MATHEMATICA Policy Research

REPORT

FINAL REPORT

Evaluation of the Burkina Faso Agriculture Development Project: Design Report

June 6, 2017

Christopher Ksoll
Chantal Toledo
Seth Morgan
Anca Dumitrescu
Kristen Velyvis

Submitted to:
Millennium Challenge Corporation
875 Fifteenth St., NW
Washington, DC 20005
Project Officer: Jack Molyneaux
Contract Number: MCC-16-CON-0029

Submitted by:
Mathematica Policy Research
1100 1st Street, NE
12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Christopher Ksoll
Reference Number: 50284.01.240.032.000

This page has been left blank for double-sided copying.

CONTENTS

FIGURE

TABLE

This page has been left blank for double-sided copying.

ACRONYMS

ADP	Agriculture Development Project
AEA	American Economic Association
APD	*Agence de Partenariat pour le Développement* (post-Compact successor to MCA-BF)
AMVS	*Autorité de la Mise en Valeur de la Vallée du Sourou* (Authority for the Development of the Sourou Valley)
ARF	Access to Rural Finance
BRIGHT	Burkinabé Response to Improve Girls' Chances to Succeed
CATG	*Centre d'Appui Technique et de Gestion* (private consulting firm that provides technical assistance to water user associations)
CLE	*Comité local de l'eau* (local water committee)
CRD	Comparative Regression Discontinuity Design
DA	Diversified Agriculture (Activity)
DRAAH	*Direction Régionale de l'Agriculture et des Aménagements Hydrauliques* (Regional Directorate of Agriculture and Hydraulic Installations)
ERR	Economic Rate of Return
FAO	Food and Agriculture Organization (United Nations)
CFA	Franc CFA
CFE	*Contribution Financière en matière d'Eau* (water usage fee)
FGD	Focus group discussion
GDP	Gross Domestic Product
INERA	*Institut de l'Environnement et de Recherches Agricoles*
IRB	Institutional Review Board
IRIS	Center for Institutional Reform and the Informal Sector, University of Maryland (defunct)
IWRM	Integrated Water Resource Management
LONAB	*Loterie Nationale Burkinabè* (National Lottery of Burkina)

MAAH	*Ministère de l'Agriculture et des Aménagements Hydrauliques* (Ministry of Agriculture and Hydraulic Installations)
MRAH	*Ministère des Ressources Animales et Halieutiques* (Ministry of Livestock and Fishing)
MCA-BF	Millennium Challenge Account–Burkina Faso
MCC	Millennium Challenge Corporation
MDI	Minimum Detectable Impact
MIS	Market Information System
O&M	Operations and Maintenance
OLS	Ordinary Least Squares
PAP	*Personne affectés par le project* - Persons affected by the project. In the case of the Di perimeter, these were persons who lost land to the construction of the perimeter and who later received irrigated land in compensation.
Non-PAP	Person not affected by the project
RCT	Randomized Control Trial
RD	Regression Discontinuity Design
SDAGE	*Schéma Directeur d'Aménagement et de Gestion de l'Eau* (water management plans)
SIMCA	*Système d'Information sur les Marchés Agricoles du MCA* (MCA Market Information System on Agricultural Markets)
SONABEL	*Société Nationale d'Électricité du Burkina* (Burkina Faso's Electricity Company)
TA	Technical Assistance
USAID	U.S. Agency for International Development
WMI	Water Management and Irrigation (Activity)
WSC	Within-study Comparison
WUA	Water User Association

I. INTRODUCTION

In Burkina Faso, as in much of Africa, the agriculture sector is a critical component of the economy. A large fraction of the country's population depends upon farming and other agriculture-related activities for their livelihoods and their own consumption. As of 2011, agriculture contributed nearly a third of the country's annual gross domestic product (GDP), with total production estimated at just under $3 billion annually (FAPDA 2014). The sector also employs 80 percent of Burkina Faso's workforce, primarily on small subsistence farms of five hectares or less (USAID Burkina Faso 2015; FAPDA 2014). Despite its prominent role in the country's economy, the agriculture sector is characterized by low crop and livestock productivity (USAID Burkina Faso 2015). Burkina Faso also is a net food importer (Chauvin et al 2012). Low agricultural productivity contributes to extreme poverty in Burkina Faso, which is one of the poorest countries in the world with a GDP per capita of $634 (FAPDA 2014).

Agricultural improvements are needed for economic growth and poverty reduction in Burkina Faso. However, the sector faces several challenges. In particular, the level of rainfall is low and variable (USAID Burkina Faso 2015). Annual rainfall in Burkina Faso averages around 750 millimeters, with the northern Sahelian area typically receiving less than 600 millimeters while the southern Sudanian region receives up to 1,200 millimeters. The rainy season in Burkina Faso normally lasts from April or May to September or October. However, rainfall has been gradually decreasing since the severe droughts of the 1970s (Sally et al. 2011). Inadequate rainfall necessitates irrigation for successful agriculture, yet infrastructure is poor and farmers' access to irrigated water is low (FAPDA 2014). Less than 1 percent of cultivated land in Burkina Faso is equipped for irrigation (FAO 2016). Other challenges facing the country's agriculture sector include limited knowledge and capacity among farmers, land tenure insecurity, poor roads and other transportation infrastructure, and limited access to credit. Burkina Faso's economy is also susceptible to regional trade shocks and volatile food and fuel prices (FAPDA 2014; USAID Burkina Faso 2015).

In response to the challenges facing Burkina Faso's agriculture sector, the Millennium Challenge Corporation (MCC) invested in the Agriculture Development Project (ADP) as part of the Burkina Faso Compact. The project's objectives were to improve agricultural productivity, increase incomes among farmers and livestock producers, and support economic development. The ADP was a five-year effort, implemented from 2009 to 2014, and was comprised of three activities: (1) Water Management and Irrigation (WMI), (2) Diversified Agriculture (DA), and (3) Access to Rural Finance (ARF). The ARF activity does not fall under the scope of this evaluation.[1]

The WMI activity attempted to increase agricultural incomes through several initiatives, particularly by constructing an irrigated perimeter in the Di Department and by providing capacity building and technical assistance (TA) for operations and maintenance (O&M) of the irrigation perimeters in Sourou province. The TA and support for capacity building provided in Sourou included (1) establishing and training water-user associations (WUAs) and (2) providing

[1] MCC separately contracted the evaluation of the ARF activity. A2F completed an evaluation of the ARF activity in 2015 (A2F 2015).

TA to the Autorité de Mise en Valeur de la Vallée du Sourou (AMVS) for its action plan. In addition, it aimed to preserve agricultural livelihoods by supporting integrated water resource management (IWRM) in the Mouhoun and Comoé basins.

The DA activity attempted to increase farmer incomes by improving agricultural productivity. Its components included (1) providing farmers with training on rain-fed and irrigated production, (2) providing training to producer associations and agribusinesses, (3) improving veterinary services and providing livestock training, (4) establishing a market information system (MIS) and information centers, (5) establishing and training market committees, and (6) rehabilitating rural markets. The two activities were designed to work in an integrated way to increase agricultural productivity for beneficiaries and enhance their ability to sell their crops. The five-year compact closed on July 31, 2014.

Mathematica Policy Research is designing and implementing an evaluation of the WMI and DA activities to determine their impact on the use of improved agricultural technologies, agricultural production, household income, land tenure security, maintenance of irrigation infrastructure, and IWRM. MCC contracted with Mathematica in July 2016. MCC had previously contracted with two previous evaluators, the IRIS Center followed by IMPAQ International, to evaluate the WMI and DA activities.

Mathematica staff conducted site visits to project areas in Burkina Faso in October 2016 in preparation for developing the evaluation design. The ADP evaluation, which is described in this report, will address research questions on project outcomes, implementation, and sustainability. We propose a mixed-methods evaluation that will employ quantitative and qualitative evaluation methods and will draw on a variety of quantitative and qualitative data sources. We will conduct household surveys crop cutting surveys, focus group discussions, key informant interviews, in-depth interviews, site visits, document reviews, and analysis of administrative data.

The remainder of this design report provides context for the project and presents the evaluation design of each activity in further detail. Chapter II describes the compact as well as the goals and implementation of each of the projects to be evaluated. Chapter II also discusses the process by which evaluation design activities were prioritized. Chapter III reviews relevant literature on irrigation, farmer training and land tenure and discusses gaps in the literature. Chapter IV details our overall evaluation strategy and evaluation questions, analyzes the previous evaluation designs, presents our design for each of the evaluations, and discusses potential risks and challenges. Chapter V, describes our data collection plans. Chapter VI concludes with a discussion of administrative concerns, including institutional review board (IRB) requirements, the dissemination plan, and the evaluation timeline.

II. OVERVIEW OF THE COMPACT, ADP ACTIVITIES, AND EVALUATION

This chapter provides context for Mathematica's evaluation by describing the Burkina Faso Compact and providing background on the project locations and beneficiaries targeted. We then describe the program logic for the WMI and DA activities. Finally, we provide an overview of the evaluation.

A. Overview of the Burkina Faso Compact

With the goal of reducing poverty through economic growth, MCC entered into a five-year, $480.9 million compact with the Government of Burkina Faso in July 2009. The compact attempted to reach this goal by investing in four areas: (1) agriculture, (2) land tenure, (3) roads, and (4) girls' education. Accordingly, the compact was comprised of four separate projects: (1) the ADP, which aimed to improve agricultural outcomes; (2) the Rural Land Governance Project, which aimed to improve land tenure security and land management in rural areas of Burkina Faso and to increase efficiency of land institutions and access to them; (3) the Roads Project, which aimed to enhance access to markets through investments in the road network; and (4) the BRIGHT 2 Schools Project, which aimed to increase school enrollment and retention rates among girls. By the end of the compact, over 98 percent of anticipated funds had been disbursed.

The ADP consisted of the WMI and DA activities, which Mathematica will evaluate, as well as the ARF activity, which supported a lending facility for farmers and small- and medium-sized rural agricultural enterprises and aimed to improve the capacity of financial institutions and increase access to credit. However, due to low take-up of ARF services and limited progress toward the project's targets, MCC terminated the activity in July 2013 (MCC 2016c).

B. Program logic

The program logic for the WMI and DA activities of the ADP, presented in Figure II.1, describes the problem that motivates the project; lists the activities, sub-activities, and outputs; and links them to short- and long-term outcomes and impacts.

The program logic describes how the program was designed to address low agricultural productivity in farming and livestock rearing as well as low incomes from sales of agricultural production in the Sourou Valley and the Comoé Basin, the two primary project areas. These two areas, near the country's borders with Mali and Côte d'Ivoire, respectively, are both predominantly rural areas located outside provincial capitals. Throughout the Sourou Valley, agriculture is the principal activity for over 90 percent of the population, the majority of whom also keep livestock. Cereals, legumes, and rice are the main crops, with rice being harvested primarily by female farmers (MCC 2008b). Before the construction of the Di perimeter, irrigated perimeters in the Sourou Valley covered about 3,817 hectares, primarily near Niassan. With the exception of rice, farmers in the Comoé Basin grow largely the same crops; however, some farmers are also involved in livestock, fishery, and forestry. Agriculture in the area has traditionally been rain-fed, but government programs and nongovernmental organizations (NGOs) began introducing irrigation infrastructure and other new technologies in the late 1990s (MCC 2008a).

Underlying the logic for the ADP is the need for multiple approaches to supporting agriculture at each step of the value chain, including land tenure, irrigation, animal health, farming and livestock techniques, and market opportunities. The various components of the ADP were designed to work together to address the varied challenges facing farmers in Burkina Faso. We discuss the assumptions underlying the program logic in further detail in the Evaluability Assessment Report (Ksoll and Toledo 2016)

This page has been left blank for double-sided copying.

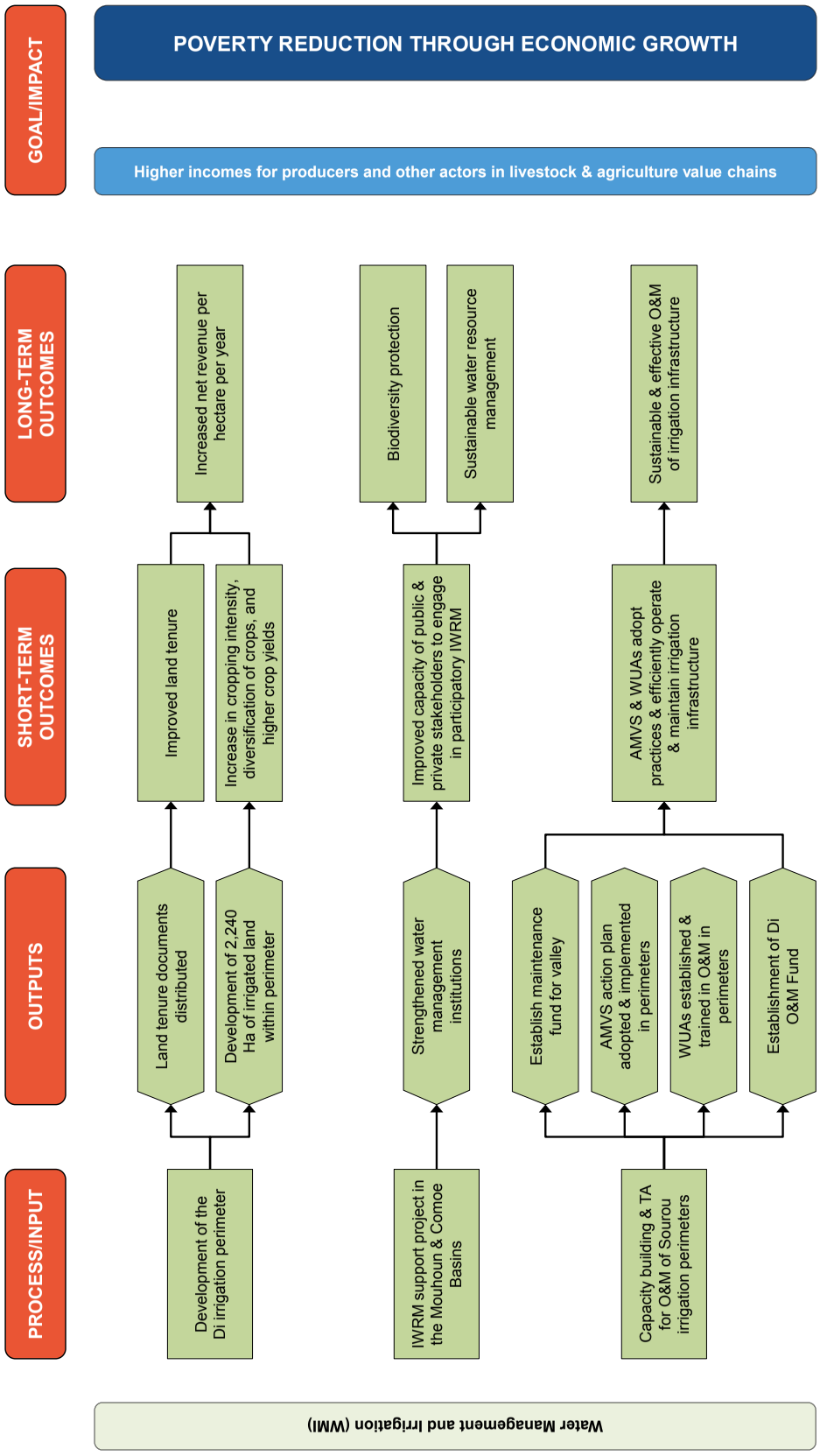
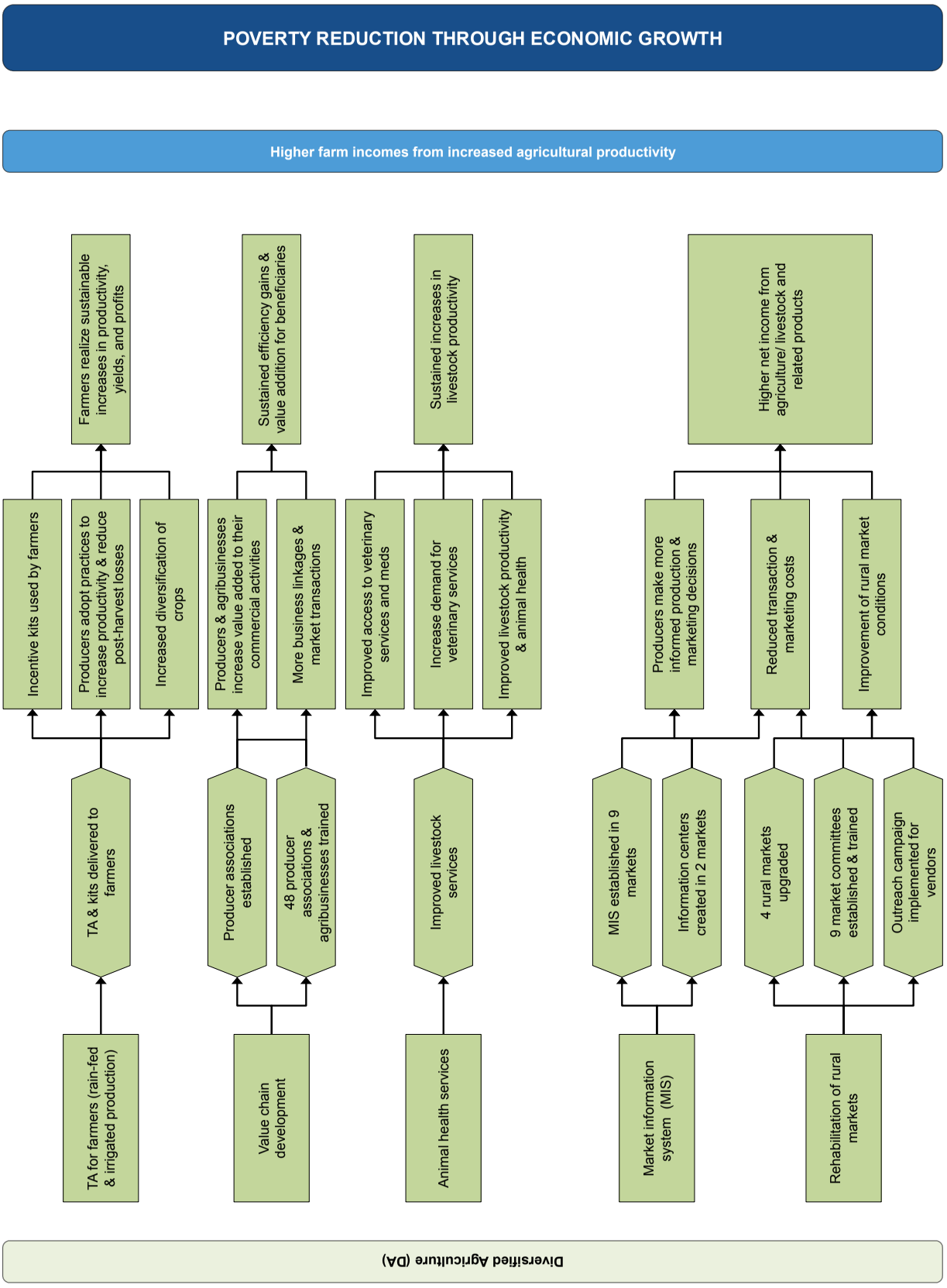


Figure II.1. Program logic


POVERTY REDUCTION THROUGH ECONOMIC GROWTH
Higher farm incomes from increased agricultural productivity
Farmers realize sustainable increases in productivity, yields, and profits
Sustained efficiency gains & value addition for beneficiaries
Sustained increases in livestock productivity
Higher net income from agriculture/ livestock and related products
Incentive kits used by farmers
Producers adopt practices to increase productivity & reduce post-harvest losses
Increased diversification of crops
Producers & agribusinesses increase value added to their commercial activities
More business linkages & market transactions
Improved access to veterinary services and meds
Increase demand for veterinary services
Improved livestock productivity & animal health
Producers make more informed production & marketing decisions
Reduced transaction & marketing costs
Improvement of rural market conditions
TA & kits delivered to farmers
Producer associations established
48 producer associations & agribusinesses trained
Improved livestock services
MIS established in 9 markets
Information centers created in 2 markets
4 rural markets upgraded
9 market committees established & trained
Outreach campaign implemented for vendors
TA for farmers (rain-fed & irrigated production)
Value chain development
Animal health services
Market information system (MIS)
Rehabilitation of rural markets
Diversified Agriculture (DA)

C. Overview of the WMI activity

Given the low and variable levels of rainfall in Burkina Faso and its limited irrigation infrastructure, the WMI activity was designed to improve water availability and delivery, flood control, and dam safety. The WMI activity was a $103.9 million investment, representing nearly three-quarters of the ADP.[2]

The largest investment under the WMI activity was the construction of a 2,240-hectare, irrigated perimeter in Di, a department in Sourou Province. It is located on the east bank of the Sourou River. Land at the site was previously developed into large agricultural fields and open grazing areas for livestock. The perimeter is divided into three sectors: (1) the North Sector, near the village of Bouna; (2) the Central Sector, near the village of Oué; and (3) the South Sector, near the village of Di. The land within the perimeter was designated for polyculture (75 percent), rice (23 percent), or mixed cultivation (2 percent) (Appendix Figure A.1). In addition to constructing the irrigation scheme and distributing irrigated land, this sub-activity included the distribution of land tenure documents, formalizing land rights for its beneficiaries. The combination of irrigation and land tenure was intended to increase land investments, cropping intensity, diversity of crops, and crop yields among farmers who received land.

The construction of the perimeter displaced some of the local population, known as persons affected by the project (PAPs). All of the PAPs (1,473 people) received irrigated land, which comprised about half of the total amount of land in the perimeter. The size of the plots that PAPs received in compensation was based on the estimated value of the parcel they lost. Because irrigated land has higher economic returns than the land that was lost, PAPs received a smaller plot in compensation than they had originally owned, but they received full ownership over this land. The ADP distributed additional land as leaseholds to PAPs if household size was large enough to cultivate more land.[3] The ADP provided female members of PAP families who were not PAPs themselves a 500-square meter plot; leases were provided to them through female producer groups. Additional land benefited "youths"—or household members' children who were over age 15 and were not PAPs. Leases for this land were provided through youth producer groups.[4]

[2] In addition to the investments Mathematica is evaluating, which are described in detail below, the WMI activity also included the rehabilitation of the Léry Dam on the Sourou River. This component was designed to improve the safety of the dam and associated infrastructure and to improve flood control in the Sourou Valley.

[3] Based on the leasehold documents, the leasehold is contracted for a period of fifty years. The leaseholder is obliged to pay an annual rent of 15,000 Francs CFA/ha, which the government can increase by ten percent every five years. The leaseholder can transfer the lease or sublet the land with the approval of the Ministry of Finance. This approval is automatically granted a month after the demand for approval has been submitted, unless the Ministry of Finance issues an objection to the transfer or rental. The lease can also be inherited. The lease can be terminated if the leaseholder does not pay the annual rent or the water user association fees or does not exploit the land for agricultural purposes. In addition, the leasehold must abide by the Di perimeter by-laws. The evaluation will investigate whether the terms of the leasehold are implemented in practice.

[4] Although the category of youths included both male and female youths in principle, female youths would typically have received small plots as women and would be part of the female groups. As a result, there is only one mixed-gender youth group.

There were two additional groups of beneficiaries who received land: (1) all households in disadvantaged neighboring communities that were not PAPs, called non-PAPs from disadvantaged villages, and (2) Di Lottery beneficiaries. The ADP provided 317 hectares of land to non-PAPs from disadvantaged villages. Finally, the ADP distributed 710 hectares of land via a lottery, called the Di Lottery. To be eligible for the lottery, applicants had to meet the following criteria: (1) not displaced by the construction of the perimeter, (2) age 18 or older; (3) resided in one of the six provinces of the Boucle de Mouhoun region, and (4) had at least two other people age 15 or older in the same household who could help cultivate the land. The lottery offered plots of 2 hectares for rice or 1 hectare for polyculture. About one-third of the households participating in the lottery received land (MCC 2016c). Appendix Figure A.2 shows a map of the allocation of plots by beneficiary type.

The development of the Di perimeter required the construction of irrigation and drainage canal networks, seven pumping stations, guard drains, a levee, and roads and paths throughout the perimeter. Its irrigation scheme requires 30 million cubic meters of water per year, or about 10 percent of the water available in the Sourou Valley (MCC 2008b). This significant requirement was one of the reasons why MCC included an IWRM component in the WMI.

The WMI also included the provision of capacity building and TA for sustainable and effective management of the irrigation infrastructure. It created and trained seven WUAs in the Di perimeter and one each in nine other perimeters in Sourou Valley. It also provided capacity building to AMVS—the GOBF agency in charge of maintaining primary canals in Sourou Valley and supervising the WUAs—to implement a set of reforms contained in the AMVS action plan.

The newly trained WUAs are responsible for maintaining all infrastructure inside their perimeters, including the canals and drains, internal roads, and pumping stations. They received O&M training in financial management, water distribution, and system maintenance. However, some WUAs from the Di perimeter had not completed the training by the time the compact closed. The Centre d'Appui Technique et de Gestion (CATG)—a private consulting firm that the ADP helped create—was tasked with providing the training after the compact ended. Implementation delays also affected the planned financing for the WUAs. They were supposed to receive direct deposits to fund maintenance for the first two years. However, by the time the WUAs were conducting maintenance activities, the deposits were considered unallowable financing of post-compact activities, and thus could not be provided (MCC 2016a).

The final sub-activity of the WMI activity was the IWRM support project in the Mouhoun and Comoé basins, which sought to create, strengthen, and train water management institutions, thereby improving public and private stakeholder capacity to engage in participatory IWRM. The ultimate objectives of this sub-activity were biodiversity protection and sustainable water management—in particular, rational and equitable resource allocation and reduced conflict over resources. Under this sub-activity, 10 Comité Local de l'Eau (CLEs) were formed and trained, 7 in Mouhoun and 3 in Comoé, as well as a basin committee in both areas. TA and equipment were also provided to two departments of water resources and basin-level water agencies, and a basic hydrological model was established. The basin committees received funding to develop IWRM plans for water use and protection and rehabilitation of water resources within their basin, the Schéma Directeur d'Aménagement et de Gestion de l'Eau (SDAGEs). Much of the

IWRM implementation occurred behind schedule. The development of the SDAGEs, for example, was scheduled to take 21 months, but ultimately took 38 months; the plans were not adopted until the close of the compact. Similarly, the creation of and training for the CLEs took 39 months, compared to the planned 24 months. The two basin committees were put into place two years later than scheduled. Although the IWRM sub-activity initially focused on the institutional environment, the difficulties in effecting change in this environment led to a greater focus on supporting participatory methods in IWRM, including consulting water users at local levels (MCC 2016a).

D. Overview of the DA activity

In response to several of the other challenges confronting Burkina Faso's agriculture sector, the DA activity was designed to complement the WMI. It focused primarily on beneficiaries in the Sourou Valley and the Comoé Basin. The DA investment totaled $29.7 million and supported efforts to improve farming and livestock productivity, as well as related activities throughout the agricultural value chain.

The DA activity provided training on rain-fed and irrigated production to over 12,000 farmers, about half of whom were women, from 30 villages in the Sourou Valley and Comoé Basin. The training aimed to sustainably increase agricultural productivity and incomes. It covered agricultural techniques and included extension services and demonstration farms. Training focused on corn, cassava, and vegetables in the Sourou Valley and on corn, rice, and onions in the Comoé Basin. The distribution of over 5,000 incentive kits containing agricultural inputs was designed to encourage participation in training activities. Several of the outputs for this sub-activity exceeded expectation, including numbers of demonstration farms created, producers trained, and producers adopting new practices (MCA-BF 2014b).

The DA activity also provided training to over 2,500 livestock holders in animal health and animal husbandry, with a focus on poultry and cattle. More than 1,000 women received support for improved poultry rearing, including material for building chicken coops and improved roosters. As part of the cattle activity, 1565 cows were inseminated with 182 calves born by the end of the compact. The training included model breeding farms, which served as learning centers. It was complemented by a vaccination drive covering over 1.39 million chickens and 1.43 million cows. This sub-activity also provided support to veterinary institutions by training veterinarians, providing equipment and medication, and constructing or rehabilitating veterinary schools and labs. Of the four animal health laboratories that were constructed or rehabilitated, three were located in project areas, while one was in the capital, Ouagadougou. Six rural vaccination parks were also constructed and 60 village volunteers were trained in vaccination techniques for cattle and poultry. As a whole, the animal health sub-activity was intended to sustainably increase livestock productivity by improving animal health services, as well as increasing access to and demand for these services and to livestock medications (MCA-BF 2014b).

The DA activity also included a value-chain development activity, which was meant to increase business linkages, market transactions, and producers' value added (MCA-BF 2014a). It focused on four areas: (1) inputs, (2) post-harvest techniques, (3) value-added techniques, and (4) the creation of links between producers and other market actors. The main activities were

providing training in post-harvest, value-added activities in agriculture and fishery and the establishment of 48 producer associations (MCA-BF 2014a).

Finally, the DA activity included various market-related components, designed to reduce transaction and marketing costs and thus increase agricultural incomes. The rural markets sub-activity—which sought to improve market conditions—funded the establishment and training of nine market committees to manage and maintain their markets; the rehabilitation of four of these markets; and an outreach campaign to provide vendors with information on hygiene, parking, safety, and taxes. Three of the markets that MCC selected for rehabilitation were located in the Sourou Valley and one was in the Comoé Basin.

The final sub-activity under the DA activity was the creation of a market information system, which sought to enable producers to make more informed marketing and production decisions by providing timely information on prices. To receive price information about a product, farmers send an SMS at the nominal rate of 10 CFA and receive a response free of charge. In 2013, the MIS provided price information on 28 agricultural goods (crops and livestock) for 19 regional and provincial markets, including the 9 markets that the ADP supported as part of the rural markets sub-activity. The MIS was transferred to a private operator at the end of the compact.

E. Prioritization of activities

Following the design trip, we learned that sub-activities—in particular those falling under the DA activity—had limited overlap in terms of beneficiaries and geographic coverage. As a result, only limited synergies are available in the evaluation of these sub-activities. Due to a limited evaluation budget, this lack of synergy led MCC and Mathematica—with input from the post-compact entity Agence de Partenariat pour le Développement (APD)—to agree on a prioritization of evaluation activities.

MCC uses the size of its investment in a particular activity or sub-activity and the opportunities for learning as two criteria for deciding on evaluation priorities. To facilitate prioritization, Mathematica assembled information from various project documents on the size of investments by activity and the number of beneficiaries. We then assessed the learning opportunities and incorporated feedback from MCC and APD to categorize activities as low, medium, or high priority for the evaluation. An activity was high priority if MCC's investment in it was a large fraction of overall compact expenditures or the evaluation of this activity provides for rigorous opportunities for learning. An activity was medium priority if it was an important part of overall compact expenditures or if substantial gaps in the literature could be addressed with this evaluation. An activity was low priority if none of these conditions applied. Besides these three main categories, we had two intermediate categories corresponding to medium-high and low-medium priority, which were used when activities did not neatly fit into one of the main categories.

Appendix B describes the available information on the size of the investments and the number of beneficiaries by activity and sub-activity, Mathematica's detailed assessment of learning opportunities of design options, and the discussion on prioritization. Table II.1 summarizes the information on cost, our assessment of opportunities for learning, and the

prioritization of project activities.[5] Based on our conversations with MCC, we focused this report on evaluation designs for activities with high and medium-high priority, though the activities may change as the evaluation evolves.

By far the largest expense of the ADP was the construction of the Di perimeter, which absorbed close to two-thirds of the total ADP expenditure. As a result, evaluation activities related to investigating the consequences of MCC's investments in the Di perimeter have high priority. These activities include updating the ERR analysis of the Di perimeter, assessing the effects on PAPs and the sustainability of irrigation maintenance (through the Sourou O&M activity), and estimating the causal impact of this activity on Di Lottery beneficiaries. Because the farmer training activity had the largest number of beneficiaries within the DA activity it has high priority. The evaluation of the integration of project activities—that is, the extent to which project activities were integrated and, if not, why not—and the evaluation of IWRM were of particular interest to MCC and the APD.

Table II.1. Overview of ADP project activities and summary of prioritization

Sub-activity	Cost (in $)	Summary of Mathematica's assessment of associated opportunities for learning and justification for prioritization	Priority category
Water management and irrigation (104M)			
Di perimeter (including Sourou O&M)	89.0M	There is little literature on the successes and challenges of large-scale irrigation projects in West Africa. The Di perimeter construction received the largest share of investments under the ADP. It provides significant opportunities for learning.	**High**
Di Lottery		The evaluation that assesses the impact of access to irrigated land through the Di Lottery relies on the most rigorous design possible, a randomized control trial (RCT). At a cost in U.S. dollars of about $39,700 (without compact administration costs) or $45,000 (with compact administration costs) per hectare, the benefit of winning the lottery ranks as one of the largest stakes in any RCT that we know of. (Appendix D provides cost information for the perimeter.) The Di Lottery provides significant opportunities for learning, both through the proposed RCT and the proposed methodological study.	**High**

[5] The sub-activities in Table II.1 combine certain evaluation design elements that Appendix D discusses separately when these elements are necessary to provide a rigorous assessment of a sub-activity. In particular, the Di performance evaluation also includes the estimation of agricultural outcomes for non-PAPs from disadvantaged villages, women, and youth as well as the analysis of crop prices. The Di Lottery evaluation also includes an analysis of land productivity outside the perimeter.

Table II.1. *(continued)*

Sub-activity	Cost (in $)	Summary of Mathematica's assessment of associated opportunities for learning and justification for prioritization	Priority category
Sourou O&M (cost also included in Di perimeter)	6.6M	The sustainability of MCC's entire investment in the Di perimeter depends upon whether the irrigation infrastructure is maintained. The Sourou O&M evaluation assesses whether the institutions tasked with the maintenance of the irrigation infrastructure (WUAs and AMVS) will be able to do so. The evaluation of this activity also provides significant opportunities for learning through comparisons with pre-existing perimeters, which are poorly maintained.	**High**
IWRM	5.0M	As of 2012, the African Ministers' Council on Water noted that very few sub-Saharan countries had put IWRM plans into practice. The IWRM evaluation will be informative for IWRM implementation in Burkina Faso as well as for IWRM implementation in other countries.	**Medium-High**
Diversified agriculture (30M)			
Farmer training	NA	Based on available information, the farmer training sub-activity is the DA sub-activity with the largest expenditure share (with the possible exception of the animal health sub-activity). The evaluation can rely on a pre-post design to provide evidence for whether farmers adopted the improved agricultural technologies. The sub-activity was a significant expenditure by MCC and provides significant opportunities for learning.	**High**
MIS	NA	The MIS performance evaluation provides an opportunity to analyze the sustainability of an MIS system in Africa and to understand how easily such a system could be set up to benefit farmers in two project areas within a country. The evaluation will not be able to disentangle MCC's investment from subsequent sources of funding.	**Medium**
Animal health and husbandry	NA	Because the sample of beneficiaries in the existing baseline data is too small for a pre-post analysis and is not representative of all animal health and husbandry beneficiaries, the evaluation is limited to a descriptive analysis on beneficiaries. The opportunities for learning are limited. The evaluation is primarily useful if MCC anticipates implementing animal health and husbandry activities as part of future compacts.	**Medium**
Rural markets	5.6M	The rural markets sub-activity primarily rehabilitated four markets and trained market management committees in the rehabilitated markets and in five additional markets. The proposed performance evaluation will be able to provide learning on the determinants of the functioning of rural markets from only these experiences, of which only four saw construction. The evaluation of the rural markets sub-activity will be primarily useful if MCC anticipates constructing rural markets as a part of future compacts.	**Low**

Table II.1. *(continued)*

Sub-activity	Cost (in $)	Summary of Mathematica's assessment of associated opportunities for learning and justification for prioritization	Priority category
Value-added	NA	Because no baseline data are available, the evaluation of the value-added sub-activity is limited to an implementation analysis as well as a qualitative assessment of its benefits to beneficiaries. As a result, we see little opportunity for learning.	**Low**
Evaluation option for the ADP project as a whole			
Integrated program logic		MCC developed an integrated set of activities for the ADP based on evidence that farmers in the two project areas faced multiple constraints. MCC thought that these constraints needed to be addressed simultaneously to maximize the value of investments. The lessons learned on why the integrated program logic of the ADP was not followed may be applicable to future MCC projects in which project activities are meant to complement each other.	**Medium-High**

Note: The cost for the rural markets activity does not include the training of market committees.

NA = not available.

III. LITERATURE REVIEW

For countries like Burkina Faso, whose population largely consists of subsistence farmers, growth in the agriculture sector is likely to be the most effective means for reducing poverty because poor farmers stand to gain more from growth in agriculture than in other sectors (Christiaensen et al. 2011; De Janvry and Sadoulet 2010). Over the past several decades, agricultural yields in Burkina Faso have been low, growing slowly, and vulnerable to severe weather and drought (Chauvin et al. 2012). Under the ADP, MCC implemented irrigation, IWRM, and farmer training interventions to address the constraints to agricultural productivity and growth in the country. The evaluation of the ADP will contribute to the evidence on the effectiveness of these types of agricultural interventions. To provide context for the evaluation, below we review the existing evidence relevant to these interventions and describe how the ADP evaluation will contribute to this literature.

A. Irrigation and IWRM

The existing literature suggests that irrigation can increase agricultural productivity and income, thus encouraging growth in the sector as well as poverty reduction. In a review of related empirical evidence, Hussain and Hanjra (2004) concluded that irrigation improves agricultural productivity and yields and lowers the risk of crop failure. Similarly, Kuwornu and Owusu (2012) showed that access to irrigation increased crop intensity and improved yields of rice and pepper in Ghana. Janaiah and colleagues (2004) found that irrigation interventions in Vietnam reduced the input costs of agricultural production, increased paddy crop yields from 13 to 22 percent, and had positive impacts for other crops. Similarly, Matsumoto-Izadifar (2009) reported that irrigation from the Senegal River, in combination with appropriate inputs and harvesting techniques, had the potential to increase domestic rice production in the Senegal River Valley.

Such increases in productivity due to irrigation access should lead to greater agriculture income and household consumption and the reduction of poverty. As evidence of this, Tucker and Yirgu (2010) found in their evaluation of the impact of irrigation in Ethiopia that households experienced a 20 percent increase in annual income. Similarly, Datar and Del Carpio (2009) found an increase in annual income of $220 among poor farmers with access to irrigation in Peru. Dillon (2011) reported that access to irrigation in Mali resulted in a 30 percent increase in household consumption. With respect to outcomes of irrigation on poverty, Duflo and Pande (2007) found that the construction of a dam in India was associated with a 0.15 percent decrease in the poverty headcount ratio in downstream districts. In addition, Janaiah and colleagues (2004) found that rehabilitated irrigation infrastructure and improved management of irrigation decreased poverty rates by 12 percent in Vietnam.

However, despite the evidence demonstrating its positive impacts, irrigation is still unavailable in many places where it would have the largest impact. In sub-Saharan Africa, for example, only 4 percent of arable land was irrigated in 2002 (Udry 2010). In Burkina Faso specifically, only about one-third of the 165,000 hectares of land suitable for irrigation is equipped for irrigation (Food and Agriculture Organization [FAO] 2016). Of the 54,270 hectares of land equipped for irrigation in 2011, only 46,130 hectares were in fact irrigated (FAO 2016). As Dillon

(2011) noted, irrigation infrastructure is often not maintained, which could be because a large portion of the irrigation infrastructure has fallen into disrepair.

Also discouraging, the small set of existing literature shows that IWRM interventions have been largely ineffective in promoting growth in the agriculture sector. In a case study conducted in Burkina Faso, Sally and colleagues (2011) argued that the water reform introduced in the late 1990s has had little effect on improving water management, and thus agricultural growth. They concluded that, as of 2011, the CLEs created under the IWRM law that were still operational had limited capacity and were institutionally weak. Rey and colleagues (2008) came to a similar conclusion in their audit of global IWRM implementation: water management institutions may appear operational, but are often too weak to realize IWRM objectives.

Our evaluation of the WMI activity will contribute to the literature on the effects of irrigation and IWRM interventions in developing countries in several ways. First, much of the existing literature of large-scale irrigation schemes primarily draws on the evaluation of projects implemented in Asia; related literature is lacking in particular for West Africa. Second, the randomized controlled trial (RCT) study of the Di Lottery provides a unique opportunity to provide evidence on the causal impact of receiving access to irrigated land on agricultural production, agricultural incomes, and household incomes. To our knowledge, it is also one of the largest-stakes RCTs and the only RCT in which a subset of applicants received irrigated land. Third, the literature on O&M of such schemes is rare and often not rigorous. The evaluation of O&M in Sourou will help fill this gap in evidence, in particular by contrasting the experiences of WUAs on the Di perimeter and those on the old perimeters. Finally, there is little literature on the implementation of IWRM in Africa. The evaluation of the IWRM activity will provide a case study of the successes and challenges of implementing IWRM plans.

B. Farmer training

Agricultural development projects typically include a farmer training component because it is often assumed that the impact of an agricultural intervention depends in part upon the adoption and proper implementation of new production techniques and practices. For example, interventions promoting the transition to high-value agriculture require farmers to learn how to grow new crops. The program logic hypothesizes that these farmer trainings will improve farmer productivity on the irrigated plots, thus increasing program impacts. This section reviews the literature concerning farmer training interventions and the evidence of their impacts.

Existing literature that estimates the impacts of farmer training programs is somewhat limited. Waddington and colleagues (2010) noted in their systematic review that rigorous impact evaluations of agricultural extension interventions were less common due to the difficulty of evaluating such interventions. Most of the studies they analyzed were unable to take advantage of experimental or quasi-experimental designs, or they suffered from inadequate data or selection bias. In another study in 2014, Waddington and colleagues pointed out that farmer training is often offered as one component of a larger agricultural intervention, such as a large-scale irrigation intervention. As a result, it can be difficult to separate the impact of the agricultural extension component from the impacts of the often broader intervention.

Existing rigorous evaluations of agricultural trainings that have been conducted in developing countries have reported mixed results. Just over half of the agricultural extension service interventions reviewed by the Independent Evaluation Group (IEG) of the World Bank in 2011 reported positive impacts on at least one key indicator. For example, in Myanmar, Kabir and Uphoff (2007) found that the majority of farmers in a community were using crop intensification practices three years after just one-third of farmers were trained. There was a diffusion of knowledge such that the trained farmers taught other local farmers the practices. However, some of the studies reporting positive results did not find impacts on all indicators, or the impacts were not evenly distributed across beneficiaries. For example, Benin and colleagues (2008) found that an agricultural extension program in Uganda had a positive impact on crop productivity but not on livestock. The study by Feder and colleagues (2004) evaluating a pest management training in Indonesia found no evidence of increased crop yields or positive economic gains resulting from the training. In Argentina, only previously low-yield and large-scale farmers saw yield increases from grape production extension services, while only larger producers saw improved quality (Cerdán-Infantes et al. 2008). Other studies outside of IEG's systematic review exhibited mixed results as well. For example, Larson and Lilleør (2014) find positive effects of farmer field schools on food security in Tanzania but not on poverty. The evaluations of MCC's first five farmer training activities in Armenia, El Salvador, Ghana, Honduras, and Nicaragua also found mixed evidence of impacts on practice adoption rates and farm income (MCC 2012).

Many of these evaluations highlight that farmer training interventions are unable to impact production and farm income if participants do not adopt the new techniques and practices. Jack (2013) noted that while many farmers have benefitted from the Green Revolution, the adoption of productive agricultural technologies has remained particularly low among the poor. This is particularly true in Africa, even though there is ample evidence suggesting that agricultural technologies with high expected returns do exist in sub-Saharan contexts (Zeitlin et al. 2010). Although farmer training interventions are designed to encourage farmers to adopt advanced technologies, market inefficiencies can constrain the rate of adoption (Jack 2013). For example, costly investments in infrastructure, such as irrigation and roads, may be required to make the adoption of an advanced production technique profitable. Even if such large-scale investments in infrastructure could be made, other constraints may prevent adoption, such as credit or household labor constraints (Jack 2013). Suri (2011), who examined farmers' decisions to adopt hybrid maize in Kenya, also demonstrated that benefits and costs of technologies were heterogeneous across farmers, so not every farmer will view a given technology as beneficial and adopt it.

Low adoption rates may also be due to information failures (Mobarak 2014), such as the perceived lack of credibility of an information source. To investigate the importance of information failures, a number of recent studies have focused on farmers' social networks and how they affect adoption decisions. For example, research in Ethiopia and Malawi suggests that social learning among neighbors has a stronger effect on uptake than extension agents because farmers are more likely to trust the advice and results of farmers similar to them (Krishnan and Patnam 2012; BenYishay and Mobarak 2013). Thus, adoption and dissemination incentives provided to key "contact" farmers in communities is likely an effective way to address low adoption rates due to information failures. Beaman and colleagues argue that targeting driven by social network theory should be used to identify the optimal contact farmers, whose training will have the largest effect on uptake by peers (Kondylis et al. 2014). Alternatively, for farmers

receiving information directly from extension agents, Jones and Kondylis (2016) argue that feedback mechanisms improve extension agency programs and their delivery, which has positive effects on farmers' demand for agricultural extension services—and thus possibly on their adoption rates.

Even if new technologies were adopted, their use may not always be sustained. Instability in the supply of inputs or the risks associated with new technologies may help explain the widespread failure of farmers to persistently adopt profitable technologies. Zeitlin and colleagues (2010) suggested in their study of technology use among Ghanaian cocoa farmers that persistent heterogeneity in returns to technology use may lead farmers to abandon a given technology even though average returns are high.

Our evaluation of the farmer training component of the DA activity will contribute to the literature on farmer training in several ways. First, our study will contribute to the existing research on the effects of training on practice adoption and the sustainability of this adoption for the specific innovations introduced by the compact. Via qualitative research, we will also assess if adopted practices are being implemented as intended or adapted to local contexts—two concepts not well covered in the literature. Finally, we will investigate the effects of distributing incentive kits over multiple growing seasons, an area that is not addressed in the literature.

C. Land distribution and titling

The ADP changed land rights for beneficiaries in the Di perimeter in two ways: (1) PAPs received land titles for the irrigated land they got as compensation, and (2) the ADP provided non-PAPs with rights to new land in the form of leases in addition to written leasehold contracts. To provide background on how these two changes might affect beneficiary outcomes, we review the existing evidence on the effects of land titling and land provision.

1. Impacts of the provision of land titles

Research indicates that land titling programs can prevent the negative outcomes of land insecurity. Studies typically show that land insecurity and the absence of titles can have sizeable costs. One example comes from a study of data collected in 2010 as part of MCC's independent evaluation of the Burkina Faso Compact's Rural Land Governance Project. In this study, Linkow (2016) found a potential for costly land conflicts related to both migrants and to former residents of the village returning to claim land. High levels of perceived concern over both types of conflict were associated with reductions of over 40 percent in agricultural productivity; the overall productivity impact of land tenure insecurity in the study area was 8.9 percent. Land tenure insecurity is higher for women with negative outcomes for land productivity. Goldstein and Udry (2008) showed that the lower land security of women leads to less investment on the plots controlled by women, reducing agricultural outcomes. In contrast to much of the literature on land tenure security and investment, however, Brasselle et al. (2002) found in Burkina Faso

that the traditional village order provides the basic land rights required to stimulate small-scale investment.[6]

Several reviews of the literature have shown that land titling programs can have positive impacts on tenure security and land investments, but these impacts can vary substantially depending on the features of the program and the local context (Deininger and Feder 2009; Payne et al. 2009; Besley and Ghatak 2010). For example, an impact evaluation[7] of a land regularization program in Rwanda showed that it led to a 10 percentage point increase in the use of soil conservation activities after 2.5 years (Ali et al. 2014). Deininger and Feder (2009) note, however, that the evidence for positive impacts is not uniform but rather depends on the governance environment, the effectiveness of the state apparatus, and the distribution of socioeconomic power.

2. Impacts of land provision

Ghatak and Roy (2007) and Bardhan and Mookherjee (2007) reviewed the literature on land provision, generally finding positive effects of land provision on agricultural productivity and poverty reduction. However, Deininger et al. (2008) suggest that the features of the land provision matter: provision of land without formal titles is less effective than provision with formal titles.

3. Impacts by gender

The literature highlights substantial differences in (1) land tenure security by gender and (2) the impact of land certification and land provision by gender.

A number of researchers have investigated land rights by gender in West Africa and found that these rights vary by gender and status as head of household, with important consequences for agricultural inputs, land investments, and outcomes. Using data from Burkina Faso, Udry (1996) found that plots controlled by women have significantly lower yields than similar plots within the household planted with the same crop in the same year—but controlled by men. The yield differential is attributable to significantly higher labor and fertilizer inputs per acre on plots controlled by men.[8]

Goldstein and Udry (2008) studied the difference in profits between husbands and wives on very similar plots by examining the role of ambiguous and contested land rights on investment and agricultural productivity in Ghana. They found that the entire difference between profits on husbands' and wives' plots is attributable to the longer fallow periods on men's plots, which lead to substantial increases in yields. They argue that women do not let their land be fallow as their

[6] Braselle and colleagues (2002) defined the traditional order as the persistence of indigenous practices and customs, and they captured the strength of the traditional order through a categorical variable based on an internally consistent hierarchy of rights.

[7] This evaluation was based on a geographic discontinuity design.

[8] Using an alternative data set, Akresh (2008) corroborated the finding that there are differences in outcomes across genders but notes that these differences are not ubiquitous. Specifically, Akresh found that negative rainfall shocks are correlated with increases in labor resources allocated to the wife's plots, showing that in bad years, households try to avoid losses by reallocating labor.

rights to the land are less secure, and land that is not planted and harvested is more easily claimed by other community members and relatives.

In a more recent paper in Burkina Faso, Kazianga and Wahhaj (2013) argue that it is not the gender of the person but his or her role in the household that drives these differences. They show that plots owned by the household head (who is usually male) are farmed more intensively and achieve higher yields than plots with similar characteristics controlled by other household members. However, there are no differences between plots controlled by males who are not the head of the household and female household members.

A few studies have focused on the impact of land regularization and titling on female agency and intra-household decision making. Ali et al. (2014) found that land regularization and titling improved land access for married women and reduced gender bias in the recording of inheritance rights in Ghana. Regularization and titling also increased the likelihood of a wife possessing land in her name and deciding which crops to grow on her land in Ethiopia (The Cloudburst Group 2016). In addition, regularization and titling increased participation in household decision making and reduced fertility after one to two years when females were included in the land title (Field 2003); it also decreased female time spent on chores and spending on male-favored items in China (Wang 2014).

Even when titles to land are available, women are typically at a disadvantage in having their land rights recorded, though research shows that achieving more equitable outcomes in land tenure is possible in some contexts (Ayalew et al. 2014). Using a randomized field experiment that provided price incentives for land tenure formalization in urban Tanzania, Ayalew and colleagues (2014) found that even small price incentives for female co-titling result in almost complete gender parity in land ownership, with no reduction in demand.

Overall, the literature on land provision has several important gaps. First, there are relatively few rigorous impact evaluations on this subject because land provision is rare, and instances where it does occur typically do not allow for rigorous evaluation. Second, few studies have been able to estimate accurate impacts of providing land by gender. Third, few studies document the interactions between the provision of new land and informal needs-based land rights in Africa.

Our evaluation will fill these gaps by providing a rigorous impact evaluation of land provision. It will enable us to estimate the effects of providing land for females versus males because, in our case, gender was explicitly incorporated into the land lottery. We will also be able to provide a detailed descriptive analysis of the usage of land titles.

IV. EVALUATION DESIGN

The evaluation of the ADP activities will address a wide range of questions related to improvements in water use and availability, agricultural production and technology, income, and land use. This chapter describes our proposed design for evaluating the priority ADP activities in the Boucle du Mouhoun and Cascades region. We begin by discussing the two types of evaluations that we will conduct—performance and impact evaluation—before presenting our evaluation design for each priority activity. We then describe the data sources, the data collection method, and the time frame for data collection. The chapter concludes with a discussion of the risks and challenges that these evaluations might face and our plans to address them.

A. Evaluation strategy

1. Types of evaluations: performance and impact

This section describes the methodological differences between performance evaluations and impact evaluations, both of which we will use to assess the ADP. Both are intended to be rigorous, relevant, and feasible, and in both cases, collaboration with key stakeholders is critical for the success of the evaluation.

a. Performance evaluations

A performance evaluation allows us to assess whether and the extent to which a program produces its expected output or outcome. For example, in our evaluation, we could measure whether program beneficiaries received all the land tenure instruments they were supposed to receive.

A performance evaluation provides information on the structure and implementation of a program, on key stakeholders' views of implementation and outcomes, on different actors' views of a program's outcomes, and on the interpretation of quantitative results. The questions that performance evaluations typically answer are: (1) what is the program structure and why is it structured this way? (2) how was the program implemented, and if implementation differed from the design, then why? and (3) are the program and its impacts sustainable. To understand the program structure, we collect information from different actors on funding, program logic, and expected impacts. To understand how the program was implemented, we compare final and expected results in order to discern the gaps between the program as planned and the program as implemented. This understanding forms the basis for the lessons we can learn for future implementation. To determine whether the program and its impacts are sustainable, we collect information from stakeholders on whether program activities or investments endure and on whether there is a plan for continued operation after donors no longer fund the program.

A performance evaluation relies on various data sources, including in-depth interviews and focus group discussions with key stakeholders, site visits, and document review. Key informant interviews with people who are not implementing a program but who understand the context in which a program operates provide complementary perspectives. The information from different sources is triangulated to test the strength of, and for inconsistencies in, the findings in order to draw conclusions about the program implementation that are fair and complete. Performance evaluations may also rely on quantitative data, such as administrative data, to document and/or

track program implementation. They may also include a pre-post analysis of important outcomes to measure changes over time, though these changes cannot be attributed to the specific program because of the absence of a valid counterfactual.

b. Impact evaluations

An impact evaluation allows us to measure the variation in the outcomes for a particular group that can, with a high level of confidence, be attributed to a specific program. For example, in our impact evaluation, we would be able to measure whether agricultural production and household income increased, and if this is the case, we could state that this increase a result of the program.

The ability to attribute changes to a given program is what separates an impact evaluation from a performance evaluation. To attribute changes, an impact evaluation addresses confounding factors. The most common and reliable strategy for doing so is to construct a credible control group.

An impact evaluation compares outcomes for a treatment group (the group that received the intervention) with outcomes for a control group (a group that is very similar to the treatment group but that did not receive the intervention). The control group represents the counterfactual in that it mimics what would have happened to the treatment group had it not received the treatment. Moreover, unlike a pre-post design, an impact evaluation does not rely only on the comparison of outcomes before the intervention and after the intervention for the group that received the intervention because the variable being studied could have changed over time even in the absence of the intervention. For example, agricultural production could have increased or decreased because of weather patterns even in the absence of the program.

Impact evaluations can address research questions for which it is possible to construct a credible control group. The questions that impact evaluations typically answer are (1) what is the impact of a program or policy on outputs, (2) what is the impact of a program or policy on short-, medium- and long-term outcomes, (3) are there differences in impacts for different groups of beneficiaries. Unless it is specifically designed to investigate causal mechanisms, an impact evaluation cannot typically answer the question of which causal mechanism might have led to the observed impact.

In many situations RCTs are the best option for conducting impact evaluations because the control group is expected to have the same outcomes as the treatment group in the absence of the treatment. In many other situations, however, RCTs are not feasible—either for practical, political, or ethical reasons. Consequently alternative designs are often considered. In order to help provide evidence on the efficacy of alternative designs we propose to conduct a "within-study comparison" which involves comparing results based on three alternative designs with those from the Di Lottery RCT. The alternative designs include a regression discontinuity design (RD), often considered the best non-RCT method, as well as two variations of RD that try to balance the need for rigor obtained using the RD method with the need to generalize results to larger fractions of the populations being studied than is possible using RD. These results will help inform future work based on these types of alternative designs.

An impact evaluation relies on quantitative data to measure outcomes and outputs in order to assess differences in outcomes and in outputs between the treatment and control group. The analysis of complementary qualitative data can help to interpret quantitative results from impact evaluations by providing information on the potential mechanisms.

2. Overview of the evaluation strategy

Our goal is to develop the most rigorous and feasible evaluation design that answers the research questions of interest to MCC. In total, Mathematica will conduct six evaluations: (1) the evaluation of the integration of ADP activities, (2) the Di perimeter ERR and Di PAP evaluation, (3) Di Lottery RCT, (4) the Sourou O&M evaluation, (5) the IWRM evaluation, (6) and the farmer training evaluation.

The Di Lottery evaluation will consist of an impact evaluation in which we will compare outcomes for the treatment group (lottery winners) with outcomes for the control group (eligible candidates who did not obtain a plot of land through the lottery). The remaining evaluations will be performance evaluations that will include document review, interviews, focus groups, and, when possible, pre-post analysis. Our data collection will strive to ensure representation of women in our qualitative and quantitative samples, and we will disaggregate the analysis of beneficiary outcomes and perceptions where possible. The following sections describe our evaluation design for each activity.

B. Integration of project activities conducted under the ADP

The various components of the ADP were designed to work together to address problems that farmers in Burkina Faso face along the value chain. However, according to staff from MCC, the post-compact entity APD and other implementing partners, the project was not implemented according to the integrated program logic, and there was limited overlap in the beneficiaries of different project activities.

The goal of the evaluation of the integration of project activities is to assess the extent to which these activities were integrated, and when they were not integrated, to determine why the integrated program logic was not followed during implementation. The evaluation is intended to answer the following research questions:

1. To what extent were the various project components implemented in a cohesive way, i.e., in which the components complemented each other, as anticipated in the original program logic?

The ADP was conceived as an integrated set of project activities, listed below, that are designed to address most of the challenges farmers in the project areas face along the value chain.

- Inputs: provide irrigated land, access to appropriate fertilizer (organic and inorganic), and new crop varieties
- Planting: provide training in agricultural technologies for irrigated land and prepare the land to maximize yield, including for new crop varieties
- Harvesting: provide training in harvesting practices that minimize damage to the crop
- Post-harvest processing: teach techniques to reduce losses post-harvest
- Marketing: link farmers with new markets, establish relationships between farmers/producer associations and traders

a. Are the rural markets and the MIS functioning and being used by farmers who benefitted from technical assistance or received land in the Di perimeter, as anticipated by the integrated program logic?[9]

2. If the program was not implemented as a cohesive project, according to the original logic, then why not?

To answer these questions, we will conduct a performance evaluation that will draw on qualitative and administrative data sources. In Chapter V, Table V.1 summarizes the proposed data sources, samples, data collection methods, and the areas on which the evaluation will focus. In Appendix C, Table C.1 lists the research questions for all six evaluations and the methods and type of data we will use to answer them.

To evaluate the extent to which the project components were implemented in a cohesive way (RQ1), we will review project records that document implementation. These include the strategic plan, project records such as reports from the implementers, and administrative data such as the indicator tracking table that was collected as part of the compact M&E activity. To provide a quantitative sense of overlap, we will provide a descriptive analysis of participation in ADP activities among the sample of farmer training beneficiaries who will be interviewed as part of the farmer training evaluation (see Section G of this Chapter). To address whether the MIS and rural markets are functioning (RQ1a), we will conduct site visits to the markets and test the MIS system by submitting price queries. To assess whether farmer training beneficiaries and Di beneficiaries are using the markets and MIS, we will conduct descriptive analyses of data on usage collected as part of the farmer training and Di PAP survey.

A desk review of these documents will provide evidence of the planned and actual implementation of various activities. The review will, among other things, identify *the ways in which* the program was not implemented cohesively. To evaluate *why* the program was not implemented in a cohesive way (RQ2), we will conduct in-depth interviews with people who were involved in the implementation. The interviews will focus on such potential reasons as unforeseen constraints, incorrect assumptions about the ease or utility of integration in the program logic, or unintended consequences during program implementation. The informants will include current and former staff from MCA and its successor organization, APD, along with representatives from the implementing agencies.[10]

[9] This research question also provides a limited evaluation of the rural markets and MIS activity. This evaluation could expand in depth if there are sufficient funds.

[10] We are mindful of the evaluation of the ARF activity conducted by another evaluator that addressed questions related to the lack of integration of the rural finance component with the WMI and DA activities (A2F 2015). The evaluation already identified a number of issues with overall compact management and the impact of staff turnover on the integration of ARF with the other ADP activities. We do not intend to duplicate these efforts but to focus more narrowly on the integrated program logic of the DA and WMI activities to identify the extent to which the components of these activities were integrated (or not), and if not, then why.

C. The Di performance evaluation

The objectives of the Di performance evaluation are to (1) provide an economic assessment of the value of MCC's investment in the Di perimeter and (2) study the effects of the displacement and compensation on PAP households' economic well-being, agricultural production, agricultural productivity, and land tenure security.

We will address the following key research questions:

1. How were the Di perimeter construction and associated activities (see text box on the right) implemented relative to the original plans?
2. What is the total area planted, average yield/hectare, total production and total profit on the Di perimeter for each of the focus crops: rice, corn, onions, tomatoes, soybeans, and cowpeas?
   a. Have prices for these crops changed since the completion of the perimeter?
   b. Are agricultural outcomes different for Di Lottery beneficiaries and Di PAPs? If so, why?
3. What is the economic rate of return of the Di perimeter?
4. How has PAP well-being changed? Have any PAPs been harmed (socially, economically, or politically) by the intervention? How?
5. Have PAPs received the compensation instruments (titles and/or leases and/or financial compensation) they were informed they would receive? Why or why not?
6. What are the PAPs' perceptions of the process by which compensation was determined and provided? What are the PAPs' perceptions of the compensation provided?
7. How has the PAPs' perception of land tenure security changed?
   a. Have any PAPs been involved in a land conflict on the perimeter?
8. What type of land investments do PAPs' make? Have PAPs rented or sold land from the Di perimeter? Have PAPs used land from the Di perimeter as collateral for credit?

The Di perimeter construction comprised the following project activities:

- Constructing of a perimeter with 2,240 hectares of irrigated land
- Distributing land to the following beneficiary groups:
  - PAPs
  - Non-PAPs from disadvantaged villages
  - Di Lottery beneficiaries
  - Women
  - Youth
- Providing formal titles to full ownership to PAPs for land received in compensation; providing formal leases to PAPs and other beneficiaries for noncompensation related land.
- Providing financial compensation to PAPs for lost harvest during the construction of the perimeter
- Providing both training in agricultural technologies for irrigated land and starter kits (land preparation and inputs) during first growing seasons for beneficiaries
- Setting up water user associations and CATG, and reforming AMVS (see O&M evaluation)

Below we discuss the previous evaluation designs and our proposed evaluation design.

1. Previous evaluation designs

The initial evaluation design focused on assessing the consequences of the construction of the perimeter on PAPs. It consisted of a quantitative performance evaluation based on a pre-post methodology (IRIS 2010). This design requires data on outcomes for the same households before

the implementation of project activities and data on the same outcomes after the implementation of project activities. This is not a feasible design for key outcomes such as agricultural practices, agricultural income, and overall household income because of the absence of relevant baseline data. In particular, the data collected as part of the compensation process did not capture information on the value of agricultural output before the relocation. The data that are available are (1) the area of plots lost by a household because of the construction of the perimeter, (2) whether the plot was irrigated, and (3) the crop grown during the rainy season.

A separate baseline survey did not provide information on land use within the perimeter before resettlement. It included questions only about crop production and agricultural practices on land outside of the perimeter. Because of these limitations, the second evaluator concluded that the surveys could not be used for a pre-post evaluation (IMPAQ 2014a). Instead, the second evaluator proposed that the surveys from the follow-up data collection would include questions about how the PAPs perceive their current situation compared with their situation before resettlement (IMPAQ 2014a).

We agree with the previous evaluators that the data issues in the compensation and baseline surveys preclude a true pre-post analysis for individual households. However, because the change in agricultural production, incomes, and profits from access to irrigated land is so large relative to the level of these agricultural outcomes before the construction of the perimeter, it is possible to gain an understanding of the magnitude of these changes through a mixed method study.

2. Proposed evaluation design

To answer the research questions for the Di performance evaluation, we propose a mixed-methods study that will draw on quantitative data collection, qualitative data from interviews with beneficiaries, and project documentation. In the following we describe our approach to (1) assessing the implementation of the Di perimeter construction and complementary activities (RQ1), (2) investigating perimeter-wide outcomes (RQ 2,3), and (3) evaluating the effects of the Di perimeter on PAP households (RQ 4,5,6,7,8).

a. Implementation study

Our implementation study of the Di perimeter (RQ1) covers the construction of the perimeter, the resettlement of PAPs, and the attribution of land to non-PAPs, O&M, and Di beneficiary training activities. To answer RQ1—how activities were implemented relative to the original plans—we will compare planning documentation with implementation documentation. The planning documents we intend to use are, for example, the compact, the investment memo, and documents related to the re-scoping of the perimeter. The implementation documents that we plan to use are the implementing consultants' final reports, the MCA achievement report, and the indicator tracking table. We will contrast the scope (and timing where applicable) of the planned activities with the actual implementation to determine the extent to which the project as implemented deviated from the original plans.

b. Study of perimeter wide outcomes

In order to describe agricultural outcomes on the Di perimeter (RQ2) and to recalculate the ERR (RQ3), we propose to collect and analyze quantitative data on agricultural production,

incomes, and profits on the Di perimeter. To be able to estimate total profits for the entire Di perimeter, we will draw a representative sample from the plots assigned to each of the beneficiary categories that benefited from land on the Di perimeter, with the exception of Di Lottery beneficiaries.[11] Because these beneficiaries are the focus of the Di Lottery RCT, which is described in Section D, we will survey this group in its entirety. Because beneficiaries from different beneficiary categories will be selected using different sampling probabilities, we will apply sampling weights to obtain the estimated total production.

Figure A.2 in Appendix A shows a map of the allocation of land to the different groups of beneficiaries and gives a sense of the amount of land each group received. Table IV.1 provides an overview of the number of beneficiaries and the sample size by beneficiary group. The second and third columns list the number of beneficiaries in a group and the total amount of land owned, respectively. The fourth and fifth columns show the proposed sample sizes for the household surveys and crop cuttings, respectively.[12] The final column indicates whether we will draw the sample by simple random sampling or stratify on respondent characteristics.

Table IV.1. Quantitative data collection

Beneficiary category	Number of beneficiaries	Hectares owned in Di perimeter	Survey sample size	Crop-cutting surveys	Sampling strata
PAP	846	1,099	275	110	Gender, plot acreage, plot type
Non-PAP from disadvantaged village	461	317	79	32	Gender, plot acreage, plot type
Di Lottery beneficiary	503	710	503	71	None
Women	1725	90	30	20	None
Youth	846	16	30	20	None
Other: tree nursery, National research institute (INERA), Mixed groups	17*	8	0	0	
All Di beneficiaries	**4398**	**2,240**	**917**	**253**	

Note: Information on the number of beneficiaries and the hectares owned in Di comes from the land allocation spreadsheet (MCC 2016a). Sample size is determined through optimal stratified sampling. This specifies that the optimal sample size for strata is proportional to the number of observations within strata multiplied by the relative standard deviation. We assumed that the standard deviation for a farmer's total production is proportional to the acreage of the plot(s) owned in the Di perimeter. As a result, the optimal sample size for a strata is proportional to the area of land covered by that strata. We made two exceptions: (1) we will survey all Di Lottery beneficiaries for the Di Lottery study, and (2) we will set a minimum number of surveys in a strata of 30 observations for the survey and 20 for crop cuttings to be able to provide some descriptive information on strata.

[11] We will not retain the sample of Di PAP households from the baseline survey because baseline respondents are not a representative sample of Di PAP households and therefore not of plots owned by Di PAP households because of attrition of 22.4 percent.

[12] For the PAP households and the Di Lottery beneficiaries, the number of surveys may be higher as we will interview both the initial owner of the land (to understand the consequences of the construction of the Di perimeter) and the farmer who currently cultivates the land (to assess agricultural outcomes in the Di perimeter).

To update the ERR (RQ 3), we will use the information collected to calculate agricultural profits, that is, the area planted, yields, total production, input quantities and prices, WUA payments,[13] post-harvest losses, and production sales and prices.[14] We describe the ERR and our plans to recalculate the ERR in Appendix D.

To understand whether the additional production on the Di perimeter has led to lower crop prices (RQ2a), we will analyze information on market prices collected by SIMCA, the market information system supported by MCC.[15] We will compare the prices of perishable and nonperishable goods over the course of the agricultural season at markets surrounding Di relative to other markets before and after crop production started on the perimeter.[16]

During Mathematica's site visit to Burkina Faso, stakeholders mentioned differences in agricultural outcomes (area planted, yields) and WUA payment rates between Di Lottery beneficiaries and Di PAP households, although they did not provide evidence of these differences (RQ2b). To describe the differences in agricultural productivity (area planted, yields) and in WUA payment rates between Di Lottery beneficiaries and Di PAP households, we will analyze the information collected to calculate agricultural profits separately for Di Lottery beneficiaries and Di PAP households. To understand the reasons behind these differences, if they exist, we will conduct key informant interviews with current staff at AMVS, CATG, the Ministère de l'Agriculture et des Aménagements Hydrauliques (Ministry of Agriculture and Hydraulic Installations), as well as in-depth interviews with presidents of WUAs and board members of WUAs.[17]

[13] If available, we will assess whether self-reported WUA payments are reliable by cross-checking with WUA payment records.

[14] We will collect information on agricultural outcomes even when a plot is left fallow so as not to overestimate production and yields.

[15] Most of the literature on market integration in Africa shows that markets for cereals are integrated (Dillon and Dambro 2016). As a result, the additional production on Di should only have minimal influence on prices. Two notable exceptions to this view are Aker (2010) and Essam (2013). Aker found that the degree of integration in Niger was influenced by a drought, suggesting that markets in Niger are not perfectly integrated. Essam (2013) found supporting evidence for this view, also in Niger. He combined remote-sensing information on millet production with market prices to show that prices respond to local production and that markets are better integrated in years that have negative production shocks. There is much less evidence on whether markets for perishable goods, such as vegetables, are integrated. One of the few studies on this subject is by Ddungu et al. (2015), who found that cowpea markets in Uganda are not integrated. As such, both significant price declines (especially for vegetables) and relatively stable prices are plausible.

[16] This question is not linked to the program logic. It is a potential unintended consequence of the construction of the Di perimeter that was mentioned both in the due diligence report for Di (MCC 2008a) and during Mathematica's site visit to Burkina Faso in October 2016.

[17] These differences in agricultural outcomes may be due to many factors, including differences between PAPs and Di Lottery beneficiaries in the quality of land they received, the size of their plots, land-tenure security, the quality of the training received, demographic characteristics, the lease payments that Di Lottery beneficiaries need to pay, and experience in irrigated agriculture. We will not be able to quantify the relative contribution of these agricultural differences to the potentially observed differences between PAPs and Di Lottery beneficiaries.

c. Study of the effects of the Di perimeter on PAPs

To understand whether PAP well-being has changed (RQ4), we will provide a descriptive analysis of self-reported assessments of changes in well-being collected as part of the quantitative survey. To understand how PAPs might have been harmed, we will speak with implementers, WUA presidents, WUA board members, and PAPs during the key informant interviews and focus group discussions.

During the site visit, several stakeholders noted that production in the rainy and dry seasons serve different purposes. Production in the rainy season—typically corn or rice—is used for subsistence purposes. Production in the dry season—onions and tomatoes—is mostly destined to be sold. The small number of Di PAPs we talked to were unanimous in saying that food security had risen, but it was less clear to us if this was also true for income resulting from the dry season harvests. Since food security and income are two dimensions of well-being, we will inquire about both during the interviews.

Our evaluation into PAP well-being will also investigate whether perceptions of changes in well-being vary by gender. To do this, we will select women for the in-depth interviews with PAPs and members of PAP households. We will also conduct focus group discussion separately with women beneficiaries.

To answer the question on whether PAPs have received all compensation documents (RQ5), we will first review post-compact progress reports from the APD. We will triangulate this information using self-reports by PAPs as part of the quantitative survey. If PAPs have not received compensation instruments, this issue will be addressed through the in-depth interviews with people who were involved in the implementation during the compact and with people who were tasked with the delivery of compensation instruments post-compact. These informants include the current and former staff from MCA and its successor organization APD, representatives from implementing agencies, as well as staff from the Di town hall, who are now responsible for providing titles to land in the Di perimeter.

The data we collect in the in-depth interviews and focus group discussions with Di PAPs and WUA board members will allow us to address the research questions related to the perceptions of the compensation, of the process of compensation, and of land security (RQ6, RQ7).

Finally, to investigate the research questions related to land conflict (RQ7b) and to land investment, land markets, and credit markets (RQ8), we will conduct a descriptive analysis of responses by PAPs to questions in a specific Di PAP land module that will be part of the quantitative questionnaire. The questions related to land rental and sales will also shed light on whether an active land rental and sales market has emerged.

D. The Di Lottery RCT

The objective of the Di Lottery RCT is to provide rigorous evidence of the impact of receiving access to irrigated land—in combination with training in irrigated farming technologies and start-up materials—on Di Lottery beneficiaries and their households. Key research questions for this evaluation are:[18]

The 503 Di Lottery beneficiaries received:

- Leases to land on the Di perimeter
- Training in agricultural technologies for irrigated land
- Starter kits (land preparation, materials and inputs)

The 503 Di Lottery beneficiaries were selected from among 1,528 participants and received leases for 710 hectares of land. The 1,025 lottery participants who did not win constitute the control group.

1. To what extent did Di Lottery beneficiaries receive all benefits they were meant to receive (formal lease documents, training in agricultural technologies, starter kits)?
2. What impact does winning the Di Lottery have on agricultural practices, production, total agricultural income, and overall household income of the Di Lottery beneficiaries?
3. What are the impacts of winning the Di Lottery on land tenure security?
   a. Have Di Lottery beneficiaries been involved in a land conflict on or off the perimeter?

In addition to evaluating the impact of the lottery, we propose a methodological study—called a within-study comparison (WSC)—that compares the estimated impacts of the Di Lottery RCT with the impacts estimated through a regression discontinuity design (RD). In addition, we will investigate the performance of two recently developed RD-related methods that estimate impacts for beneficiaries who are away from the discontinuity (Wing and Cook 2013, Angrist and Rokkanen 2015). Evidence on the performance of these two methods is particularly limited, and the WSC will answer the following research questions:

4. To what extent are the estimated impacts from the RD similar to those from the RCT?
5. To what extent can methods that use the discontinuity to estimate impacts away from the threshold recover the average treatment effect of the Di Lottery?

Below we describe the Di Lottery and the beneficiary selection process, discuss the previous evaluation designs, assess of the previous evaluation, and describe our proposed evaluation design. We also outline the proposed WSC.

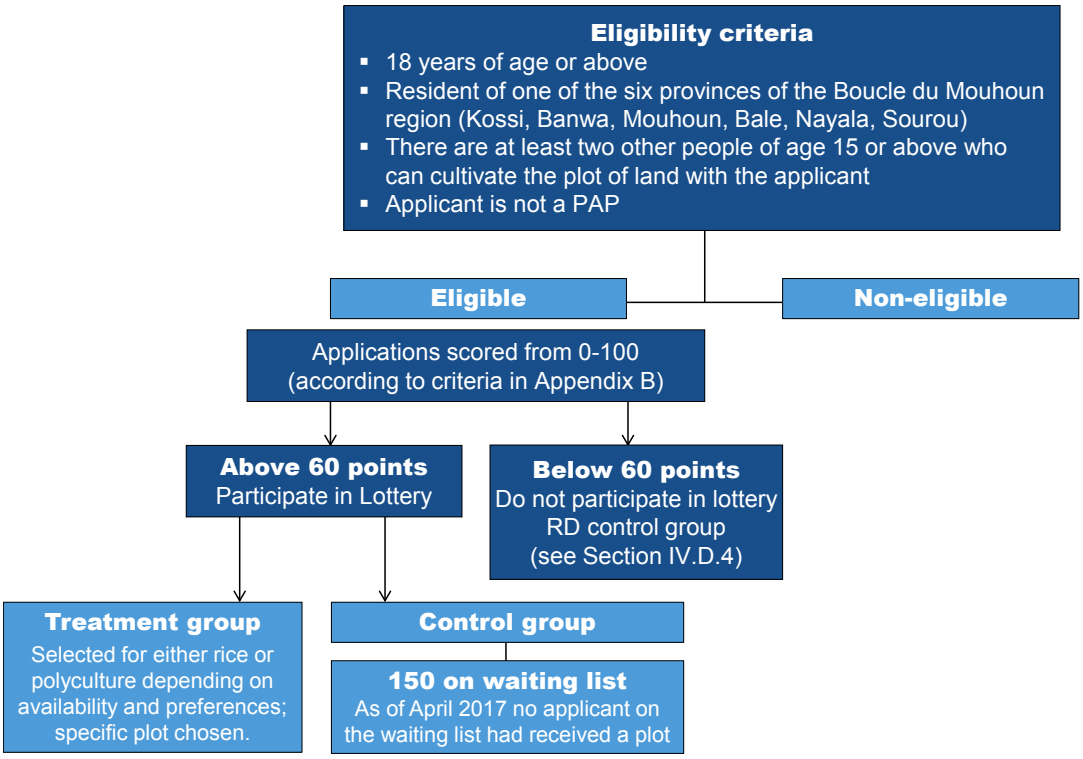
1. The Di Lottery beneficiary selection process

Recruiting applicants and selecting Di Lottery beneficiaries was a multi-step process (see Figure IV.1). The lottery was announced over the radio, with posters and through postings in local administrations (town halls). The ADP invited individuals to apply for the lottery who were (1) not PAPs; (2) 18 years of age or older; (3) residents of one of the six provinces of the Boucle du Mouhoun region (Kossi, Banwa, Mouhoun, Balé, Nayala, Sourou); and (4) able to list at least

[18] The Di Lottery survey for the interim or final data collection will also include questions to assess individual-level outcomes by gender—for example, expenditures in certain expenditure categories, control over resources, and male and female education.

two other people age 15 or older who could cultivate the plot of land with the applicant. The MCA-BF contractor SHER-GRET (also known under its contract number AD7) initially determined that 2,178 of the applicants met all four criteria. MCA-BF, however, allowed applicants to contest the decision and ask for a review. As a result, an additional 51 applicants were deemed eligible, bringing the total to 2,229 eligible applicants (IMPAQ 2014).

Figure IV.1. Di Lottery beneficiary selection process



Among eligible applicants, admission to the lottery was determined through a points-based system developed by MCC and MCA-BF, and designed to (1) select applicants with higher expected benefits (for example, applicants received more points when they owned certain machinery); and (2) meet distributional objectives (for example, female applicants and younger applicants received additional points). Appendix E includes the scoring sheet that was used to score applicants.

Applicants provided the following information as part of their application package:

- Number of household members aged 15 and above who would be available to could help with work on the land
- Applicant's experience with irrigation
- Any ownership of irrigated land in other AMVS perimeters
- Participation in MCC-sponsored training activity

- Age and gender of the applicant
- Type of agricultural equipment owned by the applicant
- Level of debt
- Location of residence

Applicants were aware of the four eligibility criteria, the information they were scored on, and the associated scores, but they were not aware of the exact threshold that would determine participation in the lottery. In order to make the selection process transparent, all application information was made public in multiple locations (for example, at local town halls). The accuracy of the application documents was verified by, for example, cross-checking debt with farmers' cooperatives and land ownership with water-user associations.[19]

The Commission pour l'Attribution de la Terre, in collaboration with MCA-BF and MCC, set the cut-off for participation in the lottery at 60 points. Given the number of male and female applicants scoring 60 points or more, this cutoff made it highly probable that at least 20 percent of beneficiaries would be females. Applicants with a score of 60 and above participated in the public lottery, which was held in February 2014. The lottery was conducted by the national lottery company LONAB (Loterie Nationale Burkinabè) and validated by a cabinet of lawyers.

Because there were two standard types of plots for which leases were distributed, applicants had to rank the two types of plots as their first and second choices. One plot was two hectares of land suitable for cultivating rice, and the other was one hectare that was suitable for polyculture.[20] Just under 10 percent of the applicants stated that they would not accept a second choice.[21] Table IV.2 shows the frequency of the applicants' various preferences.[22]

[19] AD7 verified this information.

[20] Most rice plots are two hectares, and most polyculture plots are one hectare. However, the sectors in the perimeter are not exactly rectangular. As a result, 23 percent of rice plots are a little bit smaller or larger than two hectares, with a range from 1.5 to 2.375 hectares of land. The sizes of polyculture plots range from 0.75 to 1.25 hectares, though only 4 percent of polyculture plots are larger or smaller than one hectare.

[21] Thirty applicants did not state whether they would accept the second choice. We include them here as if they had refused their second choice because that is how the lottery would have treated them. We are unsure whether applicants knew that if they failed to say whether they would accept a second choice, this would be interpreted as a refusal. Twenty-four of these applicants listed polyculture as their first choice, and six listed a rice plot as first choice. According to MCA-BF, the latter applicants would have been treated as if they had refused the alternative option, though the case was never relevant.

[22] Based on information we received from APD, applicants who did not state their preferences were assigned to polyculture plots only, so in the analysis, we treated them as if they had stated polyculture as first preference and gave no second preference.

Table IV.2. Preferences for plot types in the Di Lottery sample

	Number	Percent
First choice polyculture plot; accept rice plot	1,139	74.5%
First choice rice plot; accept polyculture plot	213	13.9%
Accept polyculture plot only	123	8.0%
Accept rice plot only	7	0.5%
First choice polyculture; not clear if accept rice plot	24	1.6%
First choice rice plot; not clear if accept polyculture plot	6	0.4%
No ranking over plot types	16	1.0%
	1,528	**100.0%**

To conduct the lottery, LONAB used a standard tombola set-up: each lottery participant was assigned a number, and balls with these numbers were put into a large container. In addition, two additional tombolas contained numbered balls identifying the available rice and polyculture plots. We call these the rice and polyculture tombolas.

In order to choose lottery beneficiaries, the following process was repeated until there were no plots remaining. To start the process, a volunteer from the public would draw the name of a lottery participant from the main tombola. If the selected participant's first choice of plot was available, a plot of this type was selected for the participant from the rice or polyculture tombola, whichever was the participant's first choice. If the selected participant's first choice was not available, but he or she had indicated a second choice, a volunteer drew an available plot from the tombola containing plots that were the participant's second choice.

If the first choice was not available, and the applicant had not indicated a second choice, the person did not receive a plot of land, and he or she was instead on the waiting list.[23] In this case, the selection process started from the beginning. Of 1,528 participants, 508 beneficiaries were selected.

After all plots had been assigned, 150 additional names were drawn to constitute a waiting list. If an applicant did not start working on the plot he or she received, applicants from the waiting list were meant to benefit from the unoccupied land. At the time of this writing, however, no applicants from the waiting list have benefited from a vacant plot; because of the political transition in Burkina Faso, the institution that re-assigns plots was only created in 2016 and has yet to re-assign plots.

2. Previous evaluation

Both previous evaluators proposed to evaluate the impact of the Di Lottery on the Di Lottery beneficiaries by using the framework of an RCT (IRIS 2010, IMPAQ 2014). This methodology involves comparing the outcomes of the lottery participants who were randomly selected to

[23] Unfortunately, we do not know how many lottery participants were placed on the waiting list because their first choice was unavailable. We do know, however, that these lottery participants appear first on the waiting list and we will make further inquiries to be able to identify them.

receive a plot with the outcomes of the lottery participants who were not selected to receive a plot.

An RCT framework is the most rigorous evaluation design possible for analyzing the impact of being a Di Lottery beneficiary. The strength of this design lies in the proper randomization of individuals to a treatment or a control group such that the two groups are comparable in every way except for being selected as a Di Lottery beneficiary. The public nature of the lottery and its implementation through the national lottery provide strong support for credible randomization. The previous evaluators did not, however, conduct formal statistical tests to confirm whether the treatment and control groups were balanced along the characteristics captured by the baseline survey and the scoring criteria. We provide formal tests as part of this design report (see Table IV.3 for the balance tests related to the scoring variables, and see Appendix F, Table F.1 for the balance tests related to the survey baseline variables).

In addition, our design addresses a number of issues in the previous design:

- The previous evaluator's proposed regression model did not account for the fact that beneficiaries ranked their preferences for the two types of available plots: rice and polyculture. Applicants who would also accept their second choice of a plot have a higher probability of selection than those who accept only their first choice. If these two types of applicants are different, then the treatment can be correlated with the applicants' observable and/or unobservable characteristics, and the standard RCT framework does not apply. The regression model must account for plot preferences by including preference strata fixed effects. This issue is particularly likely to invalidate the standard framework when separately estimating the impact of receiving a plot of a specific type.[24]
- The previous evaluator did not discuss an important issue regarding the reliability of information collected from Di Lottery beneficiaries. Households of some Di Lottery beneficiaries—primarily those who live far from Di—split into two households for the agricultural season: some household members work on the Di perimeter, and others work on the land they previously farmed. As a result, the information gained from a respondent on the Di perimeter may be unreliable with respect to agricultural production and yields for land owned off the perimeter. Similarly, the information gained from a respondent on land off the perimeter may be unreliable with respect to agricultural production on the perimeter.

[24] This concept is most easily explained for the polyculture plot beneficiaries. A higher proportion of applicants chose polyculture plots as their first choice rather than rice plots (85 percent versus 15 percent). As there were similar numbers of polyculture and rice plots available, at some point during the lottery, the polyculture plots were not available anymore and subsequent selected applicants could only receive rice plots. Therefore, only applicants who listed polyculture plots as their first choice (and those who listed polyculture plots as their only choice) received polyculture plots. In the control group, there are applicants whose first preference is a polyculture plot and other applicants whose first preference is a rice plot. If these two groups of applicants are different, then a comparison between the polyculture plot winners and the entire control group will not identify the treatment effect of receiving a polyculture plot.

- Access to communal land in Burkina Faso is likely to be partially need-based. As a result, Di Lottery beneficiaries may lose their customary land rights to other members of their community of origin because they have access to land on the Di perimeter.[25] These other community members would therefore be indirect beneficiaries of the Di Lottery. If we ignore this possible redistribution of land, the estimated total impact on all the Di Lottery beneficiary households would be smaller than the sum of the benefits to them and their communities.[26,27]

3. Proposed evaluation design

We also propose to analyze the Di Lottery by using the framework of an RCT. Our analysis also addresses the above-mentioned issues by (1) incorporating a lottery participant's plot preference into the analysis, (2) interviewing members of both parts of split households, and (3) administering survey questions on communal land rights. We propose two rounds of quantitative data collection. The first, or interim, round will be in September 2017. It will cover the 2016–2017 planting season, roughly between June 2016 and May 2017. The final round will be in 2019, and it will cover the 2018–2019 planting season. Because the quantitative data for the Di Lottery evaluation will be collected jointly with the data collected for the Di PAP survey and the farmer training surveys, we discuss the data collection in more detail in Chapter V.

In the following sections, we describe the quantitative analysis, show that the treatment and control group are balanced with respect to most characteristics, and present minimum detectable impacts for the main analysis and the subgroup analysis. We also describe the proposed WSC through which we will compare the estimated impacts from the RCT with those from an RD.

a. Descriptive analysis

To understand what proportion of Di Lottery beneficiaries received the benefits they were meant to receive (RQ1), we will first develop a list of benefits that were planned. To develop the list, we will rely on project documentation. We will then include a module for Di Lottery beneficiaries in our survey that asks whether the beneficiary received each identified benefit.

b. Impact analysis

The Di Lottery randomly assigned lottery participants to either the treatment or control groups. The RCT relies on this random assignment to estimate the causal impact of winning the Di Lottery and receiving land on the Di perimeter on the outcomes of interest—specifically, on the agricultural outcomes (RQ2) and outcomes related to land tenure security (RQ3).[28] To

[25] Di Lottery beneficiaries may not be planting and harvesting on the land they used before winning the lottery for a number of other reasons, for example, because they leave the land fallow, they rent it out or they let to another household without compensation.

[26] We will not measure production on the land that Di Lottery beneficiaries are not working on anymore, so we will not quantify the spillover. We will investigate whether someone else is benefitting from the land and, if so, why.

[27] The possible impacts on land right and usage is, however, only one of many potential positive or negative effects that the Di Lottery may have on non-beneficiaries.

[28] Where possible, we will collect information on agricultural and land security outcomes at the plot level. When households own too many plots, we may only be able to collect plot-level information for a sample of plots. In order

estimate the causal impact, we estimate the following regression specification using ordinary least squares (OLS):

(1) $$y_i = \alpha + \beta Treatment_i + X_i + \theta_i + \varepsilon_i$$

where y_i is the outcome data variable on which we want to test balance for individual *i*; $Treatment_i$ is an indicator equal to one if individual *i* randomly obtained irrigated land through the lottery; X_i is a vector of demographic, social, and economic characteristics of *i*; θ_i is a plot preference fixed effect for *i* (i.e., an indicator for individual *i*'s plot preferences);[29] and ε_i is a random error term. The parameter of interest is β, which captures the difference between the treatment and control groups. It is the causal estimate of the causal impact of winning the Di Lottery.

In the most basic specification, we do not include any variables as part of the vector X_i. In our preferred specification, the vector X_i includes variables such as gender and land rights that are unbalanced at baseline, as well as available baseline information on agricultural outcomes.[30] In an additional specification, we include all variables used to score the applications in addition to the variables that are unbalanced.

We will conduct subgroup analyses for groups of particular interest. Specifically, we will examine impacts by the applicant's gender, experience in irrigated agriculture, ownership of plots before the lottery, and whether the beneficiary received a rice or a polyculture plot.

Because all Di Lottery winners, in principle, received the same benefits (access to land, irrigation water, farmer training, and agricultural inputs), it is not possible to disentangle their relative contribution to potentially observed impacts.

to calculate outcomes for Di Lottery beneficiaries, for example agricultural income, we add agricultural income derived from the Di perimeter to the agricultural income derived from sources off-perimeter.

[29] In the empirical analysis, we will consider only three preference strata: applicants who accept a polyculture plot only, applicants whose first choice is polyculture plot but who will also accept a rice plot, and applicants whose first choice is a rice plot. We will take this approach because in the actual implementation of the lottery, the categories for which preferences were not clear were treated as if they had made a firm choice. For example, if it was not clear whether an applicant would accept his or her second choice, the actual lottery treated this applicant as if he or she would refuse the second choice. Applicants who stated that they would accept any type of plot were treated as if they had listed a polyculture plot as first choice. Finally, because rice plots were the less desirable plots, they were the last plots to be given out. As such, it did not make a difference whether someone stated that they would accept a rice plot only or whether they would accept a rice plot and then a polyculture plot. The inclusion of the indicator for the applicants who would only accept a polyculture plot is necessary to avoid biased estimates. We will also include an indicator to distinguish applicants who prefer a polyculture plot from those who prefer a rice plot, among applicants who said they would accept either. This indicator is not necessary to avoid bias but will be included to reduce the variance of the estimate.

[30] With few exceptions, we are not able to implement ANCOVA specifications as there is little information on agricultural and household outcomes in the baseline survey.

c. Randomization created comparable treatment and control groups

This section discusses the balance tests we conducted for this report. We used OLS to estimate the following regression specification:

(2) $$y_i = \alpha + \beta Treatment_i + \theta_i + \varepsilon_i$$

where y_i is the baseline data variable on which we want to test balance for individual *i*, $Treatment_i$ is an indicator equal to one if individual *i* randomly obtained irrigated land through the lottery, θ_i is a plot preference fixed effect for *i* (i.e., an indicator for individual *i*'s plot preferences), and ε_i is a random error term. This specification is the same as the basic specification that we will use to analyze outcomes.[31]

Table IV.3 presents the balance tests for the variables used to score the applications in this section. Appendix Table F.1 presents the full set of results. We found that the treatment and control groups are similar along most dimensions. As expected when conducting a large number of comparisons, there are some characteristics along which the two groups are not balanced. With respect to the application data, Di Lottery beneficiary households are slightly smaller on average than control households. In addition, Appendix Table F.1 shows that there are differences in baseline access to land rights; Di Lottery beneficiaries enjoy more rights to land (either by owning land or having rights to communal land), whereas individuals in the control group rent in more land.[32]

Table IV.3. Balance tests for scoring variables

Scoring criteria	Treatment group mean	Control group mean	Difference	P-value of difference
Number of active household members	4.07	4.24	-0.15	0.04**
Applicant owns one piece of agricultural equipment	0.15	0.12	0.03	0.15
Applicant owns at least two pieces of agricultural equipment	0.74	0.75	-0.01	0.80
Applicant received technical training in agriculture	0.41	0.39	0.01	0.62
Applicant has no experience in irrigated agriculture	0.25	0.28	-0.02	0.35
Applicant has less than two years of experience in irrigated agriculture	0.05	0.07	-0.02	0.14
Applicant has two years or more of experience in irrigated agriculture	0.70	0.65	0.04	0.10
Female	0.22	0.19	0.04	0.09*
Age of applicant - 18 to 30	0.40	0.43	-0.04	0.12
Age of applicant - 31 to 55	0.56	0.53	0.04	0.17

[31] The specification does not include the vector of covariates because these covariates would be the baseline variables for which we are conducting the balance tests.

[32] Because the survey was conducted before the lottery, these imbalances are not likely to be a result of false statements.

Table IV.3. *(continued)*

Scoring criteria	Treatment group mean	Control group mean	Difference	P-value of difference
Age of applicant - 56 or older	0.04	0.04	0.01	0.56
Applicant has debt	0.01	0.01	0.00	0.87
Applicant is from village in the rural Di commune	0.56	0.54	0.01	0.85
Applicant is from Sourou province	0.93	0.94	-0.02	0.23
Applicant is from Boucle du Mouhoun region	0.01	0.01	0.01	0.11
Applicant does not have title to a parcel on AMVS perimeters	0.99	0.99	0.00	0.56
Number of observations	**503**	**1,025**		

Source: Di Lottery baseline survey data

*Significantly different from zero at the .1 level, two-tailed test.

**Significantly different from zero at the .05 level, two-tailed test.

***Significantly different from zero at the .01 level, two-tailed test.

d. MDI calculations

To be successful, an impact evaluation must have a large enough sample. For the Di Lottery RCT, the sample size is limited to the number of Di Lottery beneficiaries and the control group. Through a statistical power analysis, we have calculated the magnitude of potential impacts we could detect given the Di Lottery sample size, which we refer to as minimum detectable impacts (MDIs). Table IV.4 provides MDIs for agricultural income.

Table IV.4. MDIs on agricultural income in the Di Lottery RCT for a given subgroup proportion of the entire sample

	Subgroup as proportion of entire sample			
	Full sample	Subgroup with land rights and subgroup without land rights (50% subsample)	Male subgroup (80% subsample)	Female subgroup (20% subsample)
MDI (in 2016 CFA)	78,436	111,919	88,437	177,310
MDI (% of the mean)	15.52%	22.14%	17.50%	35.08%
MDI (% of anticipated Di profit)	3.52%	3.97%	5.02%	7.95%

Note: Calculations are based on a statistical significance level of 0.05 for a two-tailed test with 80 percent power and a sample of 503 treatment households and 1,025 control households in the lottery. We assumed that attrition and nonresponse combined to be 18 percent, and the proportion of the post-intervention outcome explained by covariates (including baseline value) to be 0.3. Calculations are based on a single round of follow-up data collection. The values for the mean and standard deviation are taken from IMPAQ International's (2014a) design report. Values are adjusted for inflation between 2011 and 2016 (Trading Economics 2016). CFA refers to the West African Communauté Financière Africaine franc, the currency in Burkina Faso. "Subgroup with (without) land rights" refers to lottery participants who have (or do not have) informal or formal land rights to at least one plot at baseline. Anticipated profits on the Di perimeter are based on the ERR (MCC 2017).

Based on these power calculations, there is no evidence that the study is underpowered.[33] Given the sample sizes in the Di Lottery RCT, the second row of Table IV.4 shows that the MDI for income is 16 percent of agricultural incomes in a sample of farmers from the same region. This is a reasonable increase; for 70 percent of treatment households, the additional land received through the lottery makes up 25 percent or more of the land they cultivated at baseline. As the irrigated land is, by all accounts, much higher in quality, we found that a 16 percent change in agricultural income is plausible. In a review of the empirical evidence on the effects of irrigation on poverty, Hussain and Hanjra (2004) found several studies that reveal that a 50 percent income gain was attributable to the introduction of irrigation. This is corroborated by a rigorous evaluation in Mali that found that small-scale irrigation raises farmers' agricultural revenues by 30 percent (Dillon 2011). Both of these studies focused on projects that supported irrigation but did not make more land available to farmers for cultivation. Because the Di Lottery does give farmers more land for cultivation and also provides irrigation for this new land, the impacts found in these previous studies are likely to be the minimum level we would expect under the Di Lottery.[34] Therefore, we also provide an alternative approach to assessing the plausibility of the MDI: the third row of Table IV.4 shows the MDI as a percentage of the agricultural profit that, based on the ERR analysis, we would expect the Di Lottery beneficiaries to achieve on one hectare of Di perimeter land. Given the sample sizes in the Di Lottery RCT, the MDIs for income is less than 4 percent of this expected additional income. This suggests that the sample size is sufficient to detect even very small effects.

The power calculations will thus be primarily useful in giving us a more reliable assessment of whether subgroup analyses are possible. We intend to examine the impact of land access on several groups, particularly women and those who did not own land before the activity. The last three columns in Table IV.5 provide illustrative information on MDIs for subgroups whose size is 50, 80, and 20 percent of the sample. These subgroups are stand-ins that represent the approximate share of beneficiaries with and without land ownership at baseline (half of the sample) and the shares for the male and female subgroups (80 percent and 20 percent, respectively).[35] Given the small percentage of female recipients (22.7 percent in the treatment group and 17.9 percent in the control group), Table IV.5 shows that our analysis will be able to detect impacts on agricultural income of about 35 percent (that is, we are not confident that we will be able to detect more moderate impacts of the program on females even if such impacts exist). The MDIs for females are smaller than the impacts described in Hussain and Hanjra

[33] These values were computed from data collected for the Programme Nationale de Gestion de Terroirs, a survey that is not related to the Di Lottery study. Lottery participants had to meet certain criteria and are thus not representative of households in the Boucle du Mouhoun region overall, so these MDIs are illustrative only and provide approximate information only.

[34] There is one caveat to this assessment: if applicants—62 percent of whom had customary rights of access to community land—lose these rights as a consequence of relocating to Di, then the net increase in agricultural incomes for these individuals would be smaller than the overall increase seen in the Di perimeter. Our evaluation design will allow us to answer both questions: What is the change in agricultural incomes for the individuals who won the lottery? To what extent is there an increase in total agricultural production?

[35] The MDIs for the subgroups with and without experience in irrigated agriculture (approximately 75 percent and 25 percent of the sample, respectively) will be similar to the MDIs for the male and female subgroups. About half of the Di Lottery beneficiaries received rice, and the other half polyculture plots, so the MDIs for receiving a specific type of plot are reflected by the 50 percent sample split.

(2004) and in Dillon (2011) and are only 8 percent of the expected additional profit from one hectare of land on the Di perimeter. We therefore intend to conduct a subgroup analysis even for the relatively small subgroup of female lottery winners.

e. Submission of pre-analysis plan to a trial registry

The Di Lottery is unique in at least three ways: (1) it is the only RCT that we are aware of in which a subset of applicants receive irrigated land; (2) the value of the land received by lottery winners—around $45,000 in U.S. dollars—makes the Di Lottery one of the highest-stake RCTs that we are aware of; and (3) the impact estimates will be robust and very credible.[36] As a result of these three factors, the evaluation is very likely to make a significant contribution to the literature on irrigation, access to land, customary land rights, and large-scale transfers in development.

We propose to submit a fully specified pre-analysis plan during Option Period I to the American Economic Association (AEA) trial registry. The AEA trial registry is a dedicated trial registry for economics and other social sciences. It is an important source for meta-analyses and for information on survey instruments and data. It currently lists over 1,000 trials (AEA 2017). The trial registry allows for the submission of a pre-analysis plan to avoid specification searching and thus enhances the credibility of the estimates (Christensen and Miguel, 2016). The pre-analysis plan will specify both the outcomes for which we will conduct the analysis as well as the regression models. We will submit this pre-analysis plan after completing a thorough analysis of the baseline data.

Christensen and Miguel (2016) suggest outlining adjustments for multiple hypothesis tests as part of the pre-analysis plan. More specifically, the pre-analysis plan will separate the primary analysis from the secondary analyses, which are more exploratory in nature. In the primary analysis, we will correct the estimated standard errors for the fact that we conduct more than one statistical test. Known as multiple hypothesis correction this approach accounts for the fact that some estimates are expected to be significantly different from zero because of chance, given that more than one statistical test will be conducted. As shown in Anderson (2008), correcting for multiple hypothesis testing reduces the power of any individual statistical test. We therefore specify three primary outcomes (total agricultural production, total agricultural income, and total household income) for which we will adjust for multiple hypothesis testing. Following suggestions by Schochet (2008), we choose these three outcomes as they are priority outcomes for the Di Lottery. We consider the analysis of secondary outcomes (e.g., outputs and other outcomes) and the subgroup analysis as exploratory and will not adjust for multiple hypothesis testing, as suggested by Schochet (2008).

[36] Appendix D provides information on the calculation of the average cost per hectare based on the cost information contained in the ERR (MCC 2017).

4. Within-study comparison of estimates based on an RD with estimates from an RCT

The admission process for the Di Lottery also allows us to implement an RD, a second rigorous design for estimating the impacts of the lottery on beneficiaries. We will compare the impact estimates derived from this design—and from two recent RD-related extensions—to impact estimates derived from an RCT design in order to determine the relative performance of these methods. This type of analysis is known as a within-study-comparison (WSC) because it compares the impact estimates from two designs in which the study population is the same. The goal of a WSC is to determine whether a less rigorous design would lead to conclusions similar to those drawn from a more rigorous benchmark design. Impacts estimated through an RCT are often chosen as the benchmark because they are free of bias. In many cases, however, an RCT is not feasible—either for practical, political, or ethical reasons—so the analysis of the performance of a less rigorous design shows the extent to which policymakers can rely on the information provided by such designs.

The proposed Di Lottery WSC will make three contributions to the literature on RD. First, it will provide evidence on the validity of estimates derived from RDs in a developing country. The literature shows that RD estimates of program impacts can approximate RCT estimates. However, these findings derive mainly from studies in the U.S. In a recent meta-analysis of WSCs that contrast RD estimates with RCT estimates, Chaplin and colleagues (2017) found that most WSCs are conducted in the context of educational interventions in schools and in labor market settings in the U.S. Evidence from developing countries is more limited. We are aware of only two WSCs involving RDs in developing countries.[37] Although there is no theoretical reason to believe that the validity of an RD would be context-dependent, evidence from a study of agriculture in Burkina Faso that validates RD may be useful in supporting RD in other developing countries.

Second, the proposed Di Lottery WSC will provide evidence on an RD's validity for a nonstandard type of assignment variable. The assignment variable of the Di Lottery was created from multiple variables with mixed—and perhaps even contradictory—objectives. For instance:

- Four of the scoring variables provided more points to individuals for whom outcomes would plausibly be expected to be higher on the Di perimeter, such as individuals who have mechanized agricultural equipment and experience in irrigated agriculture.
- Three of the scoring variables, on the other hand, plausibly pursue distributional objectives (younger applicants, females, and applicants living closer to the perimeter received higher scores).
- The remaining two characteristics—whether an applicant has debt and/or exploits land on another AMVS perimeter—do not fit into either category.

[37] Barrera-Osorio, Filmer, and McIntyre (2014) compare RD and RCT estimates of the impact of a scholarship on education outcomes in Cambodia. Buddelmeyer and Skoufias (2004) compared RD and RCT estimates of the impact of a conditional cash transfer program known as PROGRESA on education outcomes in Mexico.

Common assignment variables in the RD literature are usually pre-program outcomes (test scores and income) and/or continuous variables (age). Of the studies surveyed in Chaplin and colleagues (2017) none rely on an assignment variable that incorporates mixed objectives. Yet this is an important contribution: a transparent selection process that also allows for incorporating multiple objectives might often be politically feasible while alternatives are not.

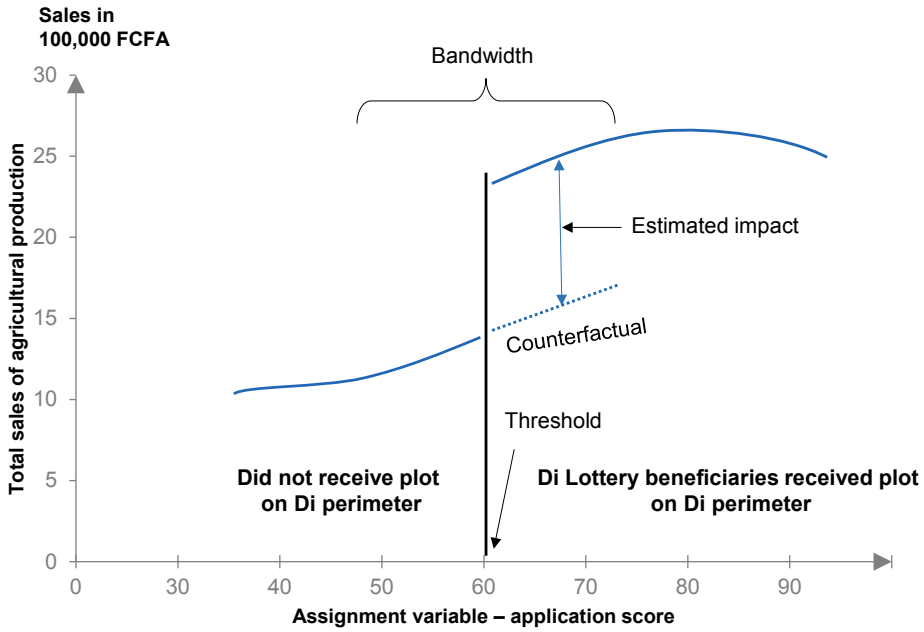
The proposed Di Lottery WSC will contribute to the literature by allowing us to investigate the performance of recently developed methods to estimate impacts for beneficiaries who are away from the threshold. The RD is, however, substantially limited: it estimates only the impact of a program on beneficiaries around the threshold. In many cases, researchers and policymakers are interested in the effect of the program as a whole and/or for a wider group of beneficiaries. We will implement two recent approaches to estimating impacts away from the RD threshold that rely on additional pre-program covariates: (1) the comparative RD (CRD) (Wing and Cook 2013) and (2) an approach developed by Angrist and Rokkanen (2015). We will compare estimates derived from both of these methods to the estimates derived from the RD and the RCT. Both of these methods rely on stronger assumptions than does an RD, and there is an empirical question with respect to how well they perform in practice.

The sections below describe the feasibility of the RD design and present power calculations for it. We then outline the WSC analysis to be conducted, which relies on the design we used for the RD and the two related methods.

a. Feasibility of the RD

As described in Section D.1, the selection of Di Lottery beneficiaries is a two-step process. In a first step, applications were scored, and only those who scored above a threshold of 60 were admitted to the lottery. In the second step, beneficiaries were drawn at random during the lottery itself. RD identifies the impact of the Di Lottery by comparing lottery beneficiaries just above the threshold (also known as discontinuity or cutoff) for admission to the lottery with applicants just below the threshold who were not admitted to the lottery. The approach is illustrated in Figure IV.2.

Figure IV.2 Illustration of the regression discontinuity design



Applicants with a score below 60 constitute the control group, and their outcomes are shown to the left of the threshold. Di Lottery beneficiaries scored above 60, and their outcomes are shown to the right of the threshold. In an RD, we would compare the outcome variable—here agricultural sales on the vertical axis between observations just below and just above the cutoff. The dashed line represents the counterfactual, which is the outcome of the treatment observations had they not been treated. The RD also relies on the assumption that observations just below the cutoff are informative about this counterfactual. Figure IV.2 illustrates an optimal bandwidth that defines which observations are considered to be close enough to the cutoff to be included in the analysis.

In their survey of RD designs, Lee and Lemieux (2010) pointed out that "[for a sharp RD] as in a randomized experiment [. . .], all observable baseline covariates will locally have the same distribution on either side of the discontinuity threshold—an empirically testable proposition." We will test the plausibility of the assumptions underlying RD as part of the analysis of the Di Lottery baseline data. In particular, we will conduct graphical analysis and formal tests, as suggested by Lee and Lemieux (2010).

b. MDI calculations

The first column of Table IV.5 provides the MDI on the value of sales of dry and rainy season agricultural production for the RD, and also presents the MDI as a percentage of the pre-intervention mean of sales and of the value of production per hectare in the Di perimeter. The second column provides MDIs for the RCT. The third column provides the MDIs within the same bandwidth as we use for the RD.

Table IV.5. MDIs on agricultural sales for the RD analysis of the Di Lottery

	Nonparametric RD with optimal bandwidth and bias-corrected standard errors	RCT	RCT within optimal RD bandwidth
MDI (in 2016 CFA)	271,186	55,400	114,262
MDI (% of the pre-intervention mean)	87.72%	17.92%	36.96%
MDI (% of anticipated value of sales per hectare at Di)	8.83%	1.80%	3.72%

Note: Calculations for the RD are based on simulations conducted in the Stata program *rdpower* (Calonico et al. 2014). The simulations account for the loss of power resulting from (1) the selection of an optimal bandwidth specified according to the method proposed by Imbens and Kalyanaraman (2012) and (2) the use of bias-corrected standard errors. The RCT and RD MDI calculations control for age, gender, and the number of household members. We assumed a statistical significance level of 0.05 for a two-tailed test with 80 percent power and the rate of attrition and nonresponse combined to be 18 percent. We censored observations at the 95th percentile. Calculations are based on a single round of follow-up data collection. CFA is the West African franc, the currency in Burkina Faso. The information on the anticipated value of sales per hectare relies on post-compact M&E information on production. Specifically, gain a sense of cash crop production by subtracting the value of rainy season food crop production from the value of total production.

Based on these MDI calculations, the RD will have sufficient power to detect anticipated increases in sales of agricultural production. The RD MDI is equivalent to about 7 percent of the anticipated average value of production per hectare in the Di perimeter, whereas the MDI for the RCT is equivalent to 1.5 percent of this value. Most of this production, in terms of value, is expected to be sold because onions make up over 90 percent of the value of production, and the onion harvest is typically almost entirely sold.[38]

The MDI, when compared with *baseline* sales, is very large. It corresponds to changes in agricultural sales of about 88 percent of the value of baseline sales compared with 18 percent for the RCT. The baseline value of sales is, however, also very low, as a significant proportion of households in the Di Lottery baseline had minimal agricultural sales and produced primarily for their own consumption. The more meaningful perspective on whether the MDI is realistic comes from the comparison with the value of sales that Di beneficiaries can expect.

The third column of Table IV.5 provides MDIs for an RCT that estimates impacts around the threshold of participation. An RD only provides estimates that are generalizable to observations that are near the threshold for program participation. In contrast, an RCT provides an estimate relevant to the entire population covered by the RCT data. In order to estimate comparable impacts in the WSC, the third column includes only observations in the RCT that fall within the optimal bandwidth for the RD.

[38] There are two reasons for this substantial increase in the MDI for the RD: (1) RDs inherently lower power, and (2) the sample for the RD control group is smaller than the sample for the RCT control group.

c. Within-study comparison

We define the RD bias for an outcome as the difference between the RD estimate and the local RCT estimate, as follows:

(3) $$Bias_y = \hat{\beta}_{y,RD} - \hat{\beta}_{y,RCTL}$$

where $Bias_y$ denotes the difference between the RD and RCT estimates for the outcome of interest y; $\hat{\beta}_{y,RD}$ is the RD estimate of the impact of the Di Lottery on outcome y, and $\hat{\beta}_{y,RCTL}$ is the RCT estimate that includes only Di Lottery participants in the neighborhood of the threshold. We will calculate this bias for the outcomes that we describe as our primary and secondary outcomes in the submission to the trial registry. In order to make comparisons across outcomes, we standardize the bias and take the absolute value, as follows:

(4) $$\widetilde{Bias_y} = abs(Bias_y) / sd(y)$$

where $\widetilde{Bias_y}$ is the standardized bias between the RD and RCT estimate for outcome y, and $sd(y)$ is the standard deviation of outcome y. *abs* denotes the mathematical operator that takes absolute value.

We will conduct a similar analysis to compare the estimates of the average treatment effect from the Di Lottery RCT and the RD with the two methods following Wing and Cook (2013) and Angrist and Rokkanen (2015) to estimate the treatment effect away from the RD threshold.

E. Evaluation of the Capacity Building and Technical Assistance for O&M in Sourou sub-activity

The goal of the performance evaluation of the Sourou O&M sub-activity is to assess whether the project created and supported institutions that are operating effectively and maintaining the irrigation infrastructure in the Sourou Valley. Specifically, we will seek to answer the following research questions:

1. How were the O&M activities implemented relative to the original plans?
2. To what extent are the Di perimeter and the old perimeters at Niassan effectively operated and maintained? Are their levels of operation and maintenance sustainable?
   a. To what extent do WUAs on these perimeters have the capacity (financial, technical, and organizational) to fully leverage and maintain the irrigation infrastructure?

The Sourou O&M sub-activity comprised the following project activities:

- Creating and training WUAs
- Providing capacity building for the Autorité de Mise en Valeur de la Vallée du Sourou (AMVS)
- Supporting the creation of the Centre d'Appui Technique et de Gestion (CATG), a private consulting firm, and building the capacity necessary to enable it to provide O&M services to WUAs beyond the end of the compact.

The purpose of this activity was to lead AMVS and WUAs to adopt practices to efficiently operate and maintain irrigation infrastructure.

b. What are the factors influencing the WUAs' level of capacity/implementation?

3. Has the Government of Burkina Faso continued to implement the AMVS reform strategic plan that was developed during the compact? If so, to what extent? If not, why not?
4. To what extent is AMVS successfully fulfilling its O&M responsibilities? What are the reasons for its success or lack thereof?
5. To what extent is the Centre d'Appui Technique et Gestion (CATG) operational? What are the reasons for its success or lack thereof?
   a. If CATG is operational, what percentage of WUAs are benefitting from the services, and what services are they accessing?
   b. What benefits do the WUAs perceive to using CATG? What are WUA perceptions of the quality of CATG services? What specific CATG services do WUAs think are most beneficial? If the percentage of WUAs using CATG is low, why are so few using it?
   c. Is the support that CATG is providing to WUAs financially sustainable?

We will conduct a mixed-methods performance evaluation in which we will use systematic and rigorous qualitative research methods to understand how the interventions have been implemented. We will draw on a variety of qualitative and quantitative data sources, including administrative data from WUAs.

To clearly understand the implementation of Sourou O&M activities and deviations from the initial plans (RQ1), we will review the planning and project implementation documentation. Since the Sourou O&M activity was a complementary activity to the Di perimeter construction, and given the overlap in beneficiaries and implementers, we plan to study the implementation of Sourou O&M activities when we study the implementation of the other Di activities. We describe our approach in Section C.

To determine whether the Di perimeter and the old perimeters at Niassan are effectively operated and maintained (RQ2), we will conduct key informant interviews with staff from AMVS, CATG and the O&M consultants hired by MCA/APD to provide technical assistance to AMVS and CATG during the compact and post-compact. In addition, we conduct in-depth interviews with board members and staff from WUAs to gain the beneficiary's perspective.[39] To gain an independent assessment, we propose to conduct site visits to observe the state of the irrigation infrastructure. To confirm if the level of operation and maintenance is sustainable, we will review WUA annual reports, which summarize water user payments, costs for operations—such as electricity and technical support—and costs associated with maintenance. To verify this information for the Di perimeter, we will also describe WUA payment records, if available.[40] To assess the levels of financial, technical, and organizational capacity (RQ 2a), we will rely on our

[39] When WUAs have not been set up in a perimeter, as is the case in some of the old perimeters, producer cooperatives manage and are responsible for the irrigation infrastructure. In that case, we would select co-op board members responsible for irrigation maintenance to provide the perspective of the "WUA board member."

[40] If WUA payment records are not available, we will use the self-reported information on WUA payments from the quantitative surveys. These may suffer from various biases and are therefore not our preferred source of information on payments.

assessment of the annual reports and on assessments from WUA members and staff, CATG, and AMVS.

To understand the factors that determine the WUAs' level of capacity (RQ 2b), we would need to conduct an impact evaluation, which is not possible in this context. It is possible, however, to explore the perspectives of WUA members and staff, AMVS, CATG, and O&M consultants on which factors they consider important determinants of WUA capacity. Such factors may include whether (1) water payments are collected and penalties are enforced, (2) maintenance is conducted, (3) water needs are fulfilled equitably through an appropriate water schedule, (4) statutory meetings are held, (5) budgets are respected, and (6) a contingency fund is created. We will use the in-depth interviews with WUA members and staff to contrast their operations with those of other sectors in the Di perimeter (for the Di WUAs) or with other old perimeters (for WUAs from the old perimeters). We will also use the in-depth interviews to assess why the operations and maintenance differs.[41] The interviews with staff from AMVS and CATG will similarly touch on their views of what determines WUA capacity.[42]

To understand whether and the extent to which the Government of Burkina Faso has continued to implement the AMVS reform strategic plan (RQ3), we will conduct key informant interviews with current and former MCA/APD staff and in-depth interviews with staff from AMVS and the O&M consultants. Since the action plan also specifies the transfer of authority for nonwater-related support—including farmer training, inputs and post-harvest activities—for farmers on irrigated perimeters in the Sourou Valley to the Ministry of Agriculture and Hydraulic Installations, we will also conduct in-depth interviews with staff from this ministry.

To understand whether AMVS is fulfilling its O&M responsibilities (RQ4), we will rely on interviews with staff from AMVS, O&M consultants, and representatives from the beneficiaries of the O&M activities—that is, the WUA representatives. In addition, we will conduct site visits to assess the state of the irrigation infrastructure under AMVS' remit. To understand the reasons behind AMVS' perceived success or lack thereof, we will, in the interviews, seek the AMVS staff's views on the constraints or enabling factors contributing to AMVS' operations.

The final research questions for the Sourou O&M evaluation are to investigate the current functioning and the sustainability of CATG. The creation of CATG was supported by MCC in order to support the WUAs on technical matters (for example, the maintenance of water pumps), operational matters (developing the water schedules), and financial matters. To understand the extent to which CATG is operational (RQ5) and the WUAs' uptake of CATG's services (RQ5b), we will conduct in-depth interviews with staff from CATG and O&M consultants. We will complement this self-assessment with information from interviews and focus group discussions with WUA leaders and members to gain the beneficiary's perspective, which will provide insight

[41] In comparing operations and maintenance for the Di perimeter with the old perimeters, we will take into account that the maintenance needs and operation costs of older infrastructure are higher.

[42] The analysis of WUA focus groups data will also include whether the female WUA members participated in the focus group discussions.

into why some WUAs are using CATG's services and some WUAs are not.[43] Finally, to understand whether CATG is financially sustainable, we will assess the costs and revenues of CATG if it shares them with us. We will conduct in-depth interviews with CATG staff and O&M consultants, and key informant interviews with staff from APD.

F. Evaluation of the IWRM sub-activity

The main objectives of the evaluation of the IWRM sub-activity are to (1) document whether and how water use and environmental plans have been implemented; (2) examine how the water management institutions created and supported by the compact, the basin committees, and CLE are functioning; and (3) assess the effects of MCC's investments on water management. Specifically, the evaluation seeks to answer the following key research questions:

The IWRM sub-activity comprises the following activities:

- Creating and training 10 local water committees, known as CLE, and 2 basin committees in Mouhoun and Comoé in IWRM
- Providing technical assistance and equipment to two Department of Water Resources and basin-level water agencies in Mouhoun and Comoé
- Establishing basin-level hydrological models
- Developing basin-level IWRM plans, known as SDAGEs

1. How were the IWRM activities implemented relative to the original plans under the ADP?
2. Have the SDAGEs been implemented as planned? What are the primary factors influencing their implementation?
    a. Have activities that were expected to be conducted under the SDAGEs been implemented?
    b. What are the perceived benefits of the SDAGEs for water users?
3. How well are the CLE and basin committee institutions functioning? What are the primary factors influencing their operation?
    a. Have activities that were expected to be conducted by the CLE and basin committee institutions been implemented?
    b. What are the perceived benefits of the CLE and basin committee institutions for water users?
4. Are the water user/polluter fees (CFE) fully defined, and to what extent are they being collected? Are the funds from these fees being directed to the CLEs and the basin committees or to the national treasury?
5. What are the effects of IWRM on (a) water resources and (b) water conflicts?

To answer these research questions, we propose a performance evaluation that will draw on qualitative data from interviews with stakeholders, focus group discussions with water users, and administrative data on CFE payments. The stakeholders include staff at the general directorate of the basin agency (DGAE) tasked with SDAGE design and the implementation of the multiyear plans that operationalize the SDAGEs, basin committee representatives, CLE representatives,

[43] To gain this insight, we will sample from both the 14 WUAs that have taken up at least some CATG services (including the 9 WUAs from the Di perimeter and 5 from the old perimeters) as well as from the remaining 2 WUAs that have not taken up these services.

WUA leaders, and staff at the Permanent Secretariat for IWRM at the Ministry of Water and Sanitation (Ministère de l'Eau et de l'Assainissement). We describe our approach to assessing deviations in implementation from the original plan in Section B and C.

To assess the extent to which SDAGEs are being implemented as planned (RQ2; RQ2a), we will review the SDAGEs, the multiyear plans, and the annual reports. To understand which factors promote or impede implementation, we will conduct in-depth interviews with basin committee representatives, staff from the basin agencies responsible for implementing the SDAGEs, and board members from CLEs who often collaborate with basin committees on specific projects that are part of the multi-year plans.

The review of the implementation of the SDAGEs will also address the question related to the functioning of the basin committee institutions (RQ3) because the main task of the basin agency is to implement the SDAGEs, and the basin committee functions as the legislative organ that oversees the basin agency. To understand the functioning of the CLEs (RQ3, part 2) we will review CLE annual reports. To the extent that the CLEs are engaged in rehabilitating water resources, we will also conduct site visits. Finally, to gain an understanding of the factors that influence the functioning of the basin committee institutions and the CLEs, we will conduct in-depth interviews with basin committee representatives, basin agency staff, and CLE staff. We will also conduct in-depth interviews with the water user representatives on the basin committee and the CLEs to ensure that we understand the perspective of beneficiaries. To further investigate the perceived benefits of the SDAGEs, the CLEs, and the basin committee institutions (RQ2b, RQ3b), we will also conduct focus groups with small and large water users.

To ascertain the extent to which CFEs are defined (RQ4), we will conduct interviews with the basin agency staff who are tasked with collecting water user fees. To understand the extent to which they are collected, we will review administrative data on CFE fee collection. To answer the question of whether CLEs and the basin agencies benefit, we will conduct in-depth interviews with staff at the basin agency and with CLE board members.

To understand the impact of IWRM on conflicts (RQ5b), we will identify situations in which the basin agencies estimated that there were not enough water resources for a group of users. We will conduct focus groups with small and large water users who depend on this water resource to understand the process by which CLEs helped (or did not help) to resolve conflicts. If the outcome was a mutually agreed upon reduction in water consumption, this will also provide an example of IWRM's effects on water resources.

G. Evaluation of farmer training

To assess the effectiveness of the training provided to farmers in agriculture, we will address the following key research questions:

1. How was the farmer training sub-activity implemented relative to the plans for this sub-activity?
2. To what extent have farmers adopted or adapted the improved production practices proposed by the project?

   a. If farmers are adopting improved farming practices, which ones have been adopted the most and the least, and why?
   b. If farmers are adapting improved practices, which ones have been modified the most and the least, and why?
   c. Have farmers continued to invest in improved seeds/fertilizers?
3. Have participating farmers used the incentive kits that they received along with the training?
4. Do participating farmers diversify crop production more than they did before the project?
5. What is the total area planted, average yield/hectare, total production, and total profit for each of the focus crops: rice, corn, onions, tomatoes, soybeans, and cowpeas?
6. Have the participating farmers' average yields/hectare increased, decreased, or remained the same for each of the focus crops, compared with the average yields/hectare before the project?
7. Have the participating farmers' overall agricultural incomes and profits increased, decreased, or remained the same compared with their incomes and profits before the project?

The compact provided technical assistance to farmers in Sourou and Comoé to improve agricultural production techniques and income. The assistance included the following:

- Compost production and use
- Pesticide and chemical fertilizer use
- Use of improved seeds
- Improved planting and harvesting techniques
- Crop rotation

The ADP provided incentive kits to participating farmers (certified seeds or plants, fertilizers, basic farm tools and sacks for post-harvest storage and selling)

Below we discuss the previous evaluation of farmer training and our proposed evaluation design.

1. Previous evaluation

By matching households in intervention areas to households in comparison areas, the previous evaluator planned to use a difference-in-differences design to evaluate the farmer training and animal husbandry components of the project (IMPAQ 2014a). This methodology could in principle provide a credible estimate of the impact of farmer training activities. However, a detailed review of documents from the previous evaluation as well as our site visit raises three major concerns about the difference-in-differences design's ability to detect unbiased impacts: (1) the intervention and comparison groups differ significantly from each other, (2) the location of the intervention communities is highly clustered, and (3) the expected take-up of farmer training was lower than expected. These issues—especially the differences between the intervention and comparison groups—prevent a matched comparison group difference-in-differences methodology from providing credible and unbiased estimates of the impact of the farmer training program. Appendix G provides a detailed assessment of these issues.

2. Proposed evaluation design

We propose a performance evaluation of the farmer training sub-activity in which quantitative and qualitative analyses will be used to answer the key research questions. For the quantitative analysis, we will use a descriptive analysis and a pre-post econometric approach. For

the qualitative analysis, we will analyze information from interviews, focus groups and site visits. Each analysis is described below.

Descriptive analysis and pre-post econometric approach. To answer research questions about agricultural practices—including the use of inputs and incentive kits—and agricultural outcomes (RQ2,3,5), we propose to use a descriptive analysis. To answer research questions about changes in agricultural practices and agricultural outcomes (RQ4,6,7), we propose to use a pre-post approach in which outcomes before the intervention are compared with outcomes after the intervention. Although we will not be able to causally attribute any observed differences to the program, we will be able to gain insight into changes in outcomes over the study period. If no other major changes occurred, we may gain some suggestive indication that the effects are linked to the program. We will compare means of variables before and after the intervention and conduct paired t-tests, which are formal statistical tests of significance. This design will allow us to understand changes in farming practices that happened between the baseline and post-intervention survey(s), such as whether farmers took up some of the techniques conveyed in the trainings and whether their selection of crops changed. In addition, we will conduct a descriptive analysis in which we will present means and standard deviations for all variables of interest, such as adopting or adapting improved techniques, use of incentive kits, and investments in improved seeds and fertilizers.

We will draw the sample of farmers from baseline survey respondents who participated in the agricultural training activities.

Power calculation for pre-post, sample size requirements. Table IV.6 shows illustrative MDIs for three sample sizes. The second column provides illustrative MDIs if we were to re-survey all farmers from the baseline who participated in the training activities. The third column presents MDI for a sample of 400 farmers. A sample of this size would be sufficient to test whether the incomes of training participants increased by 30 percent relative to baseline incomes. This is a large increase in agricultural incomes given that many of the farmer training projects reviewed in Chapter III do not detect impacts of farmer training programs on agricultural income. The fourth column presents the MDI for a 50 percent subsample. The large MDI of over 45 percent of the baseline mean shows that we will not have sufficient power to conduct a subgroup analysis of the farmer training sample. Instead, we will describe agricultural outcomes separately for various subgroups—for example, by region, by the types of training received, and by gender. We propose to collect data on 600 households.

Table IV.6. MDIs on agricultural income: Pre-post analysis in the farmer training evaluation

	Entire sample	Two-thirds subsample	One-third subsample
Sample size	600	400	200
MDI (in 2016 CFA)	117,875	144,485	218,985
MDI (% of the mean)	24.68%	30.25%	45.85%

Note: Calculations assume a two-tailed test with a 95 percent confidence level and 80 percent statistical power. The baseline-to-follow-up correlation in income is assumed to be 0.3, the attrition rate between baseline and follow-up is 15 percent, and the nonresponse rate is 10 percent. Data on agricultural income come from the IMPAQ design report (IMPAQ International 2014a). The standard deviation for the intervention group uses data from the Boucle de Mouhoun and Cascades regions (IMPAQ International 2014a).The extent of imbalance is formally the R-squared in an OLS regression of the intervention on covariates. Clustering is done as described in the text. CFA is the West African CFA franc, the currency in Burkina Faso.

Qualitative analysis. To understand the implementation of the farmer training activity we propose to conduct an implementation study. (Our approach to investigating the implementation of the sub-activities was outlined in Section B.).

To understand (1) whether the techniques learned during the training have been adapted to the local context and why, and (2) why farmers might or might not be using the techniques they have learned (RQ2). We will elicit the views of staff from the regional directorate of the Ministère de l'Agriculture et des Aménagements Hydrauliques. Focus groups with trained farmers and producer associations will provide a beneficiary perspective on the reasons for adopting and adapting the improved training. In addition, we propose to conduct site visits to observe the use (or non-use) and adaptation of the techniques received during the training. We will ensure that the perspectives of female participants are represented in the focus groups.

H. Risks and mitigation strategies.

Although our evaluation design offers the best opportunity for informing the key research questions, it is still subject to several risks. This section summarizes potential risks to our design and describes the steps we will take to mitigate those risks.

- **Staff turnover.** Many of our performance evaluations rely on key informant interviews mainly with staff in government agencies, implementing organizations, or NGOs. Because of political changes in Burkina Faso and the normal rotation in staff, there has been significant turnover in the project implementation staff, including regional heads and staff from the Ministry of Agriculture, the basin agencies, and the CLEs. Also, given that the compact has ended, it is likely that many staff who were involved in the project have had to find new positions. As a result, we anticipate that some pertinent staff will be difficult to locate or may not be willing to provide time for an interview.

  Mitigation strategy. We propose to conduct slightly more interviews with former and current staff at implementing agencies to address these concerns. We will collaborate closely with APD and attempt to interview both current and previous lead staff as well as technical staff, who are less likely to have changed.

- **Sample attrition at the follow-up survey will reduce our ability to detect statistically significant effects**. It may be difficult to find and survey respondents because the population is mobile and the baseline information for identifying them is limited. As a result, the analytic sample may be smaller.

  Mitigation strategy. We will institute data collection protocols to identify baseline respondents and to limit sample attrition. The protocols include (1) using all relevant baseline data, including gender and age, to identify respondents; (2) meeting with village chiefs to help locate households; (3) conducting surveys at a time of year and time of day that maximizes respondent availability; and (4) returning to households several times if respondents are initially absent. We will also collect additional identifying information when the interim survey is administered to ensure that we can easily find respondents in the final data collection; the information may include GPS coordinates and directions to respondents' houses. For Di crop-cutting surveys, we will sample plots from a census of plots instead of plot owners, eliminating the need to trace respondents.

- **Biases (memory, recall, and perception).** Because of the time lapse between decision making, project implementation, and the survey administration, it may be difficult for respondents to remember details about or the time at which activities took place. This bias is likely to be particularly salient for compact design decisions made nearly 10 years ago and strategic implementation decisions made 5 to 7 years ago. In addition, recall bias may be a problem for all respondents. For example, farmers may not precisely recall information on their agricultural production over the past year. Perceptions also may have changed over time, leading to inaccurate answers to retrospective questions.

  Mitigation strategy. Mathematica has extensive experience conducting retrospective evaluations and mitigating against these types of biases. When memory problems are likely to be particularly salient, we will give more weight to project documentation. To aid with recall, our interviewers will be carefully trained to help respondents reference the appropriate time frame, such as anchoring questions to national events such as political events and/or to seasons such as just after the rice harvest. Our survey module on agricultural production will also incorporate best practices in survey design, such as those found in Grosh and Glewwe (2000).

- **Social desirability.** Respondents may be hesitant to give answers that might cast the project in a negative light. In particular, farmers may be wary of expressing a negative perception out of fear of retribution or losing their land.

  Mitigation strategy. Mathematica will work only with well-trained professional interviewers who can relate to and instill confidence in the respondents. The data collection firm, as well as each individual interviewer, will be vetted to ensure that there is no conflict of interest or prior involvement in compact activities that may cause any respondent to feel uneasy. Mathematica will adhere to stringent confidentiality standards governed by an IRB and obtain informed consent from all respondents prior to data collection, ensuring that respondents feel comfortable and know that their confidentiality will be protected.

- **Availability/survey timing.** The availability of respondents can vary greatly based on the agricultural season.

  Mitigation strategy. The timing of our data collection will follow the planting and harvesting seasons in our study sites while being respectful of the farmers' availability. Interviewers can plan to stay in a village for extra time so that the interviews can take place at the respondents' convenience. In addition, the survey and protocols will be designed to minimize respondent burden.

V. SUMMARY OF DATA SOURCES

We will collect quantitative and qualitative data from a range of sources to evaluate the WMI and DA activities. Below we discuss our data collection plans for each.

A. Quantitative data collection plan

Our design calls for collection of survey data on the ADPs activities' key outcomes directly from households. Survey data will be collected by a local data collection firm procured by Mathematica. We anticipate a common ADP survey with separate modules focusing on the Di perimeter, the Di Lottery, and Farmer Training.

An integrated ADP survey leverages efficiencies across the three evaluations in survey design, testing, training, survey administration, and analysis. For instance, modules on agriculture production, crop choices, yields, irrigation access, farming inputs, and agricultural and household income will overlap between the samples. The questionnaire will then also include specialized modules that relate to specific evaluations. These include a module to capture water user payments for the Di PAP and Di Lottery surveys and a module on land rights in the area of origin for the Di Lottery survey. Table V.1 provides an overview of the sample, rounds of data collection and survey modules by evaluation. The table also highlights common modules and modules specific to each evaluation, and where applicable, the respondent category within an evaluation.

Table V.1. Primary quantitative data collection overview

Sample	Data collection round	Sample size	Modules
ADP survey			
Common modules			
Di perimeter beneficiaries (incl. Di Lottery beneficiaries);	• interim, final	917	• Agricultural practices (crop choice, area planted, input use, agricultural techniques [including particular focus on improved techniques learned under the DA activity)
Di Lottery applicants;	• Interim, final	2,178	
Farmer training beneficiaries	• interim	600	• Agricultural outcomes (production, sales, total agricultural income)
Additional modules Di perimeter			
Di PAPs	• interim, final	275	• Implementation outputs (titles, leases, training, starter kits) • Employment outcomes (self-employment, off-farm employment), household income • Perceptions of land tenure security • Land investments, land rental or sales • Use of land as collateral • Payments to WUAs • WUA labor contributions • Water availability (part of input use) • Use of rehabilitated markets and MIS
Crop cuttings	• final	253	• Crop cuttings for focus crops
Additional modules Di Lottery			
Di Lottery applicants	• interim, final	2,178	• Employment outcomes (self-employment, off-farm employment), household income • Perceptions of land tenure security • Land investments, land rental or sales • Use of land as collateral • Plot-level information on agricultural outcomes off-perimeter
Di Lottery beneficiaries	• interim, final	503	• Implementation outputs (titles, leases, training, starter kits) • Payments to WUAs • WUA labor contributions • Water availability (part of input use) • Individual outcomes by gender (expenditures, control over resources, education)
Additional modules farmer training evaluation			
Farmer training beneficiaries	• interim	600	• Implementation outputs (training, starter kits) • Employment outcomes (self-employment, off-farm employment), household income • Benefit of other ADP activities
Farmers in Sourou with plot on old perimeters	• interim	171	• Payments to WUAs • WUA labor contributions • Water availability (part of input use)

Table V.1. *(continued)*

Sample	Data collection round	Sample size	Modules
WUA data-collection			
WUA payment records	• interim	13	• WUA payments
WUA annual reports	• interim	13	• Revenues and expenditures
MIS data			
Agridata/Ecodata	• interim, final	N/A	• Prices for Di and markets surrounding the perimeter. Prices for markets further away from Di

The timing of our data collection will follow the planting and harvesting seasons in our study sites, and for onions, rice, tomatoes and corn in particular. To maximize efficiencies in travel and interviewer training, we will collect primary data at the same time in all evaluations for each agricultural season. One round of data collection will cover the 2016/2017 agricultural season during the fall of 2017, which will concentrate on medium term outcomes. To evaluate long term outcomes, we will do two rounds of data collection, one to collect information on agricultural production during the 2019 dry season and the second to cover agricultural production during the 2019/2020 rainy season. As part of the 2019/2020 data collection, we will also conduct cuttings for key crops over the course of both agricultural seasons.

We will conduct a pilot test before data collection to assess whether respondents can interpret the items as intended, whether the answer options are appropriate, and whether there is variation in responses. Mathematica will work closely with a local data collection partner to train interviewers and monitor the data collection effort. For example, if the data collection firm uses computer-assisted personal interviewing (CAPI), this would enable us to review the data and conduct consistency checks on an ongoing basis. To minimize attrition, we will also track respondents in the Di Lottery baseline survey who migrated within the Sourou Valley and to important migration destinations, such as Tougan and Ouagadougou. We will ask neighbors and local authorities for contact information for these migrants and then contact them to set up interviews.

B. Qualitative data collection plan

Our design also calls for qualitative data collection. Working with a local data collection firm procured by Mathematica—which could be the same firm that collects the survey data—we will collect qualitative data to support the six proposed evaluations. For each evaluation, we will draw on a variety of data sources, including implementers and program participants whose knowledge and perspectives differ and complement each other. This variety of sources will give us a comprehensive picture of the interventions and help us triangulate information during data collection and analysis. We will use a number of qualitative methods to gather data, choosing the method that will produce the richest and most relevant data for the questions to be answered. The qualitative data will help us understand the implementation of the various projects, the decisions made, and the successes and challenges of different aspects of the interventions.

Data sources. As shown in Table V.2, we will speak to a wide range of stakeholders. These include program implementers—such as former and current staff from MCA and APD, technical consultants, and relevant ministry staff who helped implement and oversee the project—as well

as beneficiaries and members of associations created or supported by the project. In general, interviews with implementers will focus on project implementation, whereas our interviews and focus groups with beneficiaries and association members will focus on stakeholder perceptions of implementation and impacts. In addition, we will use compact documents, reports, and administrative data to help analyze project implementation, including any deviations from the initial design.

Qualitative methods. The choice of qualitative method—including key informant interviews, focus group discussions, observations, and document review—will reflect the type of information we are seeking from each source. For example, we will use focus group discussions when we need to obtain a multitude of opinions and experiences, as the interactive nature of the group will generate ideas from many perspectives. Key informant interviews will be used when we are trying to obtain in-depth information from people who are particularly knowledgeable about certain aspects of the project or the project overall. These interviews will also be used to corroborate information from focus groups or to gather additional information from leaders of various organizations created by or affected by the subactivities. We will conduct observations along with site visits to gain a more in-depth understanding of the settings and circumstances of the interventions. Finally, we will use document review to supplement other information we have already gathered. Table V.2. summarizes these data collection methods and the evaluation and areas of focus they will inform.

Table V.2. Qualitative data collection by evaluation and source

Data source	Data collection method	Number	Evaluation	Area of focus
Project documentation				
Compact documents	Desk Review	NA	All evaluations	• Project implementation/Deviations from design
Reports from Implementers	Desk Review	NA	All evaluations	• Project implementation/Deviations from design
Monitoring data	Desk Review	NA	All evaluations	• Project implementation/Deviations from design
MCA / APD / other implementing agencies				
Former and current staff from MCA / APD	Interviews	8	All evaluations	• Project implementation/Deviations from design
			Integration of ADP project activities	• Project design
Former consultants and staff from AD7 and AD10	Interviews	5	All evaluations	• Project implementation/Deviations from design • Regional differences in implementation
Land registrar at Di town hall	Interview	1	Di perimeter evaluation	• Delivery of land tenure instruments
Former and current AMVS staff	Interviews	2	Di perimeter evaluation, including Sourou O&M	• Project implementation/Deviations from design • AMVS action plan • Irrigation maintenance on the Di perimeter and old Niassan perimeters • WUA capacity, and determinants of capacity • Life-span of irrigation infrastructure and evolution of land productivity
Staff from Regional directorate of Ministry of Agriculture	Interviews	4	Di perimeter evaluation, including Sourou O&M	• Agricultural production on Di • AMVS action plan • AMVS O&M responsibilities
			Farmer training	• Adoption and adaptation of techniques from farmer training
Staff from Ministry of Water Resources	Interview	1	IWRM	• SDAGE implementation • CFE
Basin committee members and staff from basin agencies	Interviews	6	IWRM	• SDAGE implementation • Functioning of basin institutions • CFE • Factors determining functioning of IWRM institutions

Table V.2. *(continued)*

Data source	Data collection method	Number	Evaluation	Area of focus
Beneficiaries & others				
PAPs	Focus group discussions Interviews	4 8	Di perimeter	• Potential harms to PAPs • Changes in well-being • Gender-specific changes in well-being • Perceptions of the compensation, of the process of compensation, and of land security
WUA presidents from Di perimeter	Interviews	4	Di perimeter evaluation; Sourou O&M	• Potential harms to PAPs • WUA capacity, and determinants of capacity • AMVS O&M responsibilities • Quality of CATG services
WUA board members and staff from Di	Focus group discussions	4	Di perimeter evaluation; Sourou O&M	• Potential harms to PAPs • Perceptions of the compensation, of the process of compensation, and of land security • WUA capacity, and determinants of capacity • AMVS O&M responsibilities • Quality of CATG services
CATG staff	Interviews	4	Di perimeter evaluation; Sourou O&M	• Potential harms to PAPs • WUA capacity, and determinants of capacity
WUA presidents and board members from old perimeters	Focus group discussions	3	Sourou O&M	• WUA capacity, and determinants of capacity • AMVS O&M responsibilities • Quality of CATG services
Members of CLE governing bodies	Interviews	6	IWRM	• SDAGE implementation • Functioning of basin institutions • Factors determining functioning of IWRM institutions
Representatives of large water users such as Sosuco and mining companies involved in water conflict	Interviews	3	IWRM	• Effect of CLEs on water conflicts
Small water users involved in water conflict	Focus group discussions	3	IWRM	• Effect of CLEs on water conflicts • Perceived benefits of SDAGEs, CLEs, and basin committee institutions
Farmer training beneficiaries; Members of producter associations	Focus group discussions	4	Farmer training	• Adoption and adaptation of techniques from farmer training

Table V.2. *(continued)*

Data source	Data collection method	Number	Evaluation	Area of focus
Site visits				
Rehabilitated markets	Site visits	4	Integration of ADP activities	• Integration of ADP activities • Functioning of markets • Use of markets by farmer training and Di beneficiaries
Di perimeter; Sourou irrigation infrastructure	Site visits	2	Sourou O&M	• State of Di irrigation infrastructure • Life-span of irrigation infrastructure and evolution of land productivity • State of irrigation infrastructure under AMVS responsibility
CLE locations	Site visits	2	IWRM	• CLE involvement in rehabilitation of water resources
Farms of farmer training beneficiaries	Site visits	4	Farmer training	• Adoption and adaptation of techniques from farmer training

Sample selection. We will identify our criteria for selecting participants before fielding the study. Certain key informants will be selected purposively, based on their role or experience. For example, we will attempt to interview the staff member who is most knowledgeable regarding each aspect of the implementation, but we will also strive to avoid burdening any one agency. Other participants will be randomly selected in an effort to reduce bias. For farmer training participants, we will gather lists of sample frame members in the most comprehensive and unbiased manner possible and randomly select participants from the lists; we will use selection criteria to ensure balance and variation based on factors such as geography, demographic characteristics, and so on. For members of PAP households, we will use our criteria to identify participants through contacts and choose them purposively, being as transparent as possible regarding how they were selected. The composition of the focus groups will take a number of elements into consideration, including people's demographics, experiences with the project, and geographic characteristics. The local data collection firm will handle participant selection, in conjunction with Mathematica.

Sample sizes. We have chosen sample sizes of between 6 and 12 interviews per homogeneous group and between three and six focus groups per homogeneous group. We made these decisions based on research that shows these sample sizes will result in saturation (Namey et al. 2016), which is the point when further data produce little or no new information. This is therefore the most efficient use of resources to maximize learning. We will use smaller sample sizes when we expect the variety of responses to be limited and the respondents to be able to provide great depth of information.

Data collection activities. Working with our local research consultant, Mathematica will take the lead in obtaining documents and other administrative data. This includes reports from the APD on the delivery of land tenure documents (for which we also have quantitative information from the Di perimeter and Di Lottery surveys). In preparation for focus groups and interviews, Mathematica will develop tailored data collection protocols that cover similar themes across participant types. A common set of themes across interviews and focus groups will

facilitate triangulation of findings during analysis. Evaluation team members will travel to Burkina Faso for pretesting or piloting of protocols, training, and oversight of data collection. Mathematica and the local data collector will conduct interviews jointly for those that can be done in French; the local data collector will conduct local-language interviews. The local data collection firm will transcribe the interviews and translate them into French when necessary. The firm will also clean the data, which will include reviewing transcripts for fidelity to the recordings, adding definitions of acronyms and jargon, and adding notes for context.

Data analysis. After the transcripts are cleaned, they will be transmitted to Mathematica. Mathematica will conduct coding, during which it will identify themes that emerge from the data for each research question. We will use a simple theoretical framework for this task, organizing stakeholder input into logic model categories (program design versus implementation versus results) as well as program components (Di perimeter versus farmer training versus IWRM). Coding and analysis will be conducted using NVivo, a proven data analysis software that helps identify themes across many diverse respondent groups and data collection methods. Given that responses from all participant types will be coded during this step, triangulation will be key at this stage. Once the data have been coded for the second time, we will write summaries of the themes, highlighting our findings. Finally, we will integrate the findings from all data sources into a detailed final report, which will include pervasive perspectives as well as contrary opinions and cases. Our coding and analysis processes will enable us to develop a key set of qualitative findings across respondent groups, painting a comprehensive picture of the implementation of each sub-activity and the outcomes.

Timeline. We will conduct one round of qualitative data collection in fall 2017 across all evaluations, being mindful of planting seasons and other periods in which potential respondents are less available. We will collect data on the implementation of the sub-activities and provide evidence on the medium-term evolution of the interventions post-compact. Because the compact activities have been completed, and given the overlap in data sources and the fixed costs of data collection, collecting the data simultaneously across sub-activities is the most efficient way to meet the needs of the evaluation, lessen the burden on farmers, and reduce disruptions to communities.

VI. ADMINISTRATION

Given the complexity of this multicomponent project and evaluation, careful management of the evaluation and timeline is essential. In this section, we discuss administrative issues related to the evaluation and present a timeline of evaluation activities.

A. Summary of IRB requirements and clearances

Mathematica is committed to protecting the rights and welfare of human subjects by obtaining approval from an institutional review board (IRB) for relevant research and data collection activities. IRB approval requires three sets of documents: (1) a research protocol in which we describe the purpose and design of the research and provide information about our plans for protecting study participants, their confidentiality, and their human rights, including how we will acquire consent from individuals for their participation; (2) copies of all data collection instruments and consent forms that we plan to use for the evaluation; and (3) a completed IRB questionnaire that provides information about the research protocol, how we will securely collect and store data, participants' protection, and any possible threats to participants resulting from the study or any compromise of data confidentiality. For example, we will ensure that interviewees, survey respondents, and participants in the focus groups are not identified in the reports. We expect our documents to qualify for an expedited review by the IRB because the study presents minimal risk to participants. IRB approval is valid for one year, and we will submit annual renewals for subsequent years as needed.

We will ensure that the study meets all U.S. and local research standards for ethical clearance. Mathematica will submit the research protocols and instruments to its U.S.-based IRB and the local IRB in Burkina Faso, the Conseil National de la Statistique du Burkina Faso. The local data collection firm hired by Mathematica will obtain permits or clearances from the relevant national and/or local government offices before starting field work. If either the U.S. IRB or the local IRB recommend changes to protocols or instruments, the data collection firm, MCC, and Mathematica will work together to accommodate the changes, and all parties will agree on the final protocols before the start of data collection.

B. Preparing data files for access, privacy, and documentation plan

The qualitative and quantitative data collected for this evaluation will be stored on Mathematica's secure server and will only be accessible to project team members. After producing and finalizing the interim and final evaluation reports, we will prepare corresponding de-identified data files, users' manuals, and codebooks based on the quantitative data. We understand that these files could be made available to the public, so we will de-identify these data files, user manuals, and codebooks according to the most recent guidelines set forth by MCC. Public use data files will be free of personal or geographic identifiers that would permit unassisted identification of individual respondents or their households, and we will remove or adjust variables that introduce reasonable risks of deductive disclosure of the identity of individual participants. We will also recode unique and rare data using top and bottom coding or by replacing these observations with missing values. If necessary, we will also collapse any variables that makes any individual highly visible because of geographic or other factors into less easily identifiable categories. We will not submit qualitative data as restricted or public use files, though we will submit qualitative instruments and codebooks.

C. Dissemination plan

To ensure that the results and lessons from the evaluation reach a wide audience, we will work with MCC to increase the visibility of the evaluation and findings targeted to the agricultural sector, particularly for policymakers and practitioners. During the first year of the evaluation, we will release outreach materials based on our final design report to inform and engage stakeholders in the evaluation process. We will ensure that these materials are distributed to the Ministry of Agriculture, local authorities involved in land tenure activities, and other representatives of the Government of Burkina Faso. The findings from the interim and final reports will be presented to MCC in Washington, DC, and to key stakeholders in Burkina Faso. The interim and final evaluation reports will be available on the MCC website within six months of the drafts being submitted.

We expect the broader research community to have a strong interest in the findings from the evaluation and particularly in the results of the Di Lottery and the WSC in which results from the RCT and the RD are compared. To facilitate wider dissemination of findings and lessons learned, we will collaborate with MCC and other stakeholders to identify additional forums—conferences, workshops, and publications—to disseminate the results and encourage other donors and implementers to integrate the findings into their programming.

D. Evaluation team roles and responsibilities

Our team will contribute our extensive experience and expertise to meet MCC's evaluation needs. Dr. Christopher Ksoll, the project director, will oversee the design and implementation of the evaluation. Dr. Ksoll also has primary responsibility for coordinating deliverables and for ensuring that the quality of work is high and that it is completed on time and within budget. He will also lead the quantitative analysis. He has extensive experience in conducting RCTs and is also a co-author on a recent WSC in which estimated impacts from an RCT are compared with estimated impacts from a matched comparison group design.

Dr. Kristen Velyvis will lead the performance evaluations, direct data collection activities, and lead the analysis of qualitative data. Mr. Matt Sloan will ensure that only high quality deliverables are produced by the team. Dr. Chantal Toledo, a researcher on the team, conducted the project's evaluability assessment and will assist Drs. Ksoll and Velyvis with analysis and reporting. Mr. Seth Morgan will support Dr. Ksoll and Dr. Toledo in the technical design process and in the quantitative analysis activities. Ms. Anca Dumitrescu will support the qualitative study and assist key staff in carrying out the evaluation. Mr. Zeyad El Omari manages the project internally for Mathematica and supports programming and research activities. Our team also draws on our expert consultants, Dr. Yiriyibin Bambio and Dr. Niels Hanssens, as well as other Mathematica staff.

E. Evaluation timeline & reporting schedule

The evaluation activities will be clustered into two time periods corresponding to the interim and final data collection. Interim data collection will involve household surveys for the Di perimeter and Di PAP, the Di Lottery and the farmer training surveys, as well as qualitative data collection in the third quarter of 2017. We will produce a report summarizing the findings from these data. We expect to finalize the report in the second quarter of 2018 after we have presented the draft report to stakeholders and obtained their feedback.

The final data collection will include two rounds of the household surveys—one for the dry season and one for the rainy season—for the Di PAP and Di Lottery households, as well as crop cuttings covering both seasons for a representative sample of plots on the Di perimeter. These data will inform the final evaluation report, which we will finalize by the end of our evaluation contract in the third quarter of 2019, again incorporating feedback on the draft report from stakeholders.

Figure VI.1a. Evaluation timeline and reporting schedule: Option period 1

Period of performance	Option period 1																	
Calendar year	2017							2018										
Month	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N
Task																		
Option period 1 (Phase 2)																		
Complete baseline analysis, reports, and anonymization																		
Finalize baseline report			▲															
Anonymize data				▲														
Revise interim evaluation materials																		
Draft data collection TOR	▲																	
Draft revised data collection instruments and training manuals in French and English		▲																
Summarize pilot test results			▲□															
Finalize data collection instruments and training manuals in French and English; IRB approval				▲														
Implement interim data collection																		
Oversee interim data collection and write data collection report				□		▲												
Develop interim report																		
Draft interim evaluation report										▲								
MCC and stakeholder feedback with response											Δ							
Finalize interim evaluation report												▲						
Anonymize data and prepare public use file														▲				
Disseminate interim results																		
Produce PPTs and present results to MCC and local stakeholders														Δ▲□				

Δ = Meeting with MCC; □ = Trip to Burkina Faso; ▲ = Report/deliverable

Figure VI.1b. Evaluation timeline and reporting schedule: Option period 2

Period of performance	Option period 2																	
Calendar year	2018	2019												2020				
Month	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M
Task																		
Option period 2 (Phase 3)																		
Revise final evaluation materials																		
Update evaluation design report	▲																	
Draft revised data collection instruments and training manuals in French and English				▲														
Summarize pilot test results					▲													
Finalize data collection instruments and training manuals in French and English; IRB approval					▲													
Implement final data collection																		
Crop cuttings (only Di perimeter)			—	—	—	—	—	—	—	—	—	—	—	—				
Oversee data collection and write data collection report							□	—						□	▲			
Develop final report																		
Data-analysis and draft final evaluation report									—	—	—	—	—	—	—	▲		
MCC and local stakeholder feedback with response																	Δ	
Finalize evaluation report																	—	▲
Anonymize data and prepare public use file																—	—	▲
Disseminate final results																		
Produce PPTs and present results to MCC and local stakeholders																	▲□	

Δ = Meeting with MCC; □ = Trip to Burkina Faso; ▲ = Report/deliverable

This page has been left blank for double-sided copying.

REFERENCES

A2F. "MCC Final Report. Evaluation of the Access to Rural Finance Activity in Burkina Faso." Bethesda, MD: A2F, October 2015.

AEA RCT Registry. "FAQ" https://www.socialscienceregistry.org The American Economic Association's registry for randomized controlled trials. MIT, Cambridge, MA. Accessed February 20, 2017.

Aker, J. "Rainfall Shocks, Markets, and Food Crises: The Effect of Drought on Grain Markets in Niger." Revue d'Economie de Développement, vol. 1, no. 3, 2010, pp. 71–108.

Aker, Jenny C. "Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger." *American Economic Journal: Applied Economics,* vol. 2, no. 3, 2010, pp. 46–59.

Aker, Jenny C., and Christopher Ksoll. "Can Mobile Phones Improve Agricultural Outcomes? Evidence from a Randomized Experiment in Niger." *Food Policy,* vol. 60, April 2016, pp. 44–51.

Aker, Jenny C., and Marcel Fafchamps. "Mobile Phone Coverage and Producer Markets: Evidence from West Africa." *World Bank Economic Review*. Volume 29, Issue 2 Published: July 2015, pp: 262 – 292.

Akresh, Richard. "(In) Efficiency in Intrahousehold Allocations" IZA Discussion Paper 1858. Bonn, Germany: IZA Institute of Labor Economics, 2008.

African Ministers' Council on Water (AMCOW). "Status Report on the Application of Integrated Approaches to Water Resources Management in Africa." African Ministers' Council on Water, 2012.

Ali, Daniel Ayalew, Klaus Deininger, and Markus Goldstein. "Environmental and Gender Impacts of Land Tenure Regularization in Africa: Pilot Evidence from Rwanda." *Journal of Development Economics*, vol. 110, 2014, pp. 262–275.

Anderson, Michael. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association,* vol. 103, no. 484, 2008, pp. 1481–1495.

Angrist, Joshua D., and Miikka Rokkanen. "Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away from the Cutoff. *Journal of the American Statistical Association*, vol. 110, no. 512, 2015.

Argent, J., B. Augsburg, and I. Rasul. "Livestock Asset Transfers with and Without Training: Evidence from Rwanda." *Journal of Economic Behavior & Organization,* vol. 108, 2004, pp. 19–39.

Bandiera, O., R. Burgess, N. Das, S. Gulesci, I. Rasul, and M. Sulaiman. "Can Basic Entrepreneurship Transform the Economic Lives of the Poor?" Economic Organisation and Public Policy Discussion Papers, EOPP 043. London, UK: The London School of Economics and Political Science, Suntory and Toyota International Centre for Economics and Related Disciplines, April 2013. Available at http://sticerd.lse.ac.uk/dps/eopp/eopp43.pdf. Accessed November 14, 2016.

Banerjee, A., E. Duflo, R. Chattopadhyay, and J. Shapiro. "Targeting the Hard-Core Poor: An Impact Assessment." Abdul Latif Jameel Poverty Action Lab (J-PAL), November 2011. Available at https://www.povertyactionlab.org/sites/default/files/publications/110-%20November%202011_0.pdf. Accessed March 2, 2017.

Bardhan, P., and D. Mookherjee. "Land Reform and Farm Productivity in West Bengal." Palo Alto, CA: Stanford Center for International Development, 2007.

Barrera-Osorio, Felipe, D. Filmer, and J. McIntyre. "An Empirical Comparison of Randomized Control Trials and Regression Discontinuity Estimations." SREE Conference Abstract, Spring 2014.

Bauchet, J., J. Morduch, and S. Ravi. "Failure vs. Displacement: Why an Innovative Anti-Poverty Program Showed No Net Impact in South India." *Journal of Development Economics,* vol. 116, September 2015, pp. 1–16.

Beaman, L., A. BenYishay, J. Magruder, and A.M. Mobarak. "Can Network Theory-Based Targeting Increase Technology Adoption?" June 2015. Available at http://economics.yale.edu/sites/default/files/mnw.pdf. Accessed April 12, 2017.

Benin, S., E. Nkonya, G. Okecho, J. Randriamamonjy, E. Kato, G. Lubade, M. Kyotalimye, and F. Byekwaso. "Impact Evaluation and Returns to Investment of the National Agricultural Advisory Services (NAADS) Program of Uganda." Discussion Paper 724. Washington, DC: International Food Policy Research Institute (IFPRI), 2008.

BenYishay, A., and A. Mushfiq Mobarak. "Communicating with Farmers Through Social Networks." Economic Growth Center Discussion Paper No. 1030. New Haven, CT: Yale University, August 2013.

Besley, Timothy, and Maitreesh Ghatak. "Property Rights and Economic Development." In *Handbook of Development Economics*, vol. 5, edited by Dani Rodrik and Mark Rosenzweig. Amsterdam: Elsevier, 2010.

Brasselle, Anne-Sophie, Frédéric Gaspart, and Jean-Philippe Platteau. "Land Tenure Security and Investment Incentives: Puzzling Evidence from Burkina Faso." *Journal of Development Economics*, vol. 67, no. 2, 2002, pp. 373–418.

Buddelmeyer, Hielke, and Emmanuel Skoufias. "An Evaluation of the Performance of Regression Discontinuity Design on PROGRESA." Policy Research Working Paper No. 3386. Washington, DC: The World Bank, August 2004.

Calonico, S., M.D. Cattaneo, and R. Titiunik. "Robust Data-Driven Inference in the Regression-Discontinuity Design." *Stata Journal*, vol. 14, no. 4, 2014, pp. 909–946.

Cerdán-Infantes, P., A. Maffioli, and D. Ubfal. "The Impact of Agricultural Extension Services: The Case of Grape Production in Argentina." Working Paper 5/08.Washington DC: Inter-American Development Bank, Office of Evaluation and Oversight (OVE), 2008.

Chaplin, Duncan D., Thomas D. Cook, Jelena Zurovac, Jared S. Coopersmith, Mariel M. Finucane, Lauren N. Vollmer, and Rebecca E. Morris. "The Internal and External Validity of the Regression Discontinuity Design: A Meta-Analysis of 15 Within-Study Comparisons." Working paper. Washington, DC: Mathematica Policy Research, February 2017.

Chauvin, N. D., F. Mulangu, and G. Porto. "Food Production and Consumption Trends in Sub-Saharan Africa: Prospects for the Transformation of the Agricultural Sector." Addis Ababa, Ethiopia: United Nations Development Programme, Regional Bureau for Africa, 2012. Available at http://www.undp.org/content/dam/rba/docs/Working%20Papers/Food%20Production%20and%20Consumption.pdf. Accessed September 22, 2016.

Christiaensen, Luc, Lionel Demery, and Jesper Kuhl. "The (Evolving) Role of Agriculture in Poverty Reduction—An Empirical Perspective." *Journal of Development Economics,* vol. 96, 2011, pp. 239–254.

Christensen, Garret, and Edward Miguel. "Transparency, Reproducibility, and the Credibility of Economics Research." NBER Working Paper No. 22989. Cambridge, MA: National Bureau of Economic Research, 2016.

Datar, G., and X. V. Del Carpio. "Are Irrigation Rehabilitation Projects Good for Poor Farmers in Peru?" World Bank Policy Research Working Paper No. 5154. Washington DC: World Bank, 2009.

David-Benz, Helene, G. Meijerink, I. Porgo, B. Mdachi, and T. Moyo. "How MIS Can Improve Producers' Market Power and Build Farmers' Organizations Voice?" Kent, United Kingdom: FARMAF Policy Brief, 2015.

Ddungu, S., W. Ekere, J. Bisikwa, R. Kawooya, D. Okello Kalule, and M. Biruma. "Marketing and Market Integration of Cowpea (Vigna unguiculata L. Walp) in Uganda." *Journal of Development and Agricultural Economics*, vol. 7, no. 1, 2015, pp. 1–11.

De Janvry, Alain, and Elisabeth Sadoulet. "Agricultural Growth and Poverty Reduction: Additional Evidence." *The World Bank Research Observer,* vol. 25, no. 1, pp 1-20, 2010.

Deininger, K., and J.S. Chamorro. "Investment and Income Effects of Land Regularization: The Case of Nicaragua." *Agricultural Economics*, vol. 30, no. 2, 2004, pp. 101–116.

Deininger, Klaus, and Gershon Feder. "Land Registration, Governance, and Development: Evidence and Implications for Policy." *World Bank Research Observer*, vol. 24, no. 2, 2009, pp. 233–266.

Deininger, Klaus, and Songqing Jin. "Securing Property Rights in Transition: Lessons from Implementation of China's Rural Land Contracting Law." *Journal of Economic Behavior and Organization*, vol. 70, no. 1–2, 2009, pp. 22–38.

Deininger, Klaus, Songqing Jin, and Vandana Yadav. "Impact of Land Reform on Productivity, Land Value, and Human Capital Investment: Household-Level Evidence from West Bengal." East Lansing, MI: Department of Agriculture, Food and Resource Economics, Michigan State University, 2008.

Dillon, A. "The Effect of Irrigation on Poverty Reduction, Asset Accumulation, and Informal Insurance: Evidence from Northern Mali." *World Development,* vol. 39, no. 12, 2011, pp. 2165–2175.

Dillon, B., and C. Dambro. "How Competitive Are Food Crop Markets in Sub-Saharan Africa?" Working paper. Seattle, WA: University of Washington, 2016.

Duflo, E., M. Kremer, and J. Robinson. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." Working Paper 15131. Cambridge, MA: National Bureau of Economic Research, July 2009.

Duflo, Ester, and Rohini Pande. "Dams." *Quarterly Journal of Economics,* vol. 122, no. 2, May 2007, pp. 601–646.

Essam, T. "Analyzing Millet Price Regimes and Market Performance in Niger with Remote Sensing Data." Ph.D. dissertation. College Park, MD: University of Maryland, 2013.

Feder, G., R. Murgai, and J. B. Quizon. "The Acquisition and Diffusion of Knowledge: The Case of Pest Management Training in Farmer Field Schools, Indonesia." *Journal of Agricultural Economics,* vol. 55, no. 2, 2004, pp. 221–243.

Field, Erica. "Fertility Responses to Land Titling: The Roles of Ownership Security and the Distribution of Household Assets." Durham, NC: Duke University, 2003.

Food and Agriculture Organization of the United Nations (FAO). "AQUASTAT Main Database." 2016. Available at http://www.fao.org/nr/water/aquastat/data/query/index.html?lang=en. Accessed February 19, 2017.

Food and Agriculture Organization of the United Nations (FAO). "Country Fact Sheet: Burkina Faso." Rome, Italy: Food and Agriculture Organization of the United Nations, 2016. Available at http://www.fao.org/nr/water/aquastat/data/cf/readPdf.html?f=BFA-CF_eng.pdf . Accessed September 22, 2016.

Food and Agriculture Organization of the United Nations (FAO). "Improved Animal Health for Poverty Reduction and Sustainable Livelihoods." FOA Animal Production and Health Paper 153. Rome, Italy: Food and Agriculture Organization of the United Nations, 2016. Available at http://www.fao.org/3/a-y3542e.pdf. Accessed November 14, 2016.

Food and Agriculture Policy Decision Analysis (FAPDA). "Country Fact Sheet on Food and Agriculture Policy Trends: Burkina Faso." Rome, Italy: Food and Agriculture Organization of the United Nations, 2014. Available at http://www.fao.org/docrep/field/009/i3760e/i3760e.pdf. Accessed September 22, 2016.

Ghatak, M., and S. Roy. "Land Reform and Agricultural Productivity in India: A Review of the Evidence." *Oxford Review of Economic Policy,* vol. 23, no. 2, 2007, pp. 251–269.

Goldstein, M., and C. Udry. "The Profits of Power: Land Rights and Agricultural Investment in Ghana." *Journal of Political Economy,* vol. 116, no. 6, 2008, pp. 981–1022.

Grosh, Margaret, and Paul Glewwe. "Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study. Volumes 1, 2, and 3." Washington, DC: World Bank, 2000.

Hussain, I., and M. Hanjra. "Irrigation and Poverty Alleviation: Review of the Empirical Evidence." *Irrigation and Drainage,* vol. 53, 2004, pp. 1–15.

Imbens, G., and K. Kalyanaraman. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *The Review of Economic Studies*, vol. 79, no. 3, 2012, 933–959.

IMPAQ International. "Evaluation Design Report: Impact of the Agricultural Development Project Burkina Faso." Columbia, MD: IMPAQ International, 2014a.

IMPAQ International. "MCA-Burkina Faso Agricultural Survey in the Sourou Valley and Comoé Basin: Data Quality Review Report." Columbia, MD: IMPAQ International, 2014b.

Independent Evaluation Group (IEG). "Impact Evaluations in Agriculture: An Assessment of the Evidence." Washington, DC: World Bank, 2011.

IRIS Center (University Research Corporation, International). "Deliverable #12—Preliminary Evaluation Design Report: Final Version. Impact Evaluation Design—Burkina Faso." College Park, MD: The IRIS Center, University of Maryland, 2010.

Jack, B. K. "Constraints on the Adoption of Agricultural Technologies in Developing Countries." White paper. Agricultural Technology Adoption Initiative, J-PAL (MIT) and CEGA (University of California, Berkeley), 2013.

Janaiah, A., V. Thieu, and N. T. Dung. "Poverty Reduction Impact of Public Spending on Large-Scale Irrigation Systems in Vietnam." Manila, Philippines: Asian Development Bank and the World Bank, November 2004.

Jones, M., and F. Kondylis. "Feedback Matters: Evidence from Agricultural Services." World Bank Policy Research Working Paper No. 7768. Washington, DC: World Bank, 2016.

Kabir, H., and N. Uphoff. "Results of Disseminating the System of Rice Intensification with Farmer Field School Methods in Northern Myanmar." *Experimental Agriculture,* vol. 43, no. 4, 2007, pp. 463–476.

Kazianga, Harounan, and Zaki Wahhaj. "Gender, Social Norms, and Household Production in Burkina Faso." *Economic Development and Cultural Change*, vol. 61, no. 3, 2013, pp. 539–576.

Kiiza, B., and G. Pederson. "ICT-Based Market Information and Adoption of Agricultural Seed Technologies: Insights from Uganda." *Telecommunications Policy,* vol. 36, no. 4, 2012, pp. 253–259.

Kondylis, F., V. Mueller, and S. Zhu. "Seeing Is Believing? Evidence from an Extension Network Experiment." World Bank Policy Research Working Paper No. 7000. Washington, DC: World Bank, 2014.

Krishnan, P., and Manasa Patnam. "Neighbors and Extension Agents in Ethiopia: Who Matters More for Technology Diffusion?" Working paper. London: International Growth Centre, March 2012.

Ksoll, Christopher, and Chantal Toledo. "Evaluability Assessment Report for the Burkina Faso Agriculture Development Project." Washington, DC: Mathematica Policy Research, 2016.

Kuwornu, John K. M., and Eric S. Owusu. "Irrigation Access and Per Capita Consumption Expenditure in Farm Households: Evidence from Ghana." *Journal of Development and Agricultural Economics,* vol. 4, no. 3, pp. 78-92, February 2012.

Larsen, A., and H. B. Lilleør. "Beyond the Field: The Impact of Farmer Field Schools on Food Security and Poverty Alleviation." *World Development,* vol. 64, 2014, pp. 843–859.

Lee, David S., and Thomas Lemieux. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, vol. 48, no. 2, June 2010, pp. 281–355.

Linkow, Benjamin. "Causes and Consequences of Perceived Land Tenure Insecurity: Survey Evidence from Burkina Faso." *Land Economics*, vol. 92, 2016, pp. 308–327.

Magesa, M. M., K. Michael, and J. Ko. "Agricultural Market Information Services in Developing Countries: A Review." *Advances in Computer Science: An International Journal,* vol. 3, no. 3, 2014, pp. 38–47.

Matsumoto-Izadifar, Yoshiko. "Senegal: Making Better Use of Agribusiness Potential." *OECD Journal: General Papers,* vol. 2009, no. 2, 2009, pp. 53–78.

Merrey, D., H. Lévite, and B. van Koppen. "Are Good Intentions Leading to Good Outcomes? Continuities in Social, Economic, and Hydro-Political Trajectories in the Olifants River Basin, South Africa." In *River Basin Trajectories: Societies, Environments and Development,* edited by F. Molle and P. Wester. Oxfordshire, UK: Centre for Agriculture and Bioscience International, 2009.

Millennium Challenge Account–Burkina Faso (MCA-BF). "Assistance Technique pour la Mise en Œuvre d'Activités de Diversification de l'Agriculture et d'Accès au Financement Rural dans le Cadre du Projet de Développement Agricole (AD10): Rapport Final d'Activités." Ouagadougou, Burkina Faso: Millennium Challenge Account–Burkina Faso, 2014a.

Millennium Challenge Account–Burkina Faso (MCA-BF). "Atlas des Activités et des Réalisations Entreprises par le MCA-BF Durant le Compact (2009–2014)." Ouagadougou, Burkina Faso: Millennium Challenge Account–Burkina Faso, 2014d.

Millennium Challenge Account–Burkina Faso (MCA-BF). "Burkina Faso Post-Compact Monitoring and Evaluation Plan." Ouagadougou, Burkina Faso: Millennium Challenge Account–Burkina Faso, 2014c.

Millennium Challenge Account–Burkina Faso (MCA-BF). "Rapport d'Achèvement du Compact du Burkina Faso." Ouagadougou, Burkina Faso: Millennium Challenge Account–Burkina Faso, 2014b.

Millennium Challenge Corporation (MCC). "Burkina Faso Agricultural Development Project: Indicator Tracking Table." Washington, DC: Millennium Challenge Corporation, 2014.

Millennium Challenge Corporation (MCC). "Burkina Faso Compact." Washington, DC: Millennium Challenge Corporation, 2016b. Available at https://www.mcc.gov/where-we-work/program/burkina-faso-compact. Accessed September 22, 2016.

Millennium Challenge Corporation (MCC). "Di Land Allocation (as of July 8 2014)." Excel spreadsheet. Washington, DC: Millennium Challenge Corporation, 2016a.

Millennium Challenge Corporation (MCC). "Final Report: Comoé." Washington, DC: Millennium Challenge Corporation, 2008a.

Millennium Challenge Corporation (MCC). "Final Report: Di." Washington, DC: Millennium Challenge Corporation, 2008b.

Millennium Challenge Corporation (MCC). "Issue Brief: MCC's First Impact Evaluations: Farmer Training in Five Countries." Washington, DC: Millennium Challenge Corporation, October 2012. Available at https://www.mcc.gov/resources/doc/issuebrief-2012002119501-ag-impact-evals-pdf. Accessed February 15, 2017.

Millennium Challenge Corporation (MCC). "Project Description: Agriculture Development Project (ADP)." Washington, DC: Millennium Challenge Corporation, 2016a.

Millennium Challenge Corporation (MCC). "Burkina Faso Agricultural Development Project, Di Irrigated Agriculture Closeout ERR." Washington, DC: MCC, 2017. Available at https://assets.mcc.gov/content/uploads/2017/04/mcc-err-burkina-di-irrigation-close.xls. Accessed April 26, 2017.

Mobarak, M. "Using Social Networks to Improve Agricultural Extension Services." Presentation to the Agricultural Technology Adoption Initiative, May 2014. Available at https://www.atai-research.org/using-social-networks-to-improve-agricultural-extension-services/. Accessed April 12, 2017.

Namey, Emily, Greg Guest, Kevin McKenna, and Mario Chen. "Evaluating Bang for the Buck: A Cost-Effectiveness Comparison Between Individual Interviews and Focus Groups Based on Thematic Saturation Levels." *American Journal of Evaluation*, vol. 37, no. 3, 2016, pp. 425–440.

Payne, Geoffrey, Alain Durand-Lasserve, and Carole Rakodi. "The Limits of Land Titling and Home Ownership." *Environment and Urbanization*, vol. 21, 2009, pp. 443–462.

Rey, J., R. Silva, F. Ardino, and H. Lévite. "Mise en Place d'Indicateurs de GIRE par Approche Managériale; Une Application en Afrique de l'Ouest." Presented at the 13th World Water Congress, Montpellier, France, September 1-4, 2008.

Sakurai, Takeshi. "On the Determinants of High Productivity Rice Farming in Irrigated Areas in Senegal: The Efficiency of Large Compared with Small-Scale Irrigation Schemes." Working Paper No. 105. Tokyo, Japan: JICA Research Institute, 2015.

Sally, H., H. Lévite, and J. Cour. "Local Water Management of Small Reservoirs: Lessons from Two Case Studies in Burkina Faso." *Water Alternatives,* vol. 4, no. 3, 2011, pp. 365–382.

Schochet, Peter Z. "Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations." NCEE 2008-4018. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2008.

Schreuder, B., H. Moll, N. Noorman, M. Halimi, A. Kroese, and G. Wassink. "A Benefit-Cost Analysis of Veterinary Interventions in Afghanistan Based on a Livestock Mortality Study." *Preventive Veterinary Medicine,* vol. 26, April 1996, pp. 303–314.

Suri, T. "Selection and Comparative Advantage in Technology Adoption." *Econometrica,* vol. 79, 2011, pp. 159–209.

Tucker, J., and L. Yirgu. "Small-Scale Irrigation in the Ethiopian Highlands: What Potential for Poverty Reduction and Climate Adaptation?" Ripple Briefing Paper No. 3. Addis Ababa, Ethiopia: WaterAid Ethiopia, 2010.

Udry, Christopher. "Gender, Agricultural Production, and the Theory of the Household." *Journal of Political Economy*, vol. 104, no. 5, 1996, pp. 1010–1046.

Udry, Christopher. "The Economics of Agriculture in Africa: Notes Toward A Research Program." *AfJARE,* vol. 5, no. 1, September 2010, pp: 284-299.

USAID Burkina Faso. "Fact Sheet: Agriculture and Food Security." Washington DC: USAID, 2015. Available at https://www.usaid.gov/sites/default/files/documents/1860/BF%20Fact%20Sheet%20-%20Food%20Security.pdf . Accessed September 22, 2016.

Waddington, H., B. Snilstveit, H. White, and J. Anderson. "The Impact of Agricultural Extension Services." Washington, DC: International Initiative for Impact Evaluation, 2010.

Waddington, H., and H. White. "Farmer Field Schools: From Agricultural Extension to Adult Education—3ie Systematic Review Summary 1." London, UK: International Initiative for Impact Evaluation (3ie), 2014.

Wang, Shing-Yi. "Property Rights and Intra-Household Bargaining." *Journal of Development Economics*, vol. 107, 2014, pp. 192–201.

Windsor, P. A. "Perspectives on Australian Animal Health Aid Projects in South-East Asia." *Transboundary and Emerging Diseases,* vol. 58, no. 5, October 2011, pp. 375–386.

Wing, C., and T.D. Cook. Strengthening the Regression Discontinuity Design Using Additional Design Elements: A Within-Study Comparison. *Journal of Policy Analysis and Management*, vol. 32, 2013, pp. 853–877.

Zeitlin, A., S. Caria, R. Dzene, P. Janský, E. Opoku, and F. Teal. "Heterogeneous Returns and the Persistence of Agricultural Technology Adoption." CSAE WPS/2010-37. Oxford, UK: University of Oxford, November 2010.

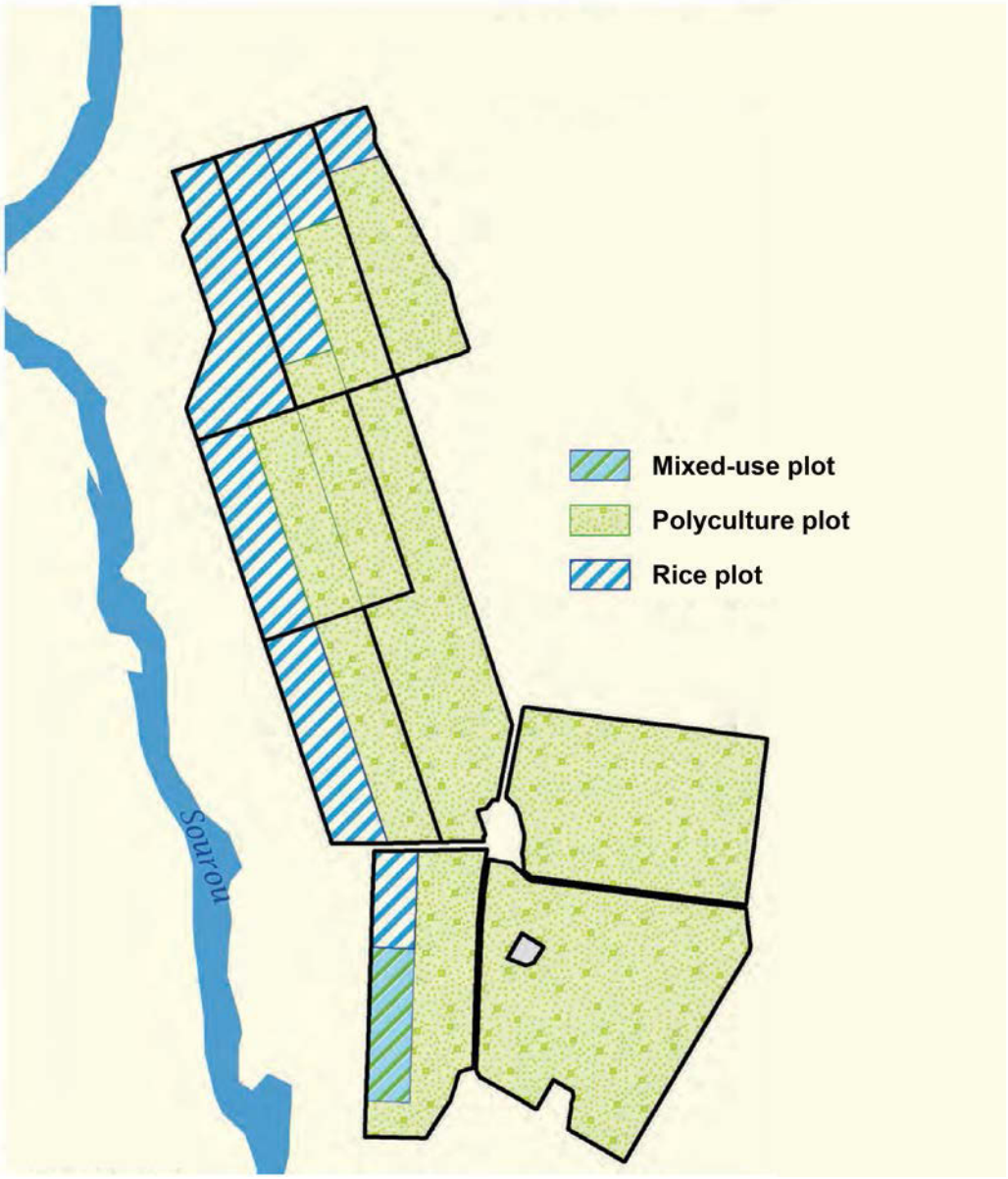
This page has been left blank for double-sided copying.

APPENDIX A

MAPS

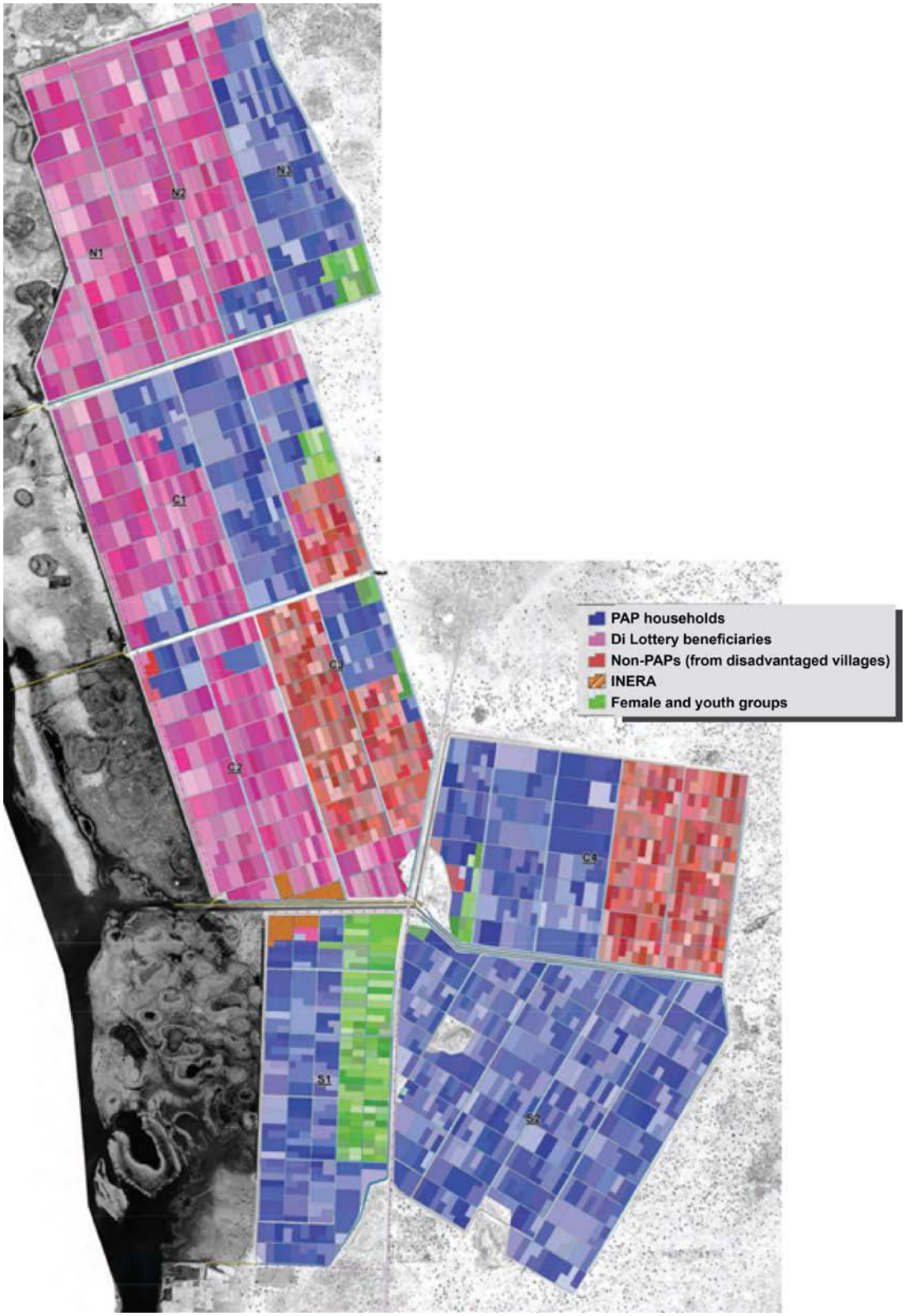
This page has been left blank for double-sided copying.

Figure A.1. Map of plot suitability for rice, polyculture, or mixed cultivation in the Di perimeter



Source: MCA (2014d).

Figure A.2. Map of land allocation in the Di perimeter by beneficiary group



Source: MCA (2014d).

APPENDIX B

PRIORITIZATION OF ACTIVITIES

This page has been left blank for double-sided copying.

This appendix describes the prioritization of evaluation design components. It first provides an overview of the available information on expenditures and beneficiaries for project activities and sub-activities. It then describes our assessment of learning opportunities that incorporate feedback from MCC and APD.

A. Overview of the activity and sub-activity budget

One criterion that MCC uses to prioritize evaluation options is the size of its investment in a particular activity or sub-activity. Table D.1 provides an overview of the available information on expenditures for project activities and sub-activities. Detailed breakdowns are not available for most of the sub-activities that are part of the DA activity.[44] The available information, although not always internally consistent as it derives from six different sources, provides an indication of the relative importance of the activities and sub-activities.[45] We note that the total expenditure for an activity is larger than the sum of the expenditures for its sub-activities. There are additional expenditures for capacity building for government institutions.

The first column of Table D.1 organizes the sub-activities by activity, while the second column provides information on the expenditures for these activities.

It is evident that the overwhelming majority of ADP investments focused on Di.[46] To gain an additional perspective on the size of project activities, the fourth column contains information on the number of beneficiaries for the different sub-activities (where available). This is particularly useful for the DA sub-activities, for which we do not have information on expenditures.

[44] Disaggregated data on compact expenditures by sub-activity were not available because, apart from the construction of the Di perimeter, the construction of the markets and the IWRM training, almost all other technical training activities were awarded as one contract (personal communication with APD).

[45] The six sources are the Atlas of Realizations (MCA-BF 2014d), the land allocation spreadsheet for Di (MCC 2016a), the ADP project description (MCC 2016b), the AD10 final report (MCA-BF 2014a), the indicator tracking table (MCC 2014), and the compact achievement report (MCA-BF 2014c).

[46] The total expenditure for the Di perimeter is even higher than it appears because beneficiaries who received land on the Di perimeter also received farmer training (though we are unable to determine the value of this training).

Table B.1. Sub-activity costs and number of beneficiaries

Sub-activity	Cost (in $)	Beneficiary category	Number of beneficiaries	
Water management and irrigation	**104M^**			**Hectares**
Di perimeter construction	89.0M	All Di beneficiaries		2,240
		PAP (HH)	846	1,099
		Non-PAP from disadvantaged village (HH)	461	317
		Di Lottery beneficiary (I)	503	710
		Women (I)	1,725	90
		Youth (I)	846	16
		Other: Tree nursery, national research institute (INERA), mixed	16 (+INERA)	8
Sourou O&M (included in Di perimeter construction)	6.6M	All Di	~4,381	2,240
		Old perimeters	NA	3,800
IWRM	5.0M^		Two regions	
(Leri dam)	6.0M	Not evaluated by Mathematica	NA	
Diversified agriculture	**30M^**			**Notes**
Technical assistance	21.6M^	Farmer training total (HH)	9,923	
		Farmer training: Comoé (HH)	3,480	*
		Farmer training: Sourou (HH)	4,350	*
		Farmer training: Di (HH)	2,729	
		Animal health	1.4M cows, 1.4M chickens, labs constructed, and vets trained	
		Animal husbandry: Chickens (I)	1,400	
		Animal husbandry: Insemination (I)	442	
		MIS (I)	4,000	+
		Value-added (I)	554	+
		Market management committees	NA	
Rural markets	5.48M		NA	

Sources: Costs are from the Atlas of Realizations (MCA-BF 2014d), except where noted. Beneficiaries of Di are from the land allocation spreadsheet. Beneficiaries of TA are from the AD10 final report, except where noted.

^Costs are from the ADP project description (MCC 2016c).

*Information is from ITT.

+Information is from the achievement report.

HH = beneficiary households; I = individual beneficiaries; NA = not available.

In addition, for the farmer training and the animal husbandry sub-activities, the project documentation contains information on the intensity of the training that beneficiaries received.

Farmer training. Farmers could participate in the farmer training sub-activity for three years, with 8 sessions of three hours each provided each year in groups of 25 to 30 participants. In total, 3,088 training sessions were conducted. Field agents made a total of around 65,000 field visits to individual farmers. Moreover, 514 model farms were set up and around 5,100 incentive kits were distributed (in addition to the kits that Di beneficiaries received).

Animal husbandry. Participants in the animal husbandry sub-activity could participate in up to 6 sessions. In total, 1,664 training sessions were conducted, with 1,400 farmers participating in the poultry training sessions and 442 farmers participating in the cattle training (at least one of their cattle was inseminated). Field agents made around 19,950 visits as part of the livestock activity.

B. Assessment of learning opportunities and evaluation priorities in the ADP

The second criterion that MCC uses to determine whether evaluations should be funded is the opportunity for learning associated with an evaluation. This section provides an assessment of learning opportunities and suggests a prioritization of evaluation activities.

Table D.2 lists activities and sub-activities as well as possible additional evaluation activities. The second column summarizes the available information on expenditures. The third column provides our assessment of the opportunities for learning from an evaluation of this sub-activity or a design addition that incorporates feedback from MCC and APD.[47] The final column provides our categorization into three priority categories—high, medium, or low—based on the size of the investment and the opportunities for learning.

Overall, it is clear that the WMI plays a central role in MCC's investments under the ADP. Within the Di perimeter, the construction of the perimeter and its complementary sub-activities make up the overwhelming share of expenditures. These projects provide substantial opportunities for rigorous learning. The expenditures for the DA activity are lower and spread over many more sub-activities. In addition, the evaluation designs for the DA activity are less rigorous and we see less opportunity for learning. Our categorization into evaluation priorities took these assessments into account to rank the evaluation activities associated with the Di perimeter as high priority. We rank the evaluation activities associated with the DA activity as lower priority. The farmer training activity is an exception, as the level of effort associated with this activity was substantial and there are substantial opportunities for learning.

[47] Some components can be thought of as having more overlap and synergies and might be analyzed together. This is true, for example, for the MIS and rural markets sub-activities.

Table B.2. Assessment of learning opportunities and evaluation priorities

Sub-activity/evaluation option	Cost (in $)	Mathematica's assessment of associated opportunities for learning and justification for prioritization	Priority category
Water management and irrigation (104M)			
Di perimeter ERR (including Sourou O&M)	89.0M	The benefit stream for the Di perimeter is the value of the additional production with the irrigated perimeter relative to the value of production without construction of the perimeter. To provide a rigorous ERR analysis, the evaluation design will need to include all beneficiaries, investigate whether prices changed, and investigate the possible effects of lower production intensity outside the perimeter.	**High**
Di Lottery		The evaluation that assesses the impact of getting access to irrigated land through the Di Lottery relies on the most rigorous design possible, an RCT. At a cost in U.S. dollars of about $39,700 (without compact administration costs) or $45,000 (with compact administration costs) per hectare, the benefit of winning the lottery ranks as one of the largest stakes in any RCT that we know of. It is also rare to find an RCT of an infrastructure project. The Di Lottery beneficiaries are the Di beneficiaries whose net benefit from the perimeter is largest, as they did not receive their land as compensation.	**High**
Sourou O&M (included in Di perimeter)	6.6M	A key question for the sustainability of the Di perimeter is whether the irrigation infrastructure is maintained. The Sourou O&M evaluation assesses whether the institutions tasked with the maintenance of the irrigation infrastructure (WUAs and AMVS) will be able to do so.	**High**
IWRM	5.0M	There is little evidence on how water management plans are implemented. As of 2012, the African Ministers' Council on Water noted that very few sub-Saharan countries had put IWRM plans into practice (AMCOW 2012). As such, an investigation of the effectiveness of the IWRM activities financed by MCC in two regions of Burkina Faso could be informative for IWRM implementation in other countries.	**Medium-High**
Non-PAPs from disadvantaged villages, women and youth	NA	Because they received about 20 percent of the land on the Di perimeter, it is important to assess whether the proportion of land planted, the crops grown, the cropping intensity and yields achieved by the non-PAPs from disadvantaged villages, women, and youth are the same, higher, or lower than the PAPs and the Di Lottery beneficiaries.	**High**
Analysis of crop prices	NA	Knowledge of the effect of additional production on the Di perimeter on local agricultural prices is important for the calculation of the ERR. If the provision of a large amount of irrigated land and the associated additional production of agricultural goods led to a decline in crop prices, this would reduce the ERR. We note that we will be able to compare the evolution of prices at Di relative to markets elsewhere in Burkina Faso.	**High**
Land productivity outside the perimeter	NA	Di beneficiaries may reassign labor that was previously used on land outside of the perimeter to the irrigated land on the perimeter. This may lead to lower productivity of the land outside the perimeter. Hence, not measuring the productivity of the land outside the perimeter could understate the benefits of access to the land inside the perimeter. We will be able to provide rigorous evidence on land reallocation for the Di Lottery beneficiaries.	**High**

Table B.2. *(continued)*

Sub-activity/evaluation option	Cost (in $)	Mathematica's assessment of associated opportunities for learning and justification for prioritization	Priority category
Qualitative analysis of customary land rights outside the perimeter	NA	Conducting much more in-depth qualitative analysis of customary land rights and beneficiaries' change in land rights due to winning the lottery would turn the current design into an exciting mixed-methods study. This would make a real contribution to the literature on customary land rights. The current research design, however, already plans for quantitative analysis of information on customary land rights gathered through the questionnaires administered to the Di Lottery beneficiaries and the control group (at little additional cost).	**Low-Medium**
Regression discontinuity design	NA	The beneficiary selection process allows for an analysis of receiving access to land on the Di perimeter by using both the RCT and RDD frameworks. The comparison of the estimated impacts is a unique opportunity to contribute to the literature on the validity of RDD in general, and more specifically within the context of the infrastructure and agriculture sectors in developing countries. A recent meta-analysis of studies that compared RCT and RDD estimates did not list a single study in these contexts (Chaplin et al. 2017. Mathematica's proposed design suggests analyzing the plausibility of the RDD assumptions as part of the baseline data analysis in Option Period I, before deciding whether to collect data on the RDD control group.	**High**
Diversified agriculture (30M)			
Farmer training	NA	Judging by available information on the implementing contractors' level of effort, the farmer training sub-activity is the DA sub-activity with the largest expenditure share (with the possible exception of the animal health sub-activity). Because baseline data are available, our research design for the performance evaluation also incorporates a pre-post analysis. The pre-post analysis will provide evidence for one of MCC's key research questions, which is whether farmers adopted the new agricultural technologies.	**High**
MIS	NA	The MIS performance evaluation provides an opportunity to analyze the sustainability of an MIS system in Africa and to understand how easily such a system could be set up to benefit farmers in two project areas within a country. As noted in Mathematica's memo to MCC on October 31, 2016, the Dutch development agency is providing subsequent funding to the MIS operator. The evaluation will not be able to disentangle the two sources of funding so that the evaluation will not be able to provide information directly on MCC's investment.	**Medium**
Animal health and husbandry	NA	Because the sample of beneficiaries in the existing baseline data is too small and is not representative of all animal health and husbandry beneficiaries, the performance evaluation will comprise an implementation analysis as well as a descriptive analysis on beneficiaries.	**Medium**
Rural markets	NA	The rural markets sub-activity primarily rehabilitated four markets and trained market management committees in the rehabilitated markets and in five additional markets. The proposed performance evaluation will be able to provide learning on the determinants of the functioning of rural markets from only these experiences, of which only four saw construction.	**Low**

Table B.2. *(continued)*

Sub-activity/evaluation option	Cost (in $)	Mathematica's assessment of associated opportunities for learning and justification for prioritization	Priority category
Value-added	NA	The value-added sub-activity can be evaluated as a performance evaluation. Because no baseline data are available, the performance evaluation of the value-added sub-activity is limited to an implementation analysis as well as a qualitative assessment of its benefits to beneficiaries. As a result, we see little opportunity for learning. Evaluating the value-added sub-activity is primarily useful if MCC anticipates implementing training on rice value-added technologies as part of future compacts. We note that the value-added sub-activity was initially part of a planned value-chain sub-activity that was meant to be heavily integrated with the other activities (see below).	**Low**
Evaluation option for the ADP project as a whole			
Integrated program logic		MCC developed an integrated set of activities for the ADP based on evidence that farmers in the two project areas faced multiple constraints. MCC thought that these constraints needed to be addressed simultaneously to maximize the value of investments. This situation can arise in other MCC projects. The lessons learned on why the integrated program logic of the ADP was not followed may be applicable to future MCC projects in which project activities are meant to complement each other.	**Medium-High**

APPENDIX C

RESEARCH QUESTIONS

This page has been left blank for double-sided copying.

Table C.1. List of research questions, the data-collection method and the research method

		Research method		Data collection method				
		Performance evaluation	Impact evaluation	Administrative	Qualitative	Quantitative	Site visit	Link to logic model
Project integration								
RQ1	To what extent were the various project components implemented in a cohesive way, i.e., in which the components complemented each other, as anticipated in the original program logic?	X		X				Yes
RQ1a	Are the rural markets and the MIS functioning and being used by farmers who benefitted from technical assistance or received land in the Di perimeter, as anticipated by the integrated program logic?	X			X	X	X	Yes
RQ2	If the program was not implemented as a cohesive project, according to the original logic, then why not?	X		X	X			Yes
Di perimeter evaluation								
RQ1	How were the Di perimeter construction and associated activities implemented relative to the original plans?	X		X				Yes
RQ2	What is the total area planted, average yield/hectare, total production and total profit on the Di perimeter for each of the focus crops: rice, corn, onions, tomatoes, soybeans, and cowpeas?	X				X		Yes
RQ2a	Have prices for these crops changed since the completion of the perimeter?	X				X		Investigates potential unintended side effect
RQ2b	Are agricultural outcomes different for Di Lottery beneficiaries and Di PAPs? If so, why?	X			X	X		No
RQ3	What is the economic rate of return of the Di perimeter?	X		X		X		Yes
RQ4	How has PAP well-being changed? Have any PAPs been harmed (socially, economically, or politically) by the intervention? How?	X			X	X		Investigates potential unintended side effect

Table C.1. *(continued)*

		Research method		Data collection method				Link to logic model
		Performance evaluation	Impact evaluation	Administrative	Qualitative	Quantitative	Site visit	
RQ5	Have PAPs received the compensation instruments (titles and/or leases and/or financial compensation) they were informed they would receive? Why or why not?	X		X	X	X		Yes
RQ6	What are the PAPs' perceptions of the process by which compensation was determined and provided? What are the PAPs' perceptions of the compensation provided?	X			X			No
RQ7	How has the PAPs' perception of land tenure security changed?	X			X			Yes
RQ7a	Have any PAPs been involved in a land conflict on the perimeter?	X				X		Yes
RQ8	What type of land investments do PAPs' make? Have PAPs rented or sold land from the Di perimeter? Have PAPs used land from the Di perimeter as collateral for credit?	X				X		Yes
Di Lottery								
RQ1	To what extent did Di Lottery beneficiaries receive all benefits they were meant to receive (formal lease documents, training in agricultural technologies, starter kits)?	X				X		Yes
RQ2	What impact does winning the Di Lottery have on agricultural practices, production, total agricultural income, and overall household income of the Di Lottery beneficiaries?		X			X		Yes
RQ3	What are the impacts of winning the Di Lottery on land tenure security?		X			X		Yes
RQ3a	Have Di Lottery beneficiaries been involved in a land conflict on or off the perimeter?		X			X		Yes
RQ4	To what extent are the estimated impacts from the RD similar to those from the RCT?		X			X		No

Table C.1. *(continued)*

		Research method		Data collection method				Link to logic model
		Performance evaluation	Impact evaluation	Administrative	Qualitative	Quantitative	Site visit	
RQ5	To what extent can methods that use the discontinuity to estimate impacts away from the threshold recover the average treatment effect of the Di Lottery?		X			X		No
Sourou O&M								
RQ1	How were the O&M activities implemented relative to the original plans?	X		X	X			Yes
RQ2	To what extent are the Di perimeter and the old perimeters at Niassan effectively operated and maintained? Are their levels of operation and maintenance sustainable?	X		X	X		X	Yes
RQ2a	To what extent do WUAs on these perimeters have the capacity (financial, technical, and organizational) to fully leverage and maintain the irrigation infrastructure?	X		X	X			Yes
RQ2b	What are the factors influencing the WUAs' level of capacity/implementation?	X			X			Yes
RQ3	Has the Government of Burkina Faso continued to implement AMVS reform strategic plan that was developed during the compact? If so, to what extent? If not, why not?	X			X			Yes
RQ4	To what extent is AMVS successfully fulfilling its O&M responsibilities? What are the reasons for its success or lack thereof?	X			X		X	Yes
RQ5	To what extent is the Centre d'Appui Technique et Gestion (CATG) operational? What are the reasons for its success or lack thereof?	X			X			No
RQ5a	If CATG is operational, what percentage of WUAs are benefitting from the services, and what services are they accessing?	X		X	X			No

Table C.1. *(continued)*

		Research method		Data collection method				
		Performance evaluation	Impact evaluation	Administrative	Qualitative	Quantitative	Site visit	Link to logic model
RQ5b	What benefits do the WUAs perceive to using CATG? What are WUA perceptions of the quality of CATG services? What specific CATG services do WUAs think are most beneficial? If the percentage of WUAs using CATG is low, why are so few using it?	X			X			No
RQ5c	Is the support that CATG is providing to WUAs financially sustainable?	X		X				No
IWRM								
RQ1	How were the IWRM activities implemented relative to the original plans under the ADP?	X		X	X			Yes
RQ2	Have the SDAGEs been implemented as planned? What are the primary factors influencing their implementation?	X		X	X			Yes
RQ2a	Have activities that were expected to be conducted under the SDAGEs been implemented?	X		X	X			Yes
RQ2b	a.b. What are the perceived benefits of the SDAGEs for water users?	X			X			Yes
RQ3	How well are the CLE and basin committee institutions functioning? What are the primary factors influencing their operation?	X		X	X		X	Yes
RQ3a	Have activities that were expected to be conducted by the CLE and basin committee institutions been implemented?	X		X	X		X	Yes
RQ3b	What are the perceived benefits of the CLE and basin committee institutions for water users?	X			X			Yes
RQ4	Are the water user/polluter fees (CFE) fully defined, and to what extent are they being collected? Are the funds from these fees being directed to the CLEs and the basin committees or to the national treasury?	X		X	X			Yes
RQ5	What are the effects of IWRM on (a) water resources and (b) water conflicts?	X			X			Yes

Table C.1. *(continued)*

		Research method		Data collection method				Link to logic model
		Performance evaluation	Impact evaluation	Administrative	Qualitative	Quantitative	Site visit	
Farmer training								
RQ1	How was the farmer training sub-activity implemented relative to the plans for this sub-activity?	X		X	X			Yes
RQ2	To what extent have farmers adopted or adapted the improved production practices proposed by the project?	X				X	X	Yes
RQ2a	If farmers are adopting improved farming practices, which ones have been adopted the most and the least, and why?	X			X	X		Yes
RQ2b	If farmers are adapting improved practices, which ones have been modified the most and the least, and why?	X			X	X	X	Yes
RQ2c	Have farmers continued to invest in improved seeds/fertilizers?	X				X		Yes
RQ3	Have participating farmers used the incentive kits that they received along with the training?	X				X		Yes
RQ4	Do participating farmers diversify crop production more than they did before the project?	X				X		Yes
RQ5	What is the total area planted, average yield/hectare, total production, and total profit for each of the focus crops: rice, corn, onions, tomatoes, soybeans, and cowpeas?	X				X		Yes
RQ6	Have the participating farmers' average yields/hectare increased, decreased, or remained the same for each of the focus crops, compared with the average yields/hectare before the project?	X				X		Yes
RQ7	6.7. Have the participating farmers' overall agricultural incomes and profits increased, decreased, or remained the same compared with their incomes and profits before the project?	X				X		Yes

This page has been left blank for double-sided copying.

APPENDIX D

DI PERIMETER ERR

This page has been left blank for double-sided copying.

MCC uses ERR models to assess whether its projects are sound investments. The ERR is a summary statistic that reflects the economic merits of an investment. Conceptually, it is the discount rate at which the benefits of an intervention are exactly equal to its costs; a higher ERR implies relatively higher benefits and lower costs.

MCC finalized the closeout ERR for the Di perimeter on March 7, 2017 (MCC 2017). MCC's ERR analysis computes the increase in agricultural profit for the land on which the Di perimeter was built.

MCC's calculations are based, in part, on realized agricultural outcomes that APD collects as part of the post-compact monitoring activities (see MCA-BF 2014c for the post-compact M&E plan). This includes information on the area planted and agricultural yields for the primary crops grown on irrigated land at Di—that is, corn, rice, cowpeas, onions, and tomatoes. Soya is used as a proxy for any other crops. Total production for a crop is calculated as the area planted with a crop multiplied by its average yield.

To estimate the value of this agricultural production, the ERR calculation makes assumptions on post-harvest losses and prices. Crop prices are assumed to be fixed across time and across season, while losses vary by season but are fixed across time. Agricultural profits subtract the cost of inputs from the value of total production. The main costs are (1) labor costs for land preparation, weeding, and harvesting; (2) the costs of fertilizer, seeds, and pesticides; (3) post-harvest and marketing costs; and (4) contributions to the WUAs. The calculation assumes that the amounts of inputs used differ across crops and dry and rainy seasons, but that input prices are constant across years and seasons.

The value of production without the perimeter is based on a similar calculation using information from the pre-compact period on the area planted by crop, the quantities produced, the inputs used, and the prices for inputs and crops.

Regarding program costs, MCC takes into account direct costs—such as costs associated with construction of the perimeter itself—and indirect costs, such as costs of design and supervision, costs related to environmental and social mitigation plans and a share of compact administration and M&E costs. These costs do not include costs spent by the post-compact entity APD after the close of the compact.[48]

Total costs per hectare amount to $39,731 U.S. dollars when compact administration costs are excluded and $45,088 U.S. dollars when they are included.[49]

[48] As not all Di beneficiaries were trained by the end of the compact, GOBF committed to funding the training that occurred during the post-compact period. In addition, GOBF also provided subsidies to CATG during a transition period. The inclusion of these costs would likely not change the overall cost of the perimeter nor the ERR substantially.

[49] We calculated the per hectare value based on the assumption that the total land area at Di perimeter comprises 2,240 hectares of land (MCA-BF 2014b). The Atlas of Realizations (MCA-BF 2014d) and the ERR calculations use a value of 2,246 hectares. We note that the Atlas of Achievements computes a cost per hectare of $37,554 U.S. dollars, using the same total cost of the perimeter. This is a clerical error.

To estimate future agricultural profits, MCC uses the values for agricultural production from 2016, assumes prices for inputs and agricultural production will remain stable, and that long-term land productivity will be about 90 percent of the 2016 value. MCC calculates that agricultural profit would increase from about 242,425 CFA to a long-term value of 1,974,184 CFA per hectare per year, resulting in an increase of 1,864,227 (about $3,000 U.S. dollars) per hectare per year.[50] The period of analysis for the ERR is 25 years; the ERR (including compact administration costs) was estimated to be 5.5 percent.

We propose to recalculate the ERR using the information on area planted, input use, input prices, agricultural production, production sales prices, and profits that we collect as part of the Di perimeter evaluation. As we describe in the chapter on data collection, we propose two rounds of data collection—interim and final quantitative data collection. We will make use of information from both rounds, as well as the crop cutting measurements, to update the ERR calculations.

As noted in Chapter IV, Section C, we also analyze whether the additional production on the Di perimeter has led to lower crop prices. If this is the case, we cannot multiply pre-compact production from land used by the perimeter with post-compact prices. We will instead multiply pre-compact production with prices that are predicted based on our relationship of the prices at Di and other markets in Burkina Faso.[51]

Our information on agricultural profits in the counterfactual scenario (that is, the scenario without the construction of the Di perimeter) was based on limited information on agricultural outcomes on the land used for the Di perimeter. We will therefore conduct a sensitivity analysis around these counterfactual profits. This may include, for example, calculating the ERR with counterfactual profits that are 50 percent higher and lower than in MCC's published ERR.

We will also provide a qualitative assessment of the state of the irrigation infrastructure by conducting a site visit to the Di perimeter. We will reference available information on soil fertility from the Bureau National des Sols at this site that might indicate whether the land productivity is declining as anticipated and whether recommended amounts of organic fertilizer are used to maintain productivity. Together with information on maintenance of primary and secondary canals, this will provide us with a plausible ranges for the life-span of the irrigation infrastructure and the evolution of land productivity on the perimeter. We will also address these questions to AMVS. We will include this information into the ERR, and conduct sensitivity analyses around this assumed lifespan and evolution of productivity.

[50] We calculated these values by dividing total profits by 2,240 hectares.

[51] To the extent that nearby markets are partially affected by the additional production at Di, we will only be able to partially address the effect of Di on lower prices.

APPENDIX E

DI LOTTERY SCORING SHEET

This page has been left blank for double-sided copying.

Table E.1. Di Lottery scoring sheet

Criteria		Points	Maximum in category
Documented number of adults or adolescents age 15 and older available to work on the land, in addition to applicant			20
If first choice is to receive plot for growing rice (2 hectares)			
	At least 4 per hectare (8 total)	20	
	At least 3 per hectare (6 total)	15	
	At least 2 per hectare (4 total)	10	
	Fewer than 2 per hectare	0	
If first choice is to receive polyculture plot (1 hectare)			
	At least 6 per hectare	20	
	At least 5 per hectare	15	
	At least 4 per hectare	10	
	Fewer than 4 per hectare	0	
Ownership of agricultural tools and draft animals			10
	None	0	
	Animal-drawn cart	5	
	Animal-drawn cart and plow	10	
Technical trainings on agricultural production attended by the applicant			5
	None	0	
	Attended at least one	5	
Applicant's technical experience in irrigated agriculture			15
	None	5	
	Less than 2 years	10	
	More than 2 years	15	
Gender			5
	Female	5	
	Male	0	
Age			5
	Between 18 and 30	5	
	Between 31 and 55	3	
	Age 56 and older	1	
Level of debt			10
	No arrears	10	
	Arrears less than or equal to 100,000 CFA	6	
	Arrears of more than 100,000 CFA	0	
Current residence			15
	Village in the rural commune of Di	15	
	Sourou Province	10	
	Mouhoun Region	5	
	Rest of the country	0	
Has a title to a plot in another AMVS perimeter			15
	Yes, at least one	0	
	No	15	
Total/maximum			100

This page has been left blank for double-sided copying.

APPENDIX F

BALANCE TESTS

This page has been left blank for double-sided copying.

Table F.1. Balance tests for scoring variables and survey baseline variables

Baseline measure	Treatment group mean	Control group mean	Difference	*p*-value of difference
Eligibility criteria				
Number of active household members	4.07	4.24	-0.15	0.04**
Applicant owns one piece of agricultural equipment	0.15	0.12	0.03	0.15
Applicant owns at least two pieces of agricultural equipment	0.74	0.75	-0.01	0.80
Applicant received technical training in agriculture	0.41	0.39	0.01	0.62
Applicant has no experience in irrigated agriculture	0.25	0.28	-0.02	0.35
Applicant has less than two years of experience in irrigated agriculture	0.05	0.07	-0.02	0.14
Applicant has two years or more of experience in irrigated agriculture	0.70	0.65	0.04	0.10
Female	0.22	0.19	0.04	0.09*
Age of applicant—18 to 30	0.40	0.43	-0.04	0.12
Age of applicant—31 to 55	0.56	0.53	0.04	0.17
Age of applicant—56 or older	0.04	0.04	0.01	0.56
Applicant has debt	0.01	0.01	0.00	0.87
Applicant is from village in the rural Di commune	0.56	0.54	0.01	0.85
Applicant is from Sourou province	0.93	0.94	-0.02	0.23
Applicant is from Boucle du Mouhoun region	0.01	0.01	0.01	0.11
Applicant does not have title to a parcel on AMVS perimeters	0.99	0.99	0.00	0.56
Information on agricultural equipment used in eligibility criteria				
Applicant owns an animal-drawn cart	0.80	0.81	-0.01	0.78
Applicant owns a plow	0.80	0.79	0.01	0.72
Applicant owns a cultivator	0.00	0.01	-0.01	0.20
Applicant owns a tractor	0.01	0.01	0.00	0.59
Information on technical experience used in eligibility criteria				
Applicant has technical experience in irrigated agriculture	0.75	0.72	0.02	0.35
Survey data				
Household land rights				
Household has land rights over plots (formal or informal)	0.69	0.60	0.09	0.00***
Household has right of access to community land	0.41	0.40	0.01	0.69
Household has right of access to community land only	0.30	0.25	0.05	0.12
Household ownership and usage of land				
Total plots household owns, has communal access to, or rents	2.85	2.61	0.24	0.08*
Number of plots owned by household	1.71	1.46	0.24	0.04**
Number of plots household has communal access to	0.77	0.66	0.11	0.15
Number of plots rented by household	0.37	0.48	-0.12	0.04**

Table F.1. *(continued)*

Baseline measure	Treatment group mean	Control group mean	Difference	*p*-value of difference
Applicant ownership and usage of land				
Number of cultivable plots owned	0.94	0.91	0.04	0.68
Number of hectares cultivated in the last 12 months	3.98	3.83	0.18	0.25
Number of plots of cultivable land rented from others	0.36	0.46	-0.11	0.03**
Number of plots of cultivable land rented to others	0.00	0.00	0.00	0.94
Total plots applicant owns, rents, or communally operates	1.66	1.73	-0.07	0.48
Number of communal plots operated by applicant	0.35	0.36	0.00	0.99
Household size				
Number of household members 15 years or older supporting ag production—less than 2	0.04	0.05	0.00	0.94
Number of household members 15 years or older supporting ag production—2 to 10	0.87	0.86	0.00	0.81
Number of household members 15 years or older supporting ag production—11 to 15	0.07	0.07	0.00	0.77
Number of household members 15 years or older supporting ag production—more than 16	0.02	0.02	0.00	0.96
Applicant profession				
Profession of lottery candidate—farmer	0.86	0.86	-0.01	0.78
Profession of lottery candidate—craft	0.02	0.01	0.00	0.62
Profession of lottery candidate—commerce	0.03	0.03	0.00	0.82
Profession of lottery candidate—laborer	0.01	0.01	0.01	0.12
Profession of lottery candidate—salaried employee	0.06	0.08	-0.01	0.41
Profession of lottery candidate— unemployed	0.01	0.01	0.00	0.45
Applicant level of training				
Level of technical training in agricultural production—none	0.58	0.56	0.02	0.51
Level of technical training in agricultural production—only one	0.18	0.19	-0.01	0.59
Level of technical training in agricultural production—more than one	0.23	0.23	0.00	0.93
Household agricultural inputs				
Agricultural inputs used—traditional seed	0.83	0.82	0.01	0.64
Agricultural inputs used—enhanced seed	0.52	0.50	0.02	0.41
Agricultural inputs used—fertilizer	0.74	0.72	0.02	0.53
Agricultural inputs used—herbicide	0.69	0.65	0.03	0.20
Agricultural inputs used—pesticide	0.66	0.64	0.01	0.60
Agricultural inputs used—compost	0.62	0.62	0.00	0.88
Agricultural inputs used—manure	0.75	0.74	0.01	0.78
Agricultural inputs used—other input	0.02	0.02	0.00	0.71

Table F.1. *(continued)*

Baseline measure	Treatment group mean	Control group mean	Difference	*p*-value of difference
Household agricultural assets				
Agricultural equipment owned—plow	0.78	0.76	0.02	0.49
Agricultural equipment owned—cart	0.74	0.74	0.00	0.93
Agricultural equipment owned—motor pump	0.09	0.09	-0.01	0.74
Agricultural equipment owned—tractor	0.02	0.02	0.00	0.54
Agricultural equipment owned— electronic equipment	0.46	0.45	0.01	0.77
Agricultural equipment owned— wheelbarrow	0.30	0.31	-0.02	0.49
Farm animals owned—traction bovine	0.71	0.71	0.00	0.92
Farm animals owned—other bovine	0.32	0.33	-0.02	0.56
Farm animals owned—traction donkey	0.62	0.62	0.00	0.89
Farm animals owned—other asinus	0.21	0.23	-0.01	0.53
Farm animals owned—traction horses	0.02	0.01	0.01	0.22
Farm animals owned—other equidae	0.01	0.01	0.00	0.72
Farm animals owned—ovis	0.52	0.54	-0.02	0.44
Farm animals owned—caprinae	0.49	0.46	0.02	0.40
Farm animals owned—suidae	0.11	0.09	0.02	0.14
Farm animals owned—chicken	0.91	0.89	0.03	0.15
Farm animals owned—guinea fowl	0.19	0.19	0.00	0.87
Farm animals owned—other poultry	0.09	0.10	0.00	0.83
Farm animals owned—other animals	0.09	0.10	0.00	0.83
House type				
Type of house walls—mudbrick	0.87	0.87	-0.01	0.77
Type of house walls—concrete	0.01	0.02	-0.01	0.23
Type of house walls—brick	0.12	0.11	0.01	0.43
Type of house roof—mudbrick	0.13	0.16	-0.03	0.13
Type of house roof—straw or other plant-based material	0.04	0.02	0.01	0.18
Type of house roof—metal sheets	0.84	0.82	0.02	0.40
Income and debt				
Income source—production sale in rainy season (unconditional)	190,581	175,375	8,637	0.78
Income source—production sale in dry season (unconditional)	352,259	478,100	-141,462	0.11
Income source—trade (unconditional)	159,888	191,277	-38,467	0.63
Income source—animal sale (unconditional)	107,932	98,338	14,624	0.47
Income source—paid labor (unconditional)	178,009	242,643	-64,491	0.37
Income source—other (unconditional)	68,896	60,717	10,497	0.58
Income source—production sale in rainy season (conditional)	309,282	317,266	-12,649	0.81
Income source—production sale in dry season (conditional)	635,420	805,530	-181,286	0.23
Income source—trade (conditional)	407,541	521,516	-140,029	0.51

Table F.1. *(continued)*

Baseline measure	Treatment group mean	Control group mean	Difference	*p*-value of difference
Income source—animal sale (conditional)	212,689	192,452	26,262	0.49
Income source—paid labor (conditional)	627,717	936,329	-330,572	0.20
Income source—other (conditional)	347,441	309,075	54,147	0.55
Total income	1,057,564	1,246,449	-210,662	0.26
Current household debt—no debt	0.89	0.87	0.02	0.28
Current household debt—less than or equal to FCFA 100,000	0.04	0.04	0.00	0.72
Current household debt—higher than FCFA 100,000	0.07	0.09	-0.02	0.12

Source: Di Lottery baseline survey data

*Significantly different from zero at the .1 level, two-tailed test.

**Significantly different from zero at the .05 level, two-tailed test.

***Significantly different from zero at the .01 level, two-tailed test.

APPENDIX G

ASSESSMENT OF MATCHED-COMPARISON GROUP DESIGN

This page has been left blank for double-sided copying.

A. Assessment of matched comparison group difference-in-differences design

This appendix describes our assessment of the matched comparison group difference-in-differences design that the previous evaluators planned to use to evaluate farmer training, value chain development, and animal health services (IMPAQ International 2014a). This was based on matching households from intervention areas to households from comparison areas. This methodology could provide a credible analysis of the impact of farmer training activities.

However, a detailed review of the previous evaluation documents as well as the implementation documents raises three major concerns about this design's ability to detect unbiased impacts: (1) intervention and comparison groups that differ significantly from each other, (2) the highly clustered location of intervention communities, and (3) the lower than expected take-up of farmer training. We discuss these issues in the following sections.

1. Intervention and comparison groups differ significantly

For the evaluation's matching procedure, the previous evaluator used data collected during a rapid household listing stage. However, the information presented in the report on data quality suggests that the matching resulted in intervention and comparison groups that were markedly different from each other (IMPAQ International 2014b). For example, in the intervention group there were almost twice as many parcels per household devoted to maize in the rainy season than there were in the comparison group. We computed the number of variables for which a test between the intervention and comparison groups was statistically significant. Of the 48 variables tested, there were significant differences between the intervention and comparison groups for 20 variables ($p < 0.05$). If the matching had led to balanced intervention and comparison groups, we would expect a number between 2 and 3. We concluded that the matching procedure failed to create comparable intervention and comparison groups and therefore any analysis would lead to biased impact estimates.

When intervention and comparison groups are unbalanced in this way, it may be possible to improve the matches by using another data source, such as the detailed baseline data collected for the matched sample. However, if the chosen villages and households also differ systematically along key unobservable characteristics, then any impact estimates from the re-matched samples would also be biased. Moreover, a downside of re-matching is that it reduces the power of the design. Because the two issues we cover next also suggest that the study is highly underpowered, it is not plausible that re-matching would lead to intervention and comparison groups that would allow for a credible impact analysis.

2. Highly clustered intervention communities

The nine intervention villages in the Comoé Basin are clustered within 40 square kilometers—about 15 square miles (MCA-BF 2014a) of each other—whereas all intervention villages in the Sourou Valley are close to the old perimeters. The old perimeters in turn are contiguously located along the Sourou Valley.[52] Because incomes in contiguous communities are

[52] The final report by the contractor who implemented the agricultural training activities notes that some neighboring areas outside the irrigated perimeters were also covered for the agro-sylvo-pastoral activities, with a

typically highly correlated due to common local factors, such as weather and planting conditions, it would be difficult to separate out these local factors from the effects of treatment in the two geographic intervention zones. As a result, the statistical analysis should group together villages that are located close to each other within the two intervention zones. During the design phase, we will conduct two site trips to assess whether any nearby villages need to be considered as one cluster and, if so, which ones.

3. Lower than expected take-up of farmer training

The initial power calculations assumed that all 1,082 respondents in the intervention areas would participate in the training activities, but only about 60 percent of the households in the intervention group had at least one household member participate in any training (IMPAQ International 2014a). Because we would only expect to observe program impacts in households that actually attended trainings, the statistical power of the existing design is thus significantly reduced; there is now a smaller treated sample in which to detect impacts.

4. Statistical power

Table E.1 presents MDIs for the evaluation. We calculated different MDIs to understand how strongly the MDI is affected by the imbalance between intervention and comparison groups and the degree of clustering in the intervention villages. The analysis is based on the sample size of 60 percent of farmers who participated in the training. The table has three horizontal panels and four columns. The columns show the effect of imbalance between treatment and control groups on the MDI. The panels show the effect on the MDI of different assumptions about how geographically clustered the intervention areas are.

In particular, the first panel is based on the assumption that all intervention villages can be considered separate units of analysis. The second and third panel are based on the more realistic case that some intervention villages will need to be considered jointly, reducing power. For Panels 2 and 3, we compute MDIs under the assumption that we need to consider the 30 intervention villages to correspond to 15 and 10 clusters, respectively.

Even when we do not account for the extent of imbalance or the clustered location of the intervention villages, the MDI for agricultural income is high—49 percent of comparison group incomes (Table E.1). As the extent of imbalance increases, the MDIs also increase. Because intervention villages are considered as fewer clusters, the MDIs also increase. Accounting for either of the two issues or both leads to MDIs that are implausibly high, given the limitation in the number of farmers in the sample who participated in the training.

total area of 590 square kilometers (MCA-BF 2014a). The area for training in irrigated agriculture is smaller, with intervention villages bordering on each other.

Table G.1. MDIs on agricultural income: Intervention and comparison group balance levels in the farmer training evaluation

MDIs with varying levels of balance between intervention and comparison groups				
Extent of balance between intervention and comparison groups	**100%**	**90%**	**80%**	**70%**
Panel 1: 30 intervention clusters				
MDI (2011 CFA)	234,102	246,766	261,734	279,806
MDI (% of the mean)	49.02%	51.67%	54.80%	58.59%
Panel 2: 15 intervention clusters				
MDI (2011 CFA)	271,032	285,693	303,023	323,945
MDI (% of the mean)	56.75%	59.82%	63.45%	67.83%
Panel 3: 10 intervention clusters				
MDI (2011 CFA)	321,104	338,473	359,005	383,793
MDI (% of the mean)	67.24%	70.87%	75.17%	80.36%

Note: Calculations assume a two-tailed test with a 95 percent confidence level and 80 percent statistical power. The baseline follow-up correlation in income is assumed to be 0.3, the attrition rate between baseline and follow-up is 15 percent, and the nonresponse rate is 10 percent. Data on household income come from the IMPAQ design report (IMPAQ International 2014a). The standard deviation for the intervention group uses data from the Boucle de Mouhoun and Cascades regions, whereas the standard deviation for the comparison group uses data from all regions (IMPAQ International 2014a). The extent of imbalance is formally the R-squared in an ordinary least squares regression of the intervention on covariates. Clustering is done as described in the text. CFA is the West African CFA franc, the currency in Burkina Faso.

This page has been left blank for double-sided copying.

This page has been left blank for double-sided copying.

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and data collection

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ TUCSON, AZ ■ WASHINGTON, DC ■ WOODLAWN, MD

MATHEMATICA
Policy Research