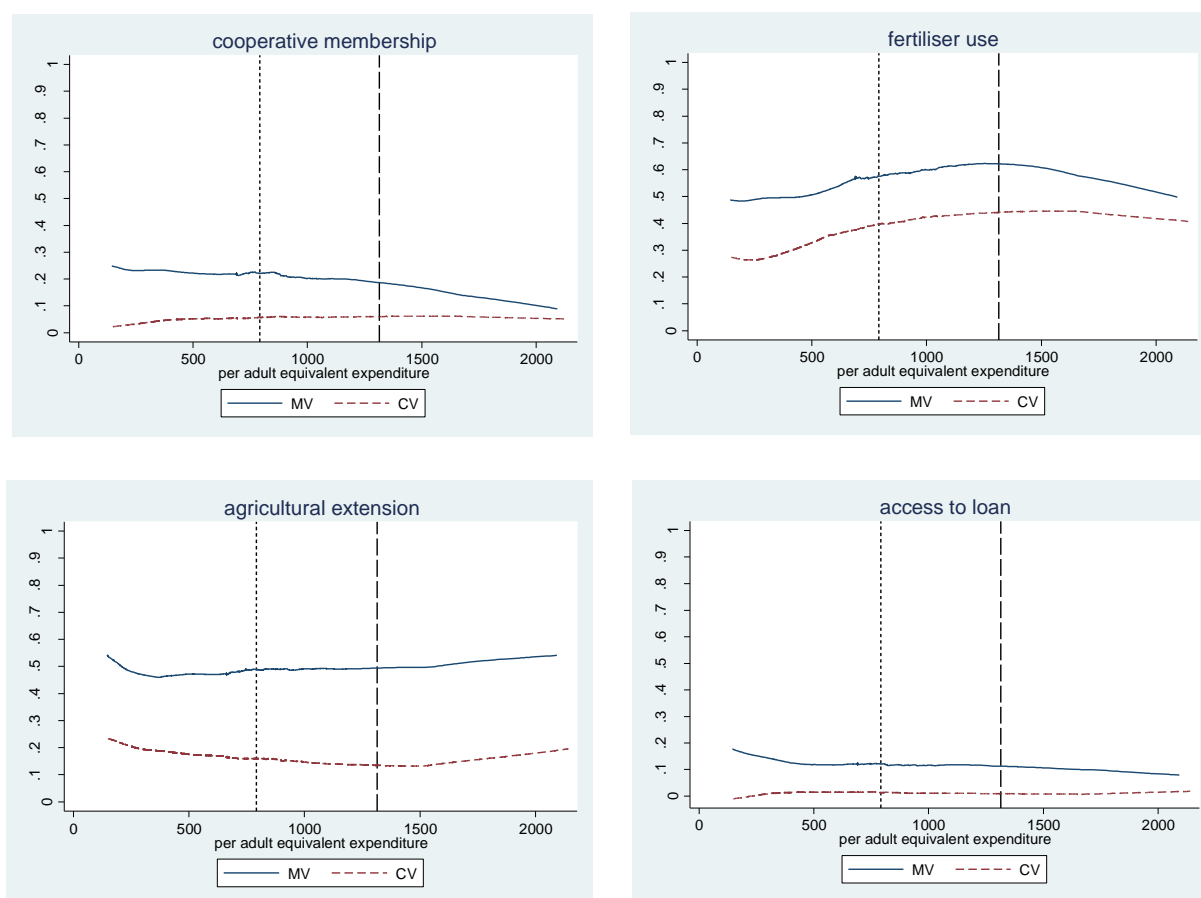
Comparator cluster no.	Characteristic	MV	CN	CF
1	Food security	Up	Up	Up
	Food expenditure on funerals	Down	Down	Up
2	Food security	Up	Down	Up
	Food expenditure on funerals	Down	Up	Down
3	Food security	Down	Down	Up
	Food expenditure on funerals	Up	Up	Up
4	Food security	Up	Down	Up
	Food expenditure on funerals	Down	No change: high	No change: high
5	Food security	Down	Down	No CFs
	Food expenditure on funerals	n/a	Down: by default	
6	Food security	Up	Down	Down
	Food expenditure on funerals	Down	Down	Down
7	Food security	Down	Down	Up
	Food expenditure on funerals	down	Down	Down

142. Food expenditure on funerals has also declined marginally more in the MVs than in the controls.

This was attributed largely to continuous education by the SADA-MVP team. The RCA team stayed in villages during a very intense funeral period with each of the team experiencing these directly. Conversations suggest that this was more active than previous years and that spending had increased. People shared how this is still an increasing burden regardless of whether they were in MV or CV areas.

143. In explaining how the MVP may have contributed to these changes, we can consider the involvement of households in project activities. This has increased in agriculture, with households in the MV areas reporting higher levels of fertiliser use, access to agricultural extension, access to loans and involvement in cooperatives than their counterparts in CV areas (Figure 9). While this is valid, the RCA study which specifically focuses on poorer households found that, with the exception of a very small number of elderly farmers,⁵² there were also instances of exclusion – with households not being invited to join groups. Either they self-excluded, were discriminated against because of concerns they could not meet the requirements of group membership or had left because of a previous bad experience. Apart from attending public showings of agricultural videos, these households had no extension support and felt excluded. The video films were much appreciated and attendance at these are large. The use of local language was much appreciated. However, the RCA team met very few in the community beyond their households who had actively experienced access to other forms of agricultural extension. The PRA team too was informed that many poor farmers were no longer eligible to receive credits, having failed to pay back the first cycle of credits owing to the failed rains. In some CVs where the Presbyterian Agricultural Station has a presence, the PRA team was informed that households are supported through their women members with farming capital in the form of high-yielding maize and soya seeds, one free acre of ploughing, fertiliser and goat breed stocks, making it easier for the poor to benefit.

⁵² Presumably included in the MV 'special groups'.

Figure 9. Involvement in project activities by household expenditure

Note: These percentages are fairly low, and access to loans is particularly so.

144. The RCA also found that some increased economic activity in communities directly relates to the MV programme; for example, there are workers installing the telephone masts and electricity connections and working on roads and there is an increase in house rental. These people have disposable incomes which are often spent in the community on food, mobile phone credit and drink – and have given a boost to some of the small enterprises. However, when the workers leave some of these enterprises close or consider closing.

5.3. Impact of the MVP on child health

145. In this section we examine the impact of MVP on three major indicators of child health: mortality, anthropometry and anaemia. First, we calculate the DDs in mortality rates between the baseline and the midterm. In order to do so we match observations in the project and comparison groups using individual and village level characteristics to remove the large difference in mortality rates between MV and CV areas that were observed at the baseline (the procedure is reported in detail in the Annex A).

146. Mortality rates are decreasing in MV areas at a higher speed than in CV areas. The PRA study appears to support an accelerated rate of change especially in increasing easier access to antenatal and postnatal education and immunisations in the MVs. In several MVs, vital healthcare staff have moved into new purpose-built or renovated accommodation. The RCA study shows that people are particularly appreciative of the midwives who have been recalled from retirement and are pleased to stay in the village (although villagers are concerned about the sustainability of

such arrangements). Both the PRA and RCA studies confirmed that there is enhanced access to emergency healthcare, particularly at night, and more provision of emergency transport. Both the RCA and PRA studies show that there is a clear shift in preference for institutional births away from using traditional birth attendants. An initiative to get traditional birth attendants to refer cases to the formal healthcare facilities (through a combination of incentives and pressure) has been a further contributory factor. The CHW initiative has likewise facilitated a speedier management of malaria cases, reportedly contributing to a reduction in morbidity and mortality. The malaria spray programmes (not MVP), supported by a number of different actors including the private sector, have had a noticeable impact on the number of mosquitos. Additionally, the supply of motorbikes and bicycles to MV-associated health workers has resulted in increased home visits and facilitated the administration of immunisations.

147. This statistical analysis, however, suffers from two main limitations. First, the size of the sample is not sufficiently large to capture small effects as those the programme might have produced. Second, mortality rates reported in Table 34 are calculated over five years before the survey. Rates calculated in this way are not appropriate for a DD analysis because the rates calculated at midterm (two years after project start) also include birth and death events that occurred up to three years before the baseline. Only a dramatic change in mortality rates after the baseline would be visible using these data. The data suggest that a drop in mortality rates occurred but better estimates will be calculated in 2016 when the fifth round of data collection becomes available and mortality rates of four or five-year intervals can be calculated and compared over time between MV and CV areas. We recognise that the analysis of mortality rates in the existing data are not very informative because of the limitations outlined above. However, we decided to conduct this analysis nevertheless and to report the results for two reasons. First, we wanted to detect the presence of any trend. The data seem to suggest an improvement in mortality indicators in MV areas, which will be appropriately tested with the endline data. Second, we wanted to ascertain the quality of the mortality data collected, and our ability to correct for initial differences in mortality rates between the two areas.

Table 34. DD analysis (five-year interval)

Mortality rate	Baseline difference	Midterm difference	DD
Neonatal	0.36	0.89	0.53
P-value	0.980	0.877	0.970
Post neonatal	0.40	-6.00	-6.40
P-value	0.962	0.635	0.600
Infant	0.76	-5.11	-5.87
P-value	0.967	0.696	0.766
Child	4.14	-6.47	-10.62
P-value	0.626	0.457	0.354
Under-5	4.68	-11.26	-15.95
P-value	0.829	0.492	0.512

Mortality rates are calculated using the synthetic cohort probability method used by the DHS. Standard errors are calculated using bootstrap replications. Observations are weighted using IPW derived from the propensity score based on the model of Table 29. *** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

148. The RCA study noted that there seems to be specific effort by non-MV actors (NGOs, INGOs, faith-based organisations, government, private philanthropists and politicians) to support the improvement of health outcomes in the CVs. So, for example, antenatal clinics were as active in MV and CV areas, new health facilities are being constructed in CV areas and uptake of exclusive breastfeeding for the first six months after birth has increased in both MV and CV areas.

5.4 Impact of the MVP on anthropometry

149. The project had a positive impact on physical growth of children. Anthropometric data were only collected at the baseline and at the midterm so that the DD analysis can only be conducted between two points in time. We look at standardised Z-scores among children under 5: height-for-age, weight-for-age and weight-for-height (Table 35). Scores were calculated using the most recent World Health Organization methodology.⁵³ The DD coefficients are adjusted for differences in baseline characteristics using inverse probability weights. Note that we corrected for baseline characteristics of households rather than baseline characteristics of children because many of the children measured at the midterm were not measured at baseline.

Table 35. Impact of the MVP on nutritional status of children under 5 (IPW method)

	Difference-in-difference	P-value	Sample size
Height-for-age	0.48***	0.001	3,392
Weight-for-age	0.24**	0.036	3,398
Weight-for-height	-0.11	0.373	3,384

*** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

150. We also analysed the MVP's impact on children's anthropometrics using the sample of children measured at both the baseline and midterm. Restricting the analysis to the sample of children observed at baseline and midterm exploits the estimation advantages of using panel data, but it comes at a cost. The first cost is the reduced sample size, which in turn reduces the statistical power and ability to identify impacts when they are present. The second cost is modifying the original sample of analysis. When using panel data, the estimates no longer refer to the population of children under-5 over the two periods, but to a subset of children that due to their age and other characteristics happened to be interviewed twice. As shown in Table 36, using panel observations reduces the sample by approximately 50%. The main reason for this is because anthropometric measurements are taken from children under-5. Children that were older than three years of age at the baseline could no longer be measured at midterm because they fall outside the specified age range. Similarly, unborn children at baseline are now in the pool of children under-5 measured at midterm. Another reason for the reduction is because there are errors and difficulties in identifying household members (discussed in Section 2.3), therefore it is not possible to enumerate exactly the same individuals each visit. Finally, even if children are correctly identified in both surveys and fall within the correct age range, they might refuse to be measured or be sick on the day they are visited, so only one observation is made over the two periods. Since the observations in the panel sample belong to a specific age cohort and might share characteristics that make them more likely to be in the panel dataset, they are not fully representative of the population, and in general we prefer to work with the cross-sectional samples observed at baseline and midterm rather than with panel data. However, to complete the analysis, an impact analysis of the intervention is performed on the sample of panel children using fixed effects and ANCOVA and adjusting for inverse probability weights as in the previous sections. The results in Table 36 show very similar results to those obtained using the cross-sectional samples. The project appears to have a strong impact on height-for-age and weight-for-age.

⁵³ We calculated Z-scores using WHO's STATA igrowup package of the WHO.

Table 36. Impact of the MVP on nutritional status of children under-5 (panel children)

	Height-for-age	Weight-for-age	Weight-for-height
Difference-in-difference (fixed effects)	0.49**	0.39**	0.06
P-value	0.002	0.001	0.576
Sample size	1,484	1,486	1,476
Difference-in-difference (ANCOVA)	0.38**	0.23**	-0.05
P-value	0.001	0.003	0.622
Sample size	719	723	722

151. The observed impacts of the project on weight-for-age and particularly on height-for-age are very large compared to a typical intervention focused on agriculture (Masset et al. 2012), and it is large even for specific nutrition interventions (IEG 2010) – although it is acknowledged that the MVP is a multi-sectoral programme and direct comparisons are not necessarily appropriate. The focus group interviews under the PRA study confirm that it is less common for children in MVs to look for piecework or to forage for food as a way of bridging their food deficit situations. There is a possibility that these effects are magnified by seasonality because data were collected at different times in MV and CV areas during the baseline (see discussion in paragraph 113). However, our analysis of secondary data and review of the literature (see Appendix H of baseline report) suggest that seasonal patterns of anthropometric indicators in Ghana are minor, if any. In addition, the largest impact of the intervention is found on height-for-age. Height is less sensitive to shocks, including seasonality – height can only increase. Height is a long-term indicator of nutritional status which is less likely to be affected by shocks and seasonal fluctuations. The fact that the intervention was particularly effective on increasing average heights, suggests that little of the observed impact is seasonal.

152. Heights and weights of children have improved in MV areas and since they have got slightly worse in comparison areas over the same period, the DD effect is magnified. Nutrition is determined by multiple factors including the size and the composition of the diet, morbidity and general health environment and parental care. All these factors may have contributed to a varying extent to achieve this result. A more detailed analysis, unpacking the determinants of these results, is in order but beyond the scope of this report. The RCA study, however, suggests that there may be nutrition and sanitation practices related to religious beliefs, which may differentially affect child morbidity and growth. For example, we found infants with conspicuous signs of *kwashiorkor*⁵⁴ in our two Mamprusi villages and not at all in the other villages. Two other practices also stand out as different in these villages compared to the others: the poor uptake of exclusive breastfeeding practices and fewer or no pigs⁵⁵ (which tend to help clean up the environment by eating faeces).

5.5 Impact of the MVP on anaemia of under-5 children

153. The baseline survey found very high prevalence rates of anaemia in children under 5, comparable with those prevailing in Ghana and other West African countries.⁵⁶ Following DHS standards, we calculate mild anaemia prevalence as the ratio of children with haemoglobin below 11 g/dL. Moderate anaemia is haemoglobin below 10 g/dL and severe anaemia is haemoglobin below 7 g/dL. We find no improvement on average haemoglobin concentration or in the distribution of anaemia prevalence (Table 37). Haemoglobin concentration is slightly higher in MV areas at

⁵⁴ The name is derived from the Ga language of coastal Ghana. It is a form of severe protein-energy malnutrition.

⁵⁵ All three villages have a Muslim majority and one has banned pigs altogether this year.

⁵⁶ Ghana Statistical Service, 2011. Ghana Multiple Indicator Cluster Survey with an Enhanced Malaria Module and Biomarker, 2011, Final Report. Accra, Ghana.

midterm and the prevalence of anaemia is slightly lower in MV areas, but these differences are not statistically significant.

154. It should be noted, however, that these results could have been biased by seasonal differences. Baseline blood tests were conducted during the dry season in the MV areas (May–June), when haemoglobin concentration is higher, and in the wet season in the CV areas when haemoglobin concentration is lower (August–September). The midterm data were collected in the period July–August in both MV and CV areas. It is possible that the improvement over time observed in the CV areas is partly the result of seasonal bias and that the improvement observed in the MV areas would have been much larger had the blood test been conducted in the wet season rather than in the dry season.

Table 37. Impact of the MVP on haemoglobin concentration and prevalence of anaemia

Indicator	Baseline			Midterm			DD	DD*
	MV	CV	Diff	MV	CV	Diff		
Haemoglobin	9.98	9.46	0.51**	10.2	9.98	0.20*	-0.31	-0.32
P-value			(0.001)			(0.054)	(0.149)	(0.171)
Mild anaemia	74.0	84.3	-10.3**	72.1	74.3	-2.1	8.1	8.2
P-value			(0.003)			(0.498)	(0.128)	(0.182)
Moderate anaemia	45.7	61.9	-16.2***	43.2	47.0	-3.7	12.4*	12.1
P-value			(0.000)			(0.324)	(0.064)	(0.102)
Severe anaemia	3.9	5.2	-1.34	1.0	2.0	-1.0	0.3	1.1
P-value			(0.464)			(0.200)	(0.859)	(0.594)

Unadjusted difference-in-differences (DD) and adjusted (DD*) are reported in the last two columns.

*** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

155. The RCA finds very little change in diets over the project period, with some families actually eating less well because of poor harvests. Those who managed to grow bambara beans, which are a good source of iron, eat these at least once per day mixed with tuo zaafi.⁵⁷ These are households which have suitable soil and sufficient water for bean production and they indicate that they eat beans more regularly than before. School feeding programmes seem to be continuing in MV schools more than CV schools and while having no effect on attendance might have an effect on nutrition outcomes. Another seasonal factor affecting haemoglobin counts may be the diet at these times. During intense agricultural periods people say they have less time for cooking, therefore foods that require longer preparation, such as dried beans, are less likely to be eaten.

5.6 Impact of the MVP on education

156. The project appears to have increased school attendance (Table 38). We first conducted a DD analysis of the MVP's impact on school attendance in primary, junior secondary and senior secondary schools.⁵⁸ We assessed the impact on school attendance using both cross-sectional and panel samples. The cross-sectional analysis looks at changes in school attendance between MV and CV areas at population levels. In the cross-sectional analysis, the children observed are different but belong to the same age group, as new children enter younger age cohorts at the

⁵⁷ Tuo Zaafi is a very thick cooked porridge consisting of maize and water.

⁵⁸ Note that the MVP intervention focuses on primary schools and does not provide infrastructure, teachers or other form of support to Junior Secondary Schools (JSS) and higher education. However, we are looking at the impact of the intervention on levels higher than primary schools for several reasons: (1) pupils attending JSS at the endline were in primary at the baseline. It is likely that their attendance of JSS and their education results are affected by the type of education received in primary; (2) as the project mobilises parents and girls through clubs and other events, it is likely that at these occasions the value of education is stressed for JSS as well, thus encouraging attendance; (3) attendance of JSS and learning outcomes may improve because pupils are beneficiaries of parallel interventions promoting, for example, better health and better nutrition. The latter have often been shown to increase school attendance and learning outcomes.

midterm while older children exit to higher age cohorts. The largest project impact occurred in primary school (8% increase in attendance). A smaller impact (5% increase) was observed in junior secondary school and no impact was found in senior secondary. The impacts appeared to have been larger in the first year of the project in primary school and in the second year of the intervention in junior secondary.

Table 38. Impact of the MVP on school attendance

	Primary	Junior secondary	Senior secondary
Average DD effect	0.08** (0.029)	0.05* (0.099)	0.00 (0.799)
DD effect second year	0.11** (0.006)	0.03 (0.371)	0.001 (0.735)
DD effect third year	0.047 (0.205)	0.08* (0.052)	0.00 (0.909)
Sample size	8,022	3,161	4,019

*** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

157. When the same analysis is applied to panel children it produces similar results (Table 39). Restricting the analysis to children that were observed at the baseline, second round and midterm reduces the sample size by about 50% because not all individuals are always enumerated each survey (see Section 2.3) and older children interviewed at baseline exit the reference age groups in later surveys. The latter reason particularly restricts the sample of children attending JHS because the reference age groups for this schooling level is very small (only three years of completed age). The impacts assessed by the cross-sectional analysis are largely confirmed by the analysis of panel data. There is a modest impact on primary school attendance, a minor impact on JHS attendance (though the small sample size does not allow the coefficient to reach statistical significance) and no effect at all on attendance of senior secondary school.

Table 39. Impact of the MVP on school attendance (panel children)

	Primary	Junior secondary	Senior secondary
Average DD effect (fixed effects)	0.08**	0.07	0.01
P-value	0.006	0.143	0.628
DD effect second year (fixed effects)	0.10***	0.06	0.00
P-value	0.000	0.276	0.962
DD effect third year (fixed effects)	0.06**	0.08	0.02
P-value	0.045	0.149	0.429
Sample size	4,014	549	1,248
Average DD effect (ANCOVA)	0.02	0.04	0.01
P-value	0.654	0.387	0.709
DD effect second year (ANCOVA)	0.04	0.03	0.00
P-value	0.413	0.510	0.981
DD effect third year (ANCOVA)	0.00	0.05	0.02
P-value	0.949	0.451	0.638
Sample size	2,676	336	832

*** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

158. The PRA team found that, overall, children in the MVs benefit more from basic schooling supplies and access to recreation facilities (although nearly all of the new recreation equipment observed during the RCA had already broken down). In several instances, the PRA and RCA teams found

that the non-MVP interventions (by Camfed, G-PASS and GPE) still include some MVs in their catchments, resulting in children from the MVs having superior access to such basic resources. More teachers have been recruited in the MVs and, relative to CVs, they are more likely to be in their classrooms and teaching, influenced in part by more regular supervision by the MVP education team. However, the RCA notes that new education facilities are also being constructed in the CVs and efforts are similarly being made to recruit more staff. While efforts to supplement teacher numbers in CVs was echoed in the PRA, the practice was not as consistent; neither were the numbers of teachers being recruited as high as in the MVs.

159. The PRA study reports a greater aversion for caning and intimidation in MVs, along with higher levels of disapproval for child labour and practices that tend to truncate a child's education such as early pregnancy or working as domestic helpers for better-off urban households. The RCA study could not, however, conclude on whether there was actually less corporal punishment in MV schools. Young teachers we spoke with were certainly less likely to use corporal punishment than their older colleagues, but this was a result of their more recent teacher training rather than emerging wider social norms. It may also be because there is more peer pressure for this in schools where there are larger numbers of newly recruited teachers. The team did not sense an increased demand for this in the community; quite the reverse was observed with very high levels of physical abuse between peers, parents to their children, children to their parents and some demands for teachers to be stricter. From the RCA direct observations in schools, the school environment has changed very little, except for new desks and chairs and occasional teaching resources such as posters. Furthermore, the PRA study came across a community school in the CV with very few resources but a vibrant learning environment.
160. The PRA study suggests that it is more common for children in CVs to look for piecework – sometimes in risky areas such as mining and stone quarrying – or to offer portage services in the markets and transport terminals. Children's involvement in such activities seemed markedly higher in faraway control sites. The RCA found that people were also reluctant to explain what sort of activities some of the younger members of their family were engaged in when they migrated for work, sometimes because of the risks associated with the legality of their work, but sometimes because they did not know. These grey areas mean that the survey may well have under-reported such issues.
161. The RCA study, however, finds that the primary pull factor for continuing school lies more with potential job prospects and widening personal networks for accessing jobs (e.g. through mobile phones) rather than the inputs provided to the school or to individuals. It further found that, especially among young teens but also their parents, a general disillusionment with farming and the increasing unpredictability of climate has encouraged young people to seek work away from their communities regardless of whether they are in MV or CV areas. Furthermore, families feel that the motivation and sacrifices made for their children's education is not intended to lead them into farming. Those who are regarded as 'school types' are encouraged by their families, job networks and older role models to continue at school and tell us they will do this irrespective of the facilities provided to them or the school. These often aspire to civil servant jobs such as teaching and nursing. Others who are perceived as 'non-school types' leave to earn their own money for their own needs as soon as they can. Boys who are struggling at school tell us that earning an independent income and being able to buy fashion accessories, toiletries, phone credit, recreation and so on is more attractive than staying in school. Seasonal work is also seen as an important means to raise the cash, which is increasingly required for agricultural inputs.
162. Nevertheless, attendance gains have not resulted in an overall improvement in learning and cognitive abilities measured by cognitive, maths and English test scores. We first look at the

impacts at population level using cross-sectional samples and then restrict the analysis to panel children who were tested at both baseline and midterm. Cognitive skills did not improve over the two-year period of intervention (Table 40). While the test scores on Raven's matrices⁵⁹ have improved in the MVs, the forward and backward digit span test⁶⁰ shows a regression in the MVs compared with the CVs. No impact is visible on simple English and maths tests, though there is a sizable improvement in the advanced English and maths tests that are administered to children attending junior secondary school. Relatively poor test results cannot be explained by the fact that the MVP brings school children in from disadvantaged backgrounds that on average perform worse than children attending school in the control group. Cognitive tests, as well as simple maths and English tests, were conducted at the household level regardless of schooling level thus this type of bias is unlikely.

Table 40. Impact of the MVP on test scores

	DD	P-value	Sample size
Raven's matrices	0.18	0.164	6,565
Forward digit span	-0.19*	0.076	6,471
Backward digit span	-0.22*	0.097	6,467
Easy English	-0.03	0.771	3,204
Easy maths	0.03	0.729	3,571
Advanced English	0.44**	0.004	0.929
Advanced maths	0.27	0.105	0.920

*** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

163. We repeated the same analysis for panel children who were tested at the baseline and midline. The sample of panel children is about 50% of the cross-sectional sample. The reduced sample size is due to difficulties with interviewing the same individuals across surveys, older baseline cohorts exiting the age group in which the tests are administered and refusal or sickness preventing a child from being tested in both surveys. Restricting the sample to panel observations changes the age composition under study and potentially further restricts the sample to a particular set of children who, for example, do not refuse the test or are not sick. For this reason, we prefer the results of the population-level estimates obtained from cross-sectional samples, but report the results of the impacts on test scores for the available panel of children tested both at baseline and midterm for completeness. Impacts are tested using a fixed effect model and a lagged variable model (ANCOVA). The observations are weighted by inverse probability weights in order to account for baseline differences in characteristics. The results are not substantively different from those obtained using the cross-sectional samples. A strange pattern regarding cognitive tests is confirmed, whereby the MVP positively affects Raven's test scores but negatively affects digit span tests scores. Effects on reading and maths tests are small and not statistically significant. Positive effects on advanced reading and maths tests that were found statistically significant using the cross-sectional sample are no longer statistically significant in the panel sample. Though it should be observed that the panel sample is nearly half the size of the cross-sectional sample and has lower statistical power to detect effects.

⁵⁹ Raven's coloured matrices are testing the ability to recognise patterns independently of mathematical and language knowledge and are believed to measure one dimension of intelligence.

⁶⁰ In the digit span test children have to repeat, in order of increasing difficulty, random series of numbers read out by the enumerator. In the backward digit span test the respondent is asked to repeat the series in reverse order. These tests are measuring short-term memory, which has been associated with attention, iron deficiency and nutritional status.

Table 41. Impact of the MVP on test scores (panel children)

	Fixed effects			ANCOVA		
	DD	P-value	Sample size	DD	P-value	Sample size
Raven's matrices	0.25***	0.000	3,442	0.08**	0.015	1,721
Forward digit span	-0.18**	0.001	3,348	-0.15*	0.098	1,673
Backward digit span	-0.31***	0.000	3,344	-0.21	0.136	1,672
Easy English	0.01	0.910	1,144	0.07	0.703	572
Easy maths	-0.03	0.737	1,542	0.11	0.151	771
Advanced English	0.11	0.556	244	0.15	0.592	122
Advanced maths	0.24	0.265	238	0.21	0.413	119

*** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

164. Just as the inputs being provided by Presbyterian Agricultural Station (PAS) are influencing a closer outcome gap in agriculture (Para 120), the distribution of educational inputs in non-MV communities by Camfed, G-PASS and GPE is narrowing the input gap. This, in turn, appears to be contributing to closing the schooling outcomes gap between MVs and CVs. That learning outcomes remain unimpressive, despite these inputs, may be partly attributable to the persisting low levels of teacher competence. Indeed, learning achievements cannot be expected to rise considerably if teachers (both regular teachers and volunteer CEWs) lack the skill needed to motivate children to learn and impart knowledge to them.

165. Interestingly, the RCA found that children in MV areas were much more confident about sharing their learning from school than they had been at the baseline. They enthusiastically engaged researchers in language and number games and songs in a way which we felt was not simply due to familiarity with the researchers.

5.7 Impact heterogeneity

166. In this section we analyse the impact of the intervention by district and sex. The population reached by the project is equally distributed between the original districts of Builsa and West Mamprusi.⁶¹ After the creation of new districts in 2012, the MVP villages are distributed in the following way: 23 in Builsa South, seven in West Mamprusi and five in Mamprugu Moagduri. The split of the Builsa district into Builsa North and Builsa South has not particularly affected our study design as all the project villages are located in Builsa South. The split of the West Mamprusi district on the other hand has divided project communities almost equally between the two newly created districts, which are administered in different ways. In our analysis, however, we employ the original subdivision pre-2012 between Builsa and West Mamprusi districts. This subdivision has only a partial validity in terms of administrative and political differences between the two areas but it does characterise areas that are quite homogeneous in terms of language spoken, ethnic groups, socio-economic characteristics and social and political organisation.

⁶¹ Note: This section refers to two districts as it discusses the original design and analysis, which at the time was based on two districts. Since this has now changed with the division of districts, the rest of the report refers to three districts in line with the present reality.

Table 42. Impact of the MVP by district

	Builsa	West Mamprusi	F-test
Per-adult equivalent expenditure	0.097 (0.202)	-0.020 (0.799)	1.65 (0.202)
Per capita income	0.518*** (0.000)	0.313*** (0.000)	2.91* (0.091)
Food security (1)	-0.307*** (0.000)	-0.346*** (0.000)	0.22 (0.642)
Food security (2)	-2.411 (0.197)	-1.225 (0.369)	0.62 (0.431)
Height-for-age	0.673*** (0.000)	0.374*** (0.000)	2.27 (0.135)
Weight-for-age	0.292** (0.025)	0.218 (0.133)	0.22 (0.641)
Weight-for-height	-0.135 (0.150)	-0.088 (0.608)	0.08 (0.778)
Haemoglobin concentration	-0.168 (0.443)	-0.436 (0.199)	0.59 (0.443)
Mild anaemia	0.052 (0.439)	0.104 (0.218)	0.31 (0.579)
Moderate anaemia	0.087 (0.286)	0.147 (0.143)	0.31 (0.579)
Severe anaemia	-0.001 (0.951)	0.021 (0.141)	0.78 (0.378)
Attendance primary school	0.030 (0.655)	0.106** (0.003)	1.12 (0.293)
Attendance junior secondary school	0.013 (0.746)	0.083* (0.052)	1.77 (0.187)
Attendance senior secondary school	0.015 (0.417)	-0.003 (0.881)	0.67 (0.415)
Raven's test	0.224* (0.052)	0.188 (0.364)	0.03 (0.857)
Digit forward test	-0.135 (0.322)	-0.169 (0.189)	0.04 (0.838)
Digit backward test	-0.361** (0.005)	-0.185 (0.311)	0.86 (0.357)
Easy maths test	0.089 (0.453)	0.034 (0.745)	0.22 (0.642)
Easy English	-0.019 (0.871)	0.174 (0.289)	1.19 (0.278)
Advanced maths	-0.030 (0.881)	0.718*** (0.000)	13.06** (0.001)
Advanced English	0.355* (0.094)	0.769*** (0.000)	3.23* (0.076)

The table reports DD coefficients of regression analysis based on cross-sectional models adjusted by IPW. *** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

167. In order to assess whether the project had a different impact in the two districts we considered the Builsa and West Mamprusi MVs as two different interventions. As usual we calculated DD effects using regression analysis after weighting observations for their selection probability. Table 39 shows the DD effects estimated for the two districts in relation to the outcomes analysed in the previous sections, together with F statistics that test the difference between the coefficients. The project appears to be more effective in Builsa than in West Mamprusi for most economic and health indicators, although a statistical test of the difference between the coefficients is significant only in the case of per capita income. It is interesting to note that the project appears to be more effective in West Mamprusi than in Builsa for education indicators. The difference in favour of West Mamprusi is apparent for tests administered to junior secondary school students, it is much less obvious for tests administered to primary school students and nearly non-existent in the case of cognitive tests. After applying the same False Discovery Ratio adjustment already applied to the results of Table 23, the only difference in advanced maths tests between the two districts remained statistically significant, while the other statistically significant effects found are more likely to be attributable to chance. Overall the data in Table 42 suggests that the project was moderately more effective in Builsa in promoting child health and economic activities, while it was more successful in promoting higher education in West Mamprusi. Note: it should be remembered that more conservative statistical testing rules out differences between districts with the exception of advanced maths test results.
168. We conducted the same type of analysis disaggregating the impact of the intervention by sex for all individual level outcomes considered in the previous sections (Table 40). We found no differences in project impact on nutritional status between boys and girls. Interestingly, the programme appears to be more effective in improving anaemia among girls than boys – although it is unclear why this is so. Also note that this effect fails the False Discovery Ratio test and it is more likely to be attributable to chance. On the other hand, the programme appears to be more effective in improving primary school attendance of boys. Overall the data suggests that project impact is quite evenly distributed between boys and girls and that differences emerge in those cases where boys and girls have different initial status. For example, more girls are normally attending school than boys, which might explain the project impact on school attendance of boys. However, the PRA team was frequently informed in the MVs that *older* girls in particular are more confident about attending school *regularly* following the construction of sanitary facilities and provision of sanitary protection in several MV schools. While some of this is also happening in CV areas, it is not as routine as in the MVs. The RCA study found that older girls are increasingly inspired by role models who are in paid employment and feel it is possible to emulate them. In the same study, boys told the team that there are well-paying options that do not require school certificates.

Table 43. Impact of the MVP by gender (boys and girls)

	Boys	Girls	F-test
Height-for-age	0.483** (0.003)	0.510** (0.001)	0.04 (0.849)
Weight-for-age	0.264* (0.053)	0.233* (0.060)	0.07 (0.794)
Weight-for-height	-0.084 (0.539)	-0.128 (0.297)	0.21 (0.648)
Haemoglobin concentration	-0.448* (0.086)	-0.158 (0.491)	3.97* (0.049)
Mild anaemia	0.117 (0.100)	0.038 (0.559)	1.82 (0.180)
Moderate anaemia	0.155* (0.054)	0.079 (0.307)	1.95 (0.166)
Severe anaemia	0.017 (0.515)	0.002 (0.936)	0.52 (0.471)
Attendance primary school	0.097** (0.020)	0.038 (0.313)	2.55 (0.113)
Attendance junior secondary school	0.057* (0.080)	0.050 (0.277)	0.03 (0.864)
Attendance senior secondary school	0.010 (0.562)	-0.07 (0.813)	0.37 (0.547)
Raven's test	0.257 (0.111)	0.149 (0.309)	0.98 (0.324)
Digit forward test	-0.212* (0.053)	-0.098 (0.381)	2.63 (0.108)
Digit backward test	-0.287* (0.033)	-0.221 (0.122)	0.69 (0.409)
Easy maths test	-0.015 (0.886)	0.148 (0.183)	1.76 (0.187)
Easy English	0.002 (0.991)	0.211** (0.047)	1.70 (0.195)
Advanced maths	0.265 (0.127)	0.341* (0.084)	0.22 (0.642)
Advanced English	0.586** (0.005)	0.518** (0.015)	0.06 (0.803)

The table reports DD coefficients of regression analysis based on cross-sectional models adjusted by IPW. *** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

169. The RCA continues to show, as in the baseline, that more girls than boys attend school. It also points out that there are many overlapping incentives for girls, and boys tell us that they can feel 'second class' and prefer to leave school to become 'adults' and earn for themselves.

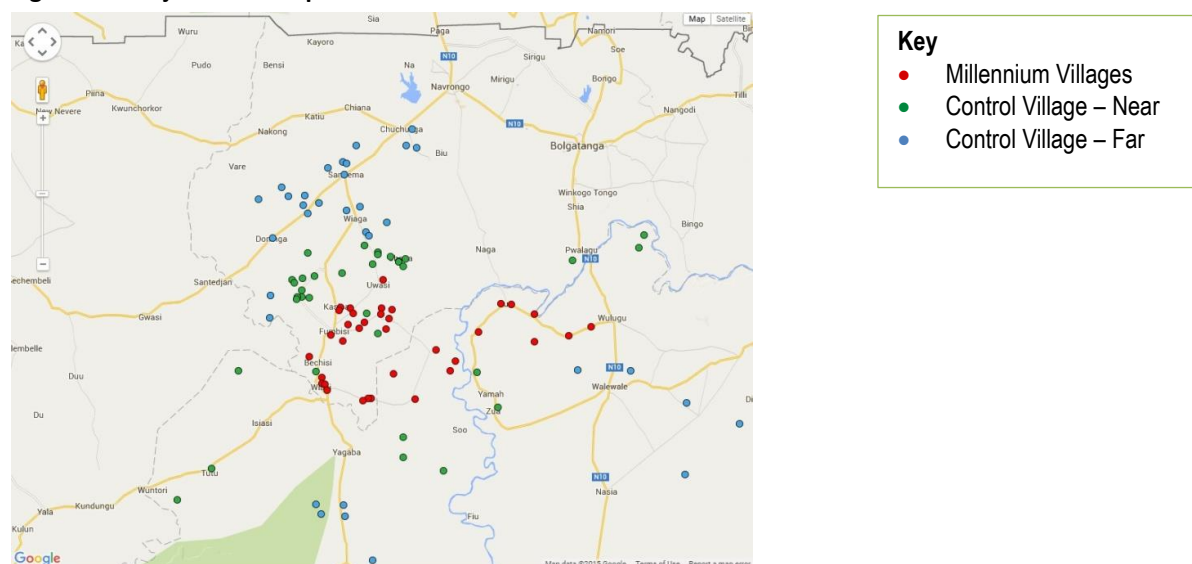
Chapter 6. Spill-over effects

170. In this section we investigate the presence of spill-over effects, that is, whether the MVP had an impact beyond the targeted geographical area of intervention. The detection of spill-over effects would imply that the impacts estimated in the previous sections are underestimates of true project effects. As discussed in the Analysis Plan, this analysis is exploratory as we have not yet formulated a conceptual framework describing the mechanisms through which spill-over effects should operate.⁶² There have been reports that families residing in non-project areas are attracted to access health clinics run by the project, and several channels could be hypothesised through which the programme could have an economic impact outside the areas of intervention on, for example, markets and prices.

6.1 Geographic spill-over effects

171. We used the distance of control group villages from MVs as a proximate determinant of spill-over effects but we do not have any guiding principle to establish whether any given distance should be considered 'near' or 'far'. The original sample was stratified by distance, whereby villages were selected in equal proportions (within each district) from villages within an area 15–20km distant from the Millennium Villages and from villages beyond this distance. This stratification is depicted in Figure 10, where the red dots are MVs, the green dots are 'nearby' comparison villages and the blue dots are 'faraway' comparison villages. This stratification was not very accurate because it was not based on global positioning system (GPS) coordinates (which were not available at the time). Hence, in addition to the sample design's original subdivision between 'near' and 'far' villages, we considered a second subdivision which employed GPS coordinates to calculate distances and that weight distances by population density in the MV villages. This second subdivision consisted of a weighted average of the distance of each comparison village from all other MVs weighted by their population. The idea behind the weighting down of poorly populated MV villages is that any spill-over effect should be larger from a larger project community than from a smaller one.

Figure 10. Project and comparison communities⁶³



⁶² Following the analysis plan, as for other impacts, we do not investigate here the mechanisms determining the outcomes that we reserve to do at a later stage.

⁶³ Source: <http://www.copypastemap.com/> to plot coordinates from the Earth Institute on Google Maps.

172. To calculate spill-over effects we split the sample of comparison villages into two equal groups based on the two criteria of 'near' and 'far' outlined earlier. We then estimate the project effects in the usual way: using DD analysis and correcting for differences in baseline characteristics by inverse probability weights. We focus first on outcomes for which we know, from the previous analysis, that the project had an impact: per capita income and anthropometry. Table 41 shows DD coefficients for MV areas and 'nearby' areas using the four definitions described above. Impacts in MV areas and 'near' comparison villages are estimated using 'far' comparison villages as the comparison group. The impact of MV on income is larger once neighbouring villages are regarded as part of the intervention area. The impact in neighbouring villages (spill-over effects) appears to be about half the size of project effects. The impact is transmitted through all income sources except income from transfers. The two different distance measures employed seem to capture the same phenomena and the differences between the coefficients are rather small.

Table 44. Impact of the MVP on income by distance

	Original 'near' and 'far' subdivision		Population weighted average distance subdivision	
	MV	Near	MV	Near
Income	0.48*** (0.000)	0.30** (0.001)	0.49*** (0.000)	0.27** (0.004)
Agricultural income	0.045*** (0.000)	0.31 (0.103)	0.46*** (0.000)	0.23 (0.257)
Livestock income	0.030** (0.001)	0.14** (0.009)	0.30** (0.001)	0.15** (0.004)
Business income	0.53*** (0.000)	0.27*** (0.000)	0.52*** (0.000)	0.23*** (0.000)
Employment income	-0.04 (0.401)	0.04 (0.442)	-0.04 (0.422)	0.05 (0.411)
Transfers income	0.11** (0.029)	0.04 (0.917)	0.10** (0.032)	-0.01 (0.901)

The table reports DD coefficients of regression analysis based on cross-sectional models adjusted by IPW. *** is statistical significance at 1%, ** is 5% significance and * is 10% significance.

173. We find no spill-over effects of the intervention on nutritional status of children. Table 45 shows DD estimates of project impact on anthropometric indicators by distance. In this case we add another distance classification based on the distance of the comparison community from the nearest clinic supported by MV (last column of Table 45). The project does not have an impact on nutritional status of children in areas in the vicinity of MV clinics or in the geographical vicinity of MV villages as defined by the classification described above. The absence of strong differences may also be partly attributable to the fact that some other organisations and initiatives⁶⁴ are making quite considerable investments in several of the 'far' control villages. A likely effect of these investments will be to bridge the difference between the MVs and 'far' villages.

⁶⁴ These include Presbyterian Agricultural Station (PAS), Camfed, Girls – Participatory Approach to Student Success in Ghana (G-PASS) and the Global Partnership for Education (GPE).

Table 45. Impact of the MVP on nutritional status by distance

	Original 'near' and 'far' subdivision		Population weighted average distance subdivision		Subdivision based on distance from nearest clinic	
	MV	Near	MV	Near	MV	Near
Height-for-age	0.50** (0.004)	0.01 (0.983)	0.49** (0.012)	-0.08 (0.679)	0.49** (0.007)	0.04 (0.983)
Weight-for-age	0.12 (0.312)	-0.25* (0.097)	0.07 (0.618)	-0.036** (0.012)	0.11 (0.456)	-0.26* (0.098)
Weight-for-height	-0.30* (0.044)	-0.38** (0.006)	-0.30** (0.027)	-0.41*** (0.000)	-0.26* (0.063)	-0.30** (0.025)

174. The PRA study similarly finds that the expected spill-over effects predicted by the MVP's programme theory do not appear to be strong thus far. One would expect, from the programme theory, that progress would be stronger in the nearby control sites than in their faraway counterparts. Of the six comparator sets for which qualitative data sets are simultaneously available for MVs, CNs and CFs, the nearby control village is ahead of the faraway control village in only one case. Furthermore, of the worst performing sites in the PRA study, two of the bottom three (CN5, CN6 and CF6) are nearby control villages. Similarly, the RCA study finds the further away CVs to be making progress at a faster rate than the near CVs.

6.2 Institutional spread of MVP practices

175. Institutional 'spill-over' was not directly explored at the midterm point because for other institutions to adopt and sustain activities introduced by a project, they need time to study them and be convinced that they provide credible solutions to their existing problems. This will be explored further at the endline evaluation in 2016–17. At this stage, the institutional study found that several of the activities being undertaken by SADA-MVP are innovations adopted from other institutions and NGOs, and in some cases adapted. The community health volunteer concept was already a mainstream practice in the Ghana Health Service. MVP has adapted this by turning them from volunteers into workers, and by expanding their scope of work and giving them a monthly stipend, to the disenchantment of the Ghana Health Service district management teams. It is also a common practice by NGOs and communities to engage community teacher volunteers and pay them a monthly token, which is similar to what MVP has adopted. The community livestock worker and the VSLA concept was brought from Care International, where the SADA-MVP team leader was a staff member prior to his appointment. Transforming farmers' groups into cooperatives is not new to Ghana. The community-led total sanitation concept is an internationally recognised and well-documented approach to sanitation and is already being practised by many NGOs in Ghana.

176. One possible spill-over is the Ministry for Employment and Labour Relations' indication of its intention to adopt the CHW model piloted by the MVP. According to a SADA-MVP staff member, the minister informed Professor Jeffrey Sachs on one of his visits to Ghana of government plans to recruit 3,000 CHWs as part of the Ghana Youth Employment and Entrepreneurial Development Agency programme. Of this number, 800 will be posted to the savannah zone. After two years the programme will be reviewed. Although, as at November 2015 this has not started, a staff member of SADA-MVP said implementation plans for the scheme have recently been sent to the district and regional directorates of health. This may be an indication of the government's intention to carry it through,⁶⁵ although possibly in the new financial year starting in January 2016.

⁶⁵ The real test will be whether the plan is supported by a budget approved by Ghana's parliament and followed through with actual allocations.

Chapter 7. Institutional sustainability

177. One of the expected project outputs noted in DFID's Business Case for the support to the Millennium Village in Northern Ghana is: *Strengthened local institutions and community capacity – to increase the involvement and commitment by the local community and officials, and help to secure sustainability of the impacts.*⁶⁶ This implies that the sustainability of SADA-MVP interventions and their impact during and after project implementation will, to a large extent, be dependent on the level of ownership and control by community and district actors and institutions. Evaluations of integrated rural development projects implemented in the 1970s and 1980s point to the importance of institutional capacity and ownership for project implementation and sustainability of the impacts.⁶⁷ For this reason one of the studies conducted by the independent impact evaluation team at baseline and midterm was an assessment of the institutional and governance arrangements in place for the management and implementation of SADA-MVP activities during the project and after it comes to an end. The primary data for the midterm institutional report is from six focus groups of district officials that took place in the district capitals of the three project districts from 11–15 May 2015. These were triangulated with and complemented by information from the qualitative participatory rural appraisal (PRA) and the reality check approach (RCA) studies, and verified by interviews with SADA-MVP staff and information from SADA-MVP documents.

178. Based on this evidence base, this chapter highlights some of the emerging challenges of institutional sustainability. As the data collected to inform the assessment was conducted during 2015, it is still too early to draw definitive judgements on the post-project sustainability of project benefits. It is with this caveat that findings are presented on:

- The model of community workers and allowances
- The challenges of sustaining interventions after the project ends
- The capacity of communities to sustain interventions and impact
- Institutional support for communities, citizen groups, NGOs and FBOs
- Local government support including fiscal decentralisation

7.1 Sustaining SADA-MVP's model of community workers and allowances

179. While the district officials acknowledge the impact that SADA-MVP has made in the areas of consumption, income, food security, health, education and infrastructure, they expressed concerns about how the institutional arrangements that made these possible are going to be sustained after the project has come to an end. The overall SADA-MVP approach of working through the existing district structures and departments is generally appreciated and considered to be relatively cost effective compared with other district level external interventions. Some district staff were of the opinion that it was not only cost effective but would lead to a greater likelihood of work sustainability. The heads or representatives of sector departments noted that community level committees pertaining to their sectors had been identified and strengthened or created by SADA-MVP.

⁶⁶ DFID. 2011. Business Case: Millennium Village in Northern Ghana. p. 2.

⁶⁷ (i) DFID. 2004. EvSum438 Synthesis of Integrated Rural Development Projects, (https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67873/ev438s.pdf); (ii) The German Federal Ministry for Economic Cooperation and Development. 2005. Country Case Study: Integrated Rural Development Programme (IRDP) North Western Province, Zambia (iii) Cynthia Clapp-Wincek. 1985. Integrated Rural Development Projects: A Summary of the Impact Evaluations, Center for Development Information and Evaluation.

180. In addition, most of the heads or representatives of sector departments said that they have seen the rekindling, adaptation or recruitment of focal people at the community level. These include community health, education and veterinary volunteers/workers. However, on the whole, district officials were ambivalent about these, as they fear that while they can help immensely with distributing workloads and increasing community access to government services, the introduction of remunerations/allowances to these volunteers will be difficult to sustain when the project ends. Moreover, they think this may in future affect the expectations of these community volunteers, reduce the level of community spiritedness and damage positive traditions of communal labour and volunteering. The same concern was expressed with the monthly top-up allowance given to community health nurses (GH¢300), midwives (GH¢500) and medical assistants (GH¢500) working in SADA-MVP-supported health facilities. For instance, depending on the number of years worked, a SADA-MVP staff member said community health nurses earn an average monthly salary of GH¢1,031 (approximately £193 at an exchange rate of GH¢5.34 to £1) and on top of this, working under the SADA-MVP project provides a top-up allowance of GH¢300 (approximately \$56), or about 36% of the standard average salary. The average salary for a principal midwife is GH¢2,062 and the SADA-MVP top-up of GH¢500 is 24% of the salary.

7.2 Sustaining interventions after the project

181. District officials are also apprehensive about how the following initiatives will be sustained post project. These include:

- Ambulances that have been acquired for use in emergencies (with fuel provided by SADA-MVP)
- The fridges to store drugs and vaccines
- The registration of community members with the NHIS for free to boost utilisation of health facilities (N.B.: MVP project staff have explained that this was meant to be a one-off free registration. The aim was that community members would then go to a health facility on the NHIS card and, by doing so, better appreciate its value to them and therefore see the need to renew it subsequently)
- The recruitment of midwives for some health facilities
- The extra allowance given to community health nurses to motivate them to work in these deprived communities (as noted above)
- School feeding programmes
- The tractor and fertiliser subsidies
- The provision of motorbikes and fuel for the extension staff of the various departments of the district assemblies to undertake health awareness promotion activities, organise farmers into cooperatives, village savings and loans associations, and to monitor project activities.

182. Similarly, while the district officials acknowledge that the numerous infrastructure projects in health, education, water, sanitation, agriculture and roads (including culverts) undertaken by SADA-MVP have had an impact on the standard of living of the people in their districts, they worry about how these are going to be maintained. The RCA study team observed many new infrastructures that were already suffering from poor maintenance (e.g. broken new culverts, broken recreational equipment in schools, damaged and cracked walls and floors in schools and health centres), some of which people felt had been due to poor initial workmanship.

183. These concerns and apprehensions about how the SADA-MVP interventions and impact will be sustained hinge on three issues. The first is the extent of community capacity and ownership of the community level initiatives that SADA-MVP has engendered. The second is the extent to which SADA-MVP has built cooperative relationships with pre-existing citizens' groups and non-governmental and faith-based organisations operating in the districts. The third and most critical is the current very incomplete state of Ghana's decentralisation framework. These have been elaborated on in the next three sections.

7.3 Community capacity to sustain interventions and impact

184. The SADA-MVP approach to community capacity building and ownership described by district officials in the focus groups, and in MVP's reports, appeared predominantly top down. It is aimed mainly at rolling out government services. None of the district officials spoke of building community capacity or supporting communities to identify and pursue their own priorities and goals, despite this approach being officially described in the local government planning process documented in L.I. 1589. Community participation thus appears to be conceived by the district and SADA-MVP officials in the study as involving communities in MVP and government schemes, rather than also supporting them to be in control of their own development, what Pretty et al. (1994) refer to as 'transformative participation' as opposed to merely 'information giving' or 'consultation'.⁶⁸

185. The SADA-MVP staff gave two examples that illustrate how this kind of transformative participation could have been the dominant approach to working with communities, rather than the exception. The first example was when additional accommodation was needed to house teachers and nurses in Kinkadina and Jadema communities. SADA-MVP said that it did not have the funds to extend what it had already provided, but could support the two communities with materials, if they took the initiative. The two communities did so, using local materials, and with the project only providing cement for plastering the walls and metal sheets for roofing. Similarly, Zamsa community also took the initiative to construct an office and warehouse for the cooperative with material support from the project. These projects, according to the SADA-MVP staff, cost a quarter of what it would have cost SADA-MVP if it had given them to contractors to construct. Moreover, in these three communities, it is likely that there will be a greater sense of ownership and responsibility towards these projects than in places where the structures were given out to contractors. This was indeed noted in the RCA report, that such projects constructed by outsiders were seen as SADA-MVP or district assembly projects and, therefore, community members were making no effort to maintain them. It is striking that even though one of the project outcomes is "*Strengthened local institutions and community capacity...*" the staff of the Department of Community Development and Social Welfare spoke of being marginalised in SADA-MVP's community level activities.

7.4 Relationships with citizens' groups, NGOs and FBOs

186. In the qualitative baseline study on well-being it was reported that:

"The Presbyterian Agricultural Station at Sandema (PAS-S) and the Friends of Widows Foundation are well regarded in several of the Buili sites visited. PAS-S provides extension services to large-scale farmers as well as those owning small ruminants and fowls such as disabled persons and poor households. World Vision Ghana also provides support in the areas

⁶⁸ Pretty et al. (1994), IIED Participation Learning and Action. A trainer's guide adapted from Adnan et al. (1992); <http://www.partnerships.org.uk/pres/fitlog/sld007.htm>; <http://www.participatorymethods.org/method/levels-participation> (both retrieved 2 August 2015).

of health, agriculture, education, water and also education finance/supplies to poor children.”⁶⁹

The midterm RCA study also revealed a growing number of development projects active in the area, with many adopting approaches that seek to work with groups and community educators. Likewise, there are self-help groups in the communities pursuing the welfare of their members. Despite all these, there was very little evidence of partnership between MVP and these civil society organisations. This suggests that many of the potential human resources that could drive development within the project districts remain underutilised or untapped.

7.5 Limited fiscal decentralisation and inconsistent and inadequate revenue

187. Numerous accounts from the district officials and the project team in this study show how, to a large extent, Ghana’s local government system is only partially functional (uneven devolution of powers, high staff movement, limited fiscal decentralisation, inadequate and inconsistent revenue, etc.). There is broad acknowledgement that until certain aspects of the decentralisation policy are realised, interventions such as the MVP can only achieve modest, localised and even temporary gains. This has implications for the extent to which the district administrations can take over interventions initiated under SADA-MVP and sustain them. Of the many shortcomings of Ghana’s decentralisation, it may be argued that the most problematic are related to finance.
188. The district officials repeatedly referred to the inadequacy of funds in relation to what they are expected to do. With regard to fiscal decentralisation, only 10% (increased in 2014 by Parliament, up from 7.5%) of total revenues of Ghana are allocated to the District Assemblies Common Fund (DACF) to be shared among the 216 district assemblies for local development. The remaining 90% is controlled by central government. Even then, only 50–60% of the DACF is available for the assemblies to use directly. Some 40–50% is reserved for the MPs’ Common Fund, the Regional Coordinating Councils, the DACF Administrator for monitoring, training of district assemblies’ staff and for priority intervention projects or programmes like school feeding, waste management and the sanitation module of the Ghana Youth Employment and Entrepreneurial Development Agency, persons with disabilities, training and funds for cured lepers.⁷⁰
189. Moreover, from the information gathered from Builsa South District, the release of the DACF to the district assemblies is regularly six to nine months in arrears,⁷¹ sometimes it is less than expected and sometimes it never comes at all. This results in serious fiscal and budgetary disruptions and renders the annual district plans out-dated and of little use by the time funds are received. This in turn discourages district officials from planning, particularly for the long term. A recent study by SEND Foundation reviewed the Medium-term Development Plans (MTDPs) of 11 district assemblies and found that only 7.3% of the planned activities contained in MTDPs were actually implemented (SEND Foundation 2014).⁷² It has been argued that the district assemblies must do more to generate revenue internally and that their dependence on DACF transfers has

⁶⁹ Masset et al. 2013c. Millennium Village Impact Evaluation, Baseline Summary Report: Appendix D. Baseline Report – Well-being and Field Group Discussions Assessment, p. 126.

⁷⁰ SEND-Ghana. 2010. Making Decentralisation Work for the Poor. p. 15.

http://www.commonfund.gov.gh/index.php?option=com_content&view=article&id=337&Itemid=437 (retrieved on 3 August 15).

⁷¹ It was not possible to verify the information on DACF provided by Builsa South on the DACF website because information on it is not up to date. The DACF Administrator is supposed to report to Parliament on how DACF funds have been distributed during the preceding financial year. On its website it has posted the executive summary of the reports for 2010 and 2011 but nothing after that. Similarly, it used to post quarterly press releases of the distribution of the common fund to the districts. There have not been any new quarterly press releases after that of the 4th quarter of 2012.

⁷² SEND Foundation. 2014. Bringing Development to the Doorstep of Citizens: The Contribution of the District Assemblies Common Fund, p. 7.

served as a disincentive to local revenue collection (Yemedi 2014).⁷³ Builsa South District (newly created in 2012) was only able to mobilise GH¢51,454 (equivalent to £9,000) for 2014. The district administrations and the technical departments spoke about having to rely on SADA-MVP for motorbikes and monthly fuel allowances to perform the roles and responsibilities agreed on in the memorandum of understanding they had signed. Thus, instead of projects like SADA-MVP lending a hand to assist the district assemblies and the departments with what they are doing, they have become lifelines that enable them to perform their most basic functions, such as monitoring and support visits to their staff in the field. Inevitably there was a sense of district officials hoping that the SADA-MVP will be extended (as one district official said) or that another donor project will come to their district when SADA-MVP ends.

⁷³ Yemedi, S. 2014. *Policy Brief Decentralization and Gender and Social Inclusion in Ghana*, STAR-Ghana.

Chapter 8. Overall findings and conclusions

190. Overall, the quantitative analysis finds that the project had moderately positive impacts at the midterm. The project did not have a sizable impact on the MDGs but several impacts are visible on other indicators of well-being:
1. The project produced a large increase in per capita income mostly driven by an increase in agricultural incomes. The increase in incomes appears to be supported by a large improvement in self-reported food security. Household expenditure increased to a lower extent, which could be due to income gains being saved or invested – though the PRA and RCA studies do not reach the same conclusion. Since poverty measures are based on household expenditure, we find little impact on poverty.
 2. There was a considerable improvement in the nutritional status of children under five. The project produced a modest reduction in mortality rates but we highlight the difficulty to calculate these effects after only two years of data. The project did not affect prevalence of anaemia.
 3. The project produced only a modest increase in school attendance rates and no improvements on the quality of education as measured by mathematics and English test scores.
 4. There are some signs that the project is having a higher impact in the Builsa district.
 5. We found no difference in impacts by sex.
 6. Our exploratory analysis is suggestive of potential spill-over effects to neighbouring communities in relation to income indicators but not in relation to nutritional status.
191. All the observed effects were analysed as ‘project effects’ independently of the mechanisms generating the outcomes. For this reason, some results may appear surprising or contradictory. We aim to resolve these contradictions and surprises by analysing more closely the determinants of the outcomes with the help of qualitative data and by using more extensively the quantitative information on participation in programme activities and on intermediate outcomes. Given the multiplicity of outcomes affected by the programme, we reserve to conduct this more detailed analysis in separate papers for selected outcomes of interest.
192. The RCA study, among other things, provides people with opportunities to share their experience of change and their aspirations for the future. Some of the underlying assumptions for the MVP such as their focus on small farms and livestock rearing are then brought into question considering people’s aspirations to leave farming and their need for reliable cash incomes. As well as people’s increasing concern about the unpredictability of the climate (particularly, late rains), high costs of inputs, declining soil fertility and risks, they also talk a lot about the increasing need for cash. In 2013, the RCA families (who are the poorer families in communities) were largely cashless but there has been quite a shift in the last two years. Our observations confirm that they have more demands for cash: in 2013, cash was needed only for the purchase of seasonings, cooking oil, batteries and agricultural inputs (although much of this was in exchange for produce). The same families are now wanting to buy phone credit (especially to keep in touch with migrant workers), health insurance, pay school ‘levies’, milling services (which they used to do at home before electricity came to the village), electricity installation and for social obligations such as funerals and naming ceremonies (increasing demand for cash not kind); as well as to meet their increasing consumer tastes for snacks, alcohol, cosmetics, skin and hair products, phone downloads and fashion clothes. This is quite a noticeable shift and one they highlight.

193. These cash demands cannot, people tell us, be met by farming. Provision of infrastructure and services is questioned as being insufficient to change behaviour and to address some deeply held cultural and traditional practices, some of which may in fact have a good basis for being preserved. Based on these observations, there seems to be insufficient attention to context in the design of the MVP interventions and the different effects of the same intervention in different situations. The 'control' villages have attracted attention from non-MVP development organisations and political patronage, and this RCA study indicates that the two 'control' villages seem poised to become quite successful as a result.
194. In large part, the PRA findings align with those of the survey work. From the PRA study, the MVs appear to be demonstrating initial gains in areas such as school enrolment, teacher attendance and pupil attendance/retention; attended deliveries, emergency care and malarial morbidity; crop yields, livestock health and food security; and school sanitation. However, when we focus on movements into and out of poverty, as perceived by the focus groups, the MVs' superiority as a predictor of improved well-being becomes less certain.
195. Another surprising finding is that the expected spill-over effects predicted by the MVP's programme theory do not appear to be strong thus far. In several areas, the faraway controls are ahead of the nearby controls. Typically too, the endowment changes that are emerging in the nearby controls thus far (mainly infrastructure, equipment and technical assistance improvements) have been initiated by relevant state institutions and/or by some NGO or project rather than through self-help (as one might expect in the nearby control villages, through a propagation effect).
196. In terms of replicability, the authors think that the degree to which the incentives being provided by the project (notably allowances, transport and dependable fuel support – along with the additional recruitments and intensive monitoring to secure the gains) can and will be sustained by the pivotal state institutions (e.g. Ghana Education Service, Ghana Health Service and the Ministry of Food and Agriculture) and will be important to watch in the future.
197. The importance of strengthened and committed local institutions and communities in sustaining a project's interventions and impacts is acknowledged in DFID's Business Case for its support to the MVP. District officials in the three project districts appreciate what has been achieved so far under the project, and its overall approach of working through existing district structures and departments. At the same time, they are concerned about the sustainability of the various initiatives that are fully funded by the MVP when the project comes to an end. They are also concerned about how the various top-up allowances to community health nurses, midwives and medical assistants, as well as payments to community health/education/livestock workers (who were previously willing to be volunteers), which have been introduced under the MVP, will be sustained. In addition, they predict that these will have a future effect on the expectations of such community volunteers, as well as the general level of community spiritedness, and the potential damage that may be done to positive traditions of communal labour and volunteering. These concerns appear justified at present because the district institutions have inadequate and inconsistent revenue due to weak fiscal decentralisation policy. Indeed, districts are themselves dependent on SADA-MVP for funds, even to fuel their vehicles for monitoring and support visits to their field staff.
198. At present, the SADA-MVP approach to working with communities seems less empowering and transformative than might have been hoped – at least from the viewpoint of the district officials interviewed. While there is much MVP work at the community level, communities could have perhaps benefited from greater support to undertake some of the community level infrastructure

projects, which would have given them a greater sense of ownership. Indeed, although there are many citizens' groups, NGOs and FBOs operating in the district, there is currently little evidence that SADA-MVP has forged partnership relationships with them in the hope that they will continue or adopt some of its initiatives when the MVP ends. Of course, these are tentative conclusions and a better assessment will be possible towards the end of the project.

Bibliography

- Anderson, M.B, D. Brown and I. Jean. 2012. *Time to Listen: Hearing People on the Receiving End of International Aid*. Cambridge MA: CDA.
- Angrist, J.D. and J-S. Pischke. 2009. *Mostly Harmless Econometrics*. Princeton, NJ: Princeton University Press.
- Barrera-Osorio, F., M. Bertrand, L.L. Linden and F. Perez-Calle. 2011. Improving the design of conditional transfer programs: evidence from a randomised education experiment in Colombia. *American Economic Journal: Applied Economics* 3 (April), 167–95.
- Becker, G. 1991. *A Treatise on the Family*. Cambridge, MA: Harvard University Press.
- Bump, Jesse B., Michael A. Clemens, Gabriel Demombynes, and Lawrence Haddad. "Concerns about the Millennium Villages project report." *The Lancet* 379, no. 9830 (2012): 1945.
- Bump, Jesse B., Michael A. Clemens, Gabriel Demombynes, and Lawrence Haddad. "Concerns about the Millennium Villages project report." *The Lancet* 379, no. 9830 (2012): 1945.
- Clemens, Michael A., and Gabriel Demombynes. "Multisector intervention to accelerate reductions in child stunting: an independent critique of scientific method." *American Journal of Clinical Nutrition* 95, no. 3 (2012): 774-775.
- Cowell, F.A. 2011. *Measuring Inequality*. Oxford: Oxford University Press.
- Deaton, A. 1997. *The Analysis of Household Surveys*. Washington, DC: The World Bank.
- Earth Institute Surveys, 2012, 2013, 2014.
- Fink, G., M. McConnell and S. Vollmer. 2014. Testing for heterogeneous treatment effects in experimental data: false discovery risks and correction procedures. *Journal of Development Effectiveness* 6 (1):44–57.
- Ghana Statistical Service. 2011. *Ghana Multiple Indicator Cluster Survey with an Enhanced Malaria Module and Biomarker, Final Report*. Accra, Ghana.
- Glewwe, P. 2000. Household roster. In M. Grosh and P. Glewwe (eds) *Designing Household Survey Questionnaires for Developing Countries*. Washington, DC: World Bank.
- Grossbard, A. 1980. The Economics of Polygamy. In J. Da Vanzo and J. Simon (eds) *Research in Population Economics Vol II*. Greenwich, CT: JAI-Press.
- Guyer, J.I. 1981. Household and community in African studies. *African Studies Review* 24 (2/3):87–137.
- Heckman, J. J., Smith, J. and N. Clements 1997. Making the Most Out Of Programme Evaluations and Social Experiments: Accounting For Heterogeneity in Programme Impacts. *Review of Economic Studies* 64: 4875–35.
- Hill, K. 2013. Direct estimation of child mortality from birth histories. In T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timmaeus and B. Zaba (eds) *Tools for demographic estimation*. Paris: International Union for the Scientific Study of Population.
- Hill, P. 1986. *Development Economics on Trial: the Anthropological Case for a Prosecution*. New York: Cambridge University Press.
- IEG. 2010. *What Can We Learn from Nutrition Impact Evaluations: Lessons from a Review of Interventions to Reduce Child Malnutrition in Developing Countries*. Washington, DC: World Bank.
- Imbens, G.W. and D.B. Rubin. 2015. *Causal Inference for Statistics, Social and Biomedical Sciences*. New York: Cambridge University Press.

- Imbens, G.W. and J.M. Wooldridge. 2009. Recent Developments in the Econometrics of Impact Evaluation. *Journal of Economic Literature* 47 (1):5–86.
- Masset, E., L. Haddad, A. Cornelius and J. Isaza-Castro. 2012. Effectiveness of interventions that aim to improve nutritional status of children: systematic review. *British Medical Journal* 344 (d8222):1–7.
- Masset, E., C. Barnett, A. Archarya and T. Dogbe. 2013a. An impact evaluation design for the Millennium Villages Project in northern Ghana. *Journal of Development Effectiveness* 5(2): 137–57.
- Masset, E., D. Jupp, D. Korboe, T. Dogbe and C. Barnett. 2013b. *Impact Evaluation of a new Millennium Village in Northern Ghana: Initial Design Document*. Hove: Itad.
- Masset, E., D. Jupp, D. Korboe, T. Dogbe and C. Barnett. 2013c., *Millennium Village Impact Evaluation, Baseline Summary Report: Appendix D. Baseline Report – Well-being and Field Group Discussions Assessment*. Hove: Itad. p. 126
- Pretty, J., I. Guijt, J. Thompson and I. Scoones (1995). *Participatory Learning and Action: A Trainer's Guide*. IIED Participatory Methodology Series, IIED: London.
- Pronyk, Paul M., Maria Muniz, Ben Nemser, Marie-Andrée Somers, Lucy McClellan, Cheryl A. Palm, Uyen Kim Huynh et al. "The effect of an integrated multisector model for achieving the Millennium Development Goals and improving child survival in rural sub-Saharan Africa: a non-randomised controlled assessment." *The Lancet* 379, no. 9832 (2012): 2179-2188.
- Remans, Roseline, Paul M. Pronyk, Jessica C. Fanzo, Jiehua Chen, Cheryl A. Palm, Bennett Nemser, Maria Muniz et al. "Multisector intervention to accelerate reductions in child stunting: an observational study from 9 sub-Saharan African countries." *American Journal of Clinical Nutrition* 94, no. 6 (2011): 1632-1642.
- Rosenbaum, P.R. 1984. The consequences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of the Royal Statistical Society. Series A (General)* 147 (5): 656–66.
- Savannah Accelerated Development Authority. 2013. Annual Report on the Millennium Village Project in Northern Ghana, December 2013.
- Savannah Accelerated Development Authority. 2014. Annual Report on the Millennium Village Project in Northern Ghana, December 2014.
- Savannah Accelerated Development Authority. 2015. Mid-Year Report on the Millennium Village Project in Northern Ghana, July 2015.
- Salmen, L.F. 1998. Toward a listening bank: review of best practices and efficacy of beneficiary assessments. *Social Development Papers* 23. Washington, DC: World Bank
- Schochet, P.Z. 2008. *Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- SEND Foundation. 2014. *Bringing Development to the Doorstep of Citizens: The Contribution of the District Assemblies Common Fund*. p. 7.
- SEND-Ghana. 2010. Making Decentralisation Work for the Poor. p. 15. Available at: http://www.commonfund.gov.gh/index.php?option=com_content&view=article&id=337&Itemid=437 (retrieved on 3 August 15).
- Shutt, C. and L. Ruedin (SDC). 2013. *How to Note Beneficiary Assessment*. Berne: Swiss Agency for Development Cooperation.

- Whitehead, A. 2006. Persistent poverty in north east Ghana. *Journal of Development Studies* 42 (2):278–300.
- Yemedi, S. 2014. *Policy Brief Decentralization and Gender and Social Inclusion in Ghana*, STAR-Ghana.
- Wanjala, Bernadette M., and Roldan Muradian. "Can big push interventions take small-scale farmers out of poverty? Insights from the Sauri Millennium Village in Kenya." *World Development* 45 (2013): 147-160.