MILLENNIUM
CHALLENGE CORPORATION
UNITED STATES OF AMERICA

Impact and Performance Evaluation of the Rehabilitation and Intensification of Rain-fed Olive Plantations of the MCA-Morocco.

DESIGN REPORT

This page was intentionally left blank

Authors and Acknowledgement

This report was prepared by NORC at the University of Chicago for the Agency of Partnership for Progress (APP/MCA-Morocco) under the ME-2 contract. This report presents the design of the impact and performance evaluation project of the Rehabilitation and Intensification of Rain-fed Olive Plantations of the MCA-Morocco. The primary authors of this report are Safaa Amer, Sarah Hughes, Mawadda Damon, Yvonne Cao, George Caldwell, Majid Benabdellah, Najib Benjelloun, Abderrahim Bentaibi and Mohammed Dehhaoui.

George Caldwell was the main author of the statistical design for this evaluation, with assistance from Safaa Amer, Fritz Scheuren and Jeff Telgarsky. Other collaborators include Abdel Wahab Chaoui and Abdellah Omari, who also supplied a significant contribution to the design of the project.

The authors are responsible for all errors. Please contact Sarah Hughes, NORC at the University of Chicago, at: hughes-sarah@norc.org. regarding any comment or question,

Table of Contents

Authors and Acknowledgement 3
I. Summary 5
II. Introduction 8
III. Context of the Intervention 10
A. Characterization of rural poverty in Morocco 10
B. The olive oil sector in Morocco 10
C. Rural young people and women: what role do they play in the olive oil program? 14
D. Farmer Associations (OPAs) 15
E. The intervention 17
IV. The Intervention: Intervention Logic and Indicators 18
V. Performance Evaluation 21
A. Objectives of the performance evaluation 21
B. Methodology 22
C. Results 24
VI. Impact Evaluation 25
A. Context 25
B. Randomization 25
C. Suggestions for improvements 26
D. Impact Variables 27
E. Impact Evaluation Analysis 28
F. Sample plan and sample size 29
Collection of Data 33
A. Design of survey tools 33
B. Ground work 37
C. Ethical considerations for the ground work 43
D. Data Management 44
VII. Calculation of the Indicators 46
VIII. Awareness raising and reinforcement of Capacities 48
IX. Bibliography 49
X. Annexes 52
A. Annex A: Organizational chart - Composition of the Teams 52
52
B. Annex B: Action Plan 54
C. Annex C: Logical Framework and Indicators 55
D. Annex D: Professional Ethics Code and Confidentiality of the data 56
E. Annex F: Analysis of statistical power of a sample 59
F. Annex E: Statistics Glossary 71
G. Annex G: Questionnaires 72

I. Summary

Context

The purpose of the MCA Morocco compact is to reduce poverty by economic growth, increasing productivity and employment in high potential sectors (MCC Morocco compact 2007). To stimulate growth in the farming sector, the MCA has created the Fruit Tree Productivity Project (PAF) the purpose of which is to encourage the transition by farmers from annual cereal crops, that are vulnerable, to more productive tree products, such as olives. By investing in an intensive training and technical assistance project in the olive oil sector, it is anticipated that the rehabilitation and intensification of olive trees will increase net farm revenue by 8.4% from now until the end of the third year of the project's implementation. Farm revenue from crops (not including livestock) should increase by 12.7% and income from the production of olives by 30.8%. It has also been anticipated that olive tree productivity will increase by 30.8% by virtue of the hypotheses of stability of actual prices. As production and crop prices go up, the project anticipates applying indirect benefits to key players through the program of values of olive production (MCC-Morocco Compact, 2007).

Evaluation questions

NORC has been hired to conduct an evaluation of the "Rehabilitation and Intensification of Rain-fed Olive Plantations" project, that includes an evaluation of the implementation of the intervention ("performance evaluation") and also an evaluation of the effects of the intervention ("impact evaluation"). NORC's performance evaluation will determine whether the implementation of the project meets the established objectives. The design of the NORC's impact evaluation is based on the measurement of the net change in farm income of olive producer households that can be attributed to the intervention. NORC will therefore seek to respond to the main following research questions defined by the APP:

- What is the impact of the component of the rehabilitation of olive plantations on the income of farmers in the target rain zones?
- Has the production of olives improved in volume and in value?
- Have techniques taught by the project been adopted by farmers?
- Has there been an improvement in the quality of the oil olive oil produced?
- Have Farmer Associations played their roles in the development of the olive oil program?
- Has knowledge generated by the project spreads through the areas in which the project was implemented?
- What are the impact differences from activities of the project according to gender, age and income?
- What are the expected impacts of the rehabilitation?

The theory of the change used in this intervention model is the following: training farmers to use improved production and transformation techniques of olives will bring about an increase in the volume and value of their production of olives and olive oil, which result in an increase in income from these activities and an overall increase in their farm income.

Evaluation Plan

The major principle of the impact evaluation is the construction of an alternative scenario, i.e. to identify what the product would be if the intervention did had not taken place. To create this alternative scenario, NORC will collect data from pairs of perimeters [one of which] has been randomly assigned to the treatment group, that receives the intervention, or to the control group, that does not receive the intervention. The evaluation is designed to determine the differences between the treatment and control perimeters, as well as the differences over time, based on the reference situation in 2010 up until the final follow-up in 2013. The evaluation plan uses the method of "differences in differences" or "double differences". To minimize the effect of a possible switchover of control perimeters into treatment areas, NORC has added a group of comparison perimeters that are not eligible for the intervention, but whose major characteristics are similar to those of the treatment and control perimeters.

The data for the analysis will be gathered from surveys with farmers in the treatment, control and comparison perimeters on four occasions: in 2010, 2011, 2012 and 2013, approximately at the same time of the year. The sampling of farmers will be drawn from lists of farmers that have parcels in the treatment, control and comparison perimeters by using a multilevel sample plan. Farmers will be selected at random in each perimeters by using a systematic sampling. A proportional allocation is then used to distribute the total sample of farmers among the various perimeters (distribution in proportion to the number of farmers in each perimeter). The sample for the follow-up surveys will include situations that have been selected according to the baseline survey (75%) and situations that have been surveyed during the reference year and for each of the follow-up surveys, thereby constituting a panel sample (25%).

The target intervention not only of all olive producers, but also Farmer Associations and private companies in the olive value supply chain. Accordingly, the evaluation plan includes the analysis of data gathered from Farmer Associations (OPAs), and from olive processing plants by surveys. Samples of these two other populations of interest are connected to the sample of farmers and the data analysis will take into account connections among these three populations within each perimeter.

The questionnaires of NORC will be created by experts of the sectors in question and by survey specialists, and in consultation with the APP, the MAPM and other key players. Questions of the questionnaires will cover the critical variables of interest for the performance and impact evaluations, including the ownership of the farmland and their legal status, levels of production, the farming techniques, the inputs, the farming income, assets, employment and other information concerning non-farm income and the composition of the household. The data will be gathered by the Moroccan subcontractor NORC and by members of the survey personnel and the Statistics Division of the Ministry of Agriculture and Maritime Fishing by using standardized forms and well-defined agreements for the work on the ground.

The performance evaluation will focus on evaluating activities performed within the context of the TC–5A contract with regard to the objectives set by said consultant. It will be careful to evaluate the degree of performance of the actions provided by the implementation of the activity as well as adhere to the timeline and the quality of actions undertaken and anticipated in the action plan of the TC–5A, the organization responsible for the implementation. The evaluation team will observe and analyze deviations between the expected results and actual results of the

activities as well as any unforeseen (positive or negative) results. Moreover, the performance evaluation will focus on determining if the activity meets the needs of the beneficiaries and if the chosen indicators are relevant and measurable. The data for the performance evaluation will be gathered from various documents concerning the intervention, qualitative conversations with members of the TC–5A Team and key players of the intervention and observations of activities if possible.

Data Analysis

The impact evaluation plan is based on detecting differences between the treatment and control perimeters in terms of net farm income and other variables of interest, and differences over time. If the random allocation indicated in this study is not compromised (it could be compromised, for example, by allowing control areas to become treatment areas), a comparison of the results between the treatment and control groups should produce an impartial estimate of the effect of the intervention. Moreover, in addition to an unvariated treatment, the data analysis could include the use of grades (or scores) for qualitative variables, and once the reference data have been examined in terms of quality and suitability, a multi-variated approach, such as Multiple Correspondence Analysis (MCA) will be considered to construct the indicators. The analysis of variance and statistical and economic models will be used to study the impact of the project considering covariables and external factors. The data analysis for the performance evaluation will focus on the categorization and the codification of data gathered from qualitative conversations and the identification of trends and relationships in the data that can explain the observed results.

II. Introduction

In December 2009, the Agency of Partnership for Progress (APP) signed a contract with the National Opinion Research Center for the design and implementation of performance and impact evaluation of the Rehabilitation and Intensification of Rain-fed Olive Plantations outputs (designated herein by "intervention"). This activity is part of the Fruit Tree Productivity Project (PAF) one of the five projects of the compact signed with the Kingdom of Morocco and the United States of America, through the Millennium Challenge Corporation for an amount of $697,500,000.

The specific objectives of the rehabilitation and intensification of olive plantations in rain-fed zones are:

- The organization, training and technical assistance of producers and their farmer associations with regard to their planning and management of farm operations.

- The training and technical assistance to olive oil production and transformation units on the improvement of production processes in view of increasing the quality level of the oil produced all while promoting health and respect of the environment at the level of these units.

The estimated results of the project were prepared by the developed element process of the Compact of the MCC. It is anticipated that the rehabilitation and intensification of the olive trees will make it possible to increase net farm income by 8.4% from now until the end of the third year of the project's implementation. Farm income from crops (not including livestock) should increase by 12.7% and income from olive production by 30.8%. It has also been anticipated that the productivity of the olive orchard will increase by 30.8% by virtue of the hypotheses of stability of actual prices. The ultimate goal of the rehabilitation program is to help reduce poverty among all households located in rain-fed zones.[1]

Considering that the APP has acknowledged that conducting surveys of income and expenses of households to evaluate the thresholds of poverty further demands resources, it has been decided to use the change in net income in farmer households as an indicator of poverty reduction. The theory of the change used in this intervention model is this: the training of farmers to use improved production and transformation techniques of olives will result in an increased volume and value of their production of olives and olive oil, which will result in the increase in income from these activities and an overall increase in their farm income.

Accordingly, NORC will conduct an performance and impact evaluation of the rehabilitation activities of olive tree plantations. The performance evaluation is intended to evaluate to what degree of success the intervention model has been implemented on the ground; the impact evaluation is intended to evaluate to what extent the intervention model has achieved the expected results – both in the short and medium term.

[1] Taken from the DPA ME-2

NORC will seek to answer the main research questions indicated below defined by the APP:

- What is the impact of the rehabilitation component of olive tree plantations on the income of farmers in the target rain zones?
- Has production of olives increased in volume and in value?
- Have techniques taught in the project been adopted by the farmers?
- Has there been an improvement in the quality of the olive oil produced?
- Have Farmer Associations played their roles in the development of the olive oil program?
- Has the knowledge generated by the project been spread to areas in which the project was implemented?
- What are the impact differences from activities of the project according to gender, age and income?
- What are the unforeseen impacts of the rehabilitation?

This report first describes the economic and agricultural context in Morocco, the rehabilitation intervention of olive tree plantations in the rain zones, and then the approach proposed by the evaluation team for the performance and impact evaluation, as well as the activities linked to the collection, treatment and analysis of data.

III. Context of the Intervention

A. Characterization of rural poverty in Morocco

With a per-capital revenue of US$ 1,700, Morocco is classified in the category of countries with intermediate income. However, social indicators remain worrisome, in particular in the rural environment. At the national level, the relative poverty rate in 2004 was 14.2% versus 16.5% in 1994, i.e. a reduction of 14% [*sic*]. The poverty rate in rural environment in 2004 remained at a very high level, i.e. 22.0% (High Commissioner to the 2004 plan). This rate hides significant disparities between the 1,298 rural communities since the poverty rate - less than under 5% in 30 communities, exceeds 30% in 348 of these communities. It is in the regions of Marrakech-Tensift-Al Haouz, Sous-Massa-Draa, Meknès-Tafilalet, Oriental, Gharb, Tadla-Azilal and Fès-Boulemane in which 90% of the poorest communities are located. These communities are the ones the PAF project is focusing on.

In addition to the 16.5% of the population considered "poor" in 1994, (i.e., whose consumption is less than the corresponding poverty threshold[2]), 39.3% of the population is considered "economically vulnerable" (those at the poverty threshold or less than 50% above the poverty threshold). This group is vulnerable due to events such as drought, which has become a structural phenomenon in Morocco, an extended family illness, or the loss of work. If the "poor" and "economically vulnerable" populations were combined, this group would represent more than half of the population to which more possibilities to participate in growth and have access to base services would need to be given.

.

B. The olive oil sector in Morocco

The olive/olive oil production is a *vital sector for Morocco*, that is the second exporter of olives on the global market and the sixth producer of olives in the world. The Moroccan olive tree culture occupies a surface area on the order of 680,000 hectares with an average production of 850,000 tons of olives in 2008/2009; i.e., 6% of world production, putting Morocco in 6th place at the global level after Spain, Italy, Tunisia, Turkey and Greece. The Moroccan Ministry of Agriculture estimates that land devoted to grooving olive trees during the 2009/2010 campaign was 700,000 hectares, i.e., around 50% of the tree growing surface (Office of Plant Production, Ministry of Agriculture and Maritime Fishing, DPV/MAPM, 2009).

Geographic Distribution:

Olive tree growing is practiced in various farm/climate zones from humid to arid, the majority of which surfaces are cultivated on hills and mountainous zones with the predominance of micro exploitation (DPV/MAPM, 2009).

Indeed, Moroccan olive tree growing, estimated at more than 68 million olive trees, is mainly distributed into 3 distinct farming zones (fig.1, DPV/MAPM, 2009):

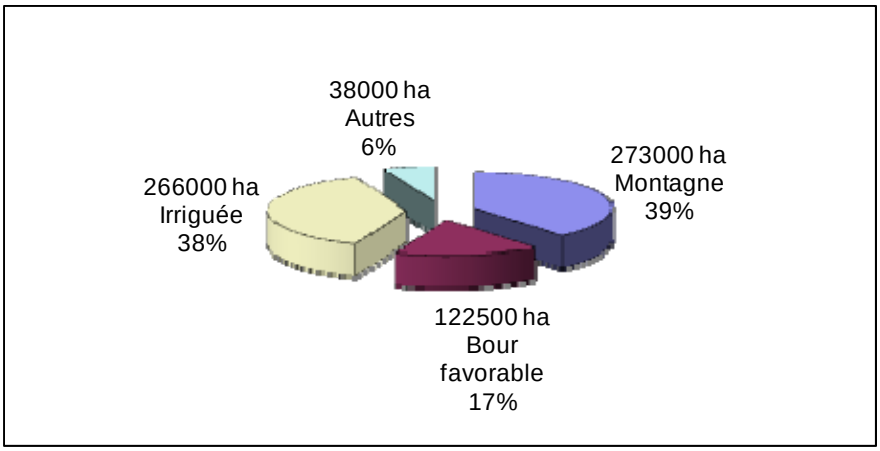
- **Irrigated zone: 266,000 hectares** (Haouz, Tadla, Souss-Massa, Moulouya, Nador, Boulemane, Oujda, El Kelaâ, Marrakech, Chichaoua, Bénimellal Ouarzazate, Tafilalet. Figuig, Essaouira).
- **Favorable Bour zone: 122,000 ha (**Séfrou, El Hajeb, Fès, Meknès, Sidi Kacem, Gharb, Loukkos, Benslimane).

2 Established at 3,922 DH in the urban environment and 3,037 DH in the rural zone (respectively 2,792 DH and 2,642 DH) for absolute poverty (World Bank and High Commissioner to the Plan, 2004).

- **Mountain zone: 273,000 ha** (Chefchaouen, Taounate, Taza, Tangiers, Tétouan, Azilal, Khénifra, Al Hoceima).
- **Other: 38,000 ha** (Safi, Settat, Khémisset and Khouribga).

The project is focused mainly on rain-fed zones where olive tree plantations do not depend on irrigation.

Fig. 1: DISTRIBUTION OF THE NATIONAL OLIVE OIL CULTURE BY ZONE

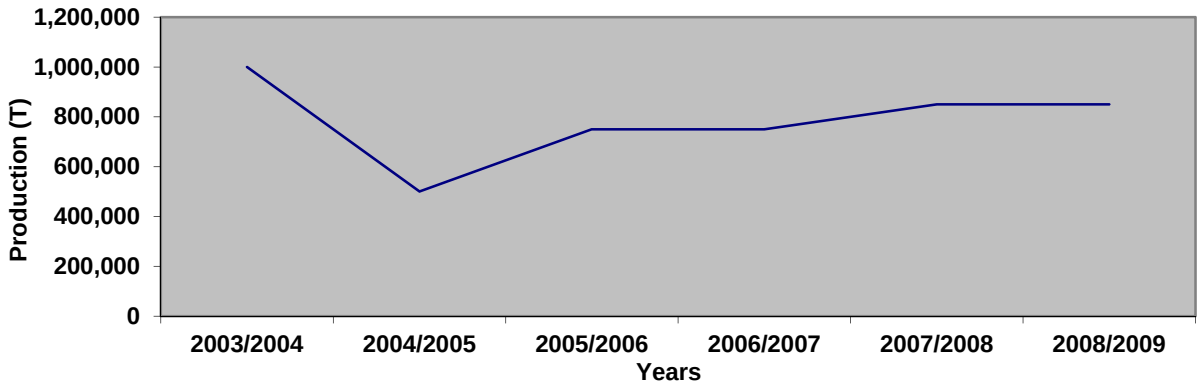


38,000 ha Other 6%
266,000 ha irrigated 273,000 ha Mountain 39%
122,500 ha Bour favorable 17%

National olive oil production

In spite of the irregular inter-annual movement of olive production due to traditional harvesting practices, production is evidencing a growing trend, from 500,000 tons in 2004-2005 to 850,000 tons in 2008-2009, i.e., an annual increase of 10%. The evolution in annual production of olives over the past 6 years (2003/2009) has varied from one campaign to another as indicated in figure 2 (DPV/MAPM, 2009).

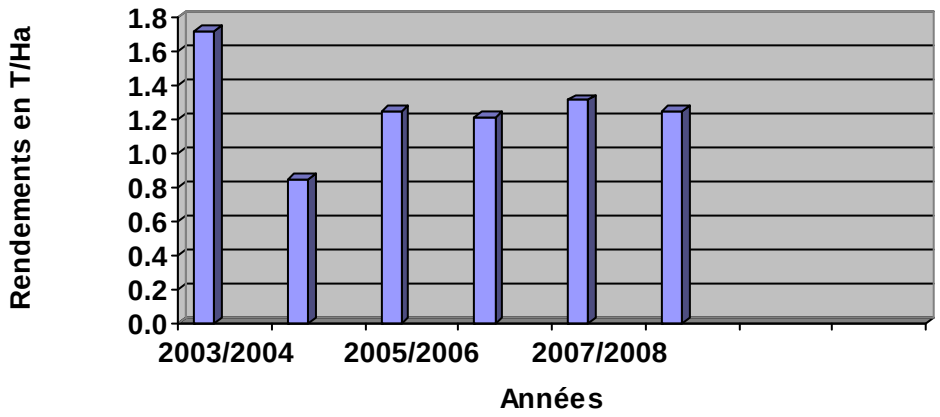
Fig. 2: EVOLUTION IN THE MOROCCAN PRODUCTION OF OLIVES OVER THE PAST 6 YEARS



Output of the olive tree:

According to figure 3, we observe that the Moroccan olive oil yield is unstable, mainly due to climate conditions, traditional harvesting conditions (beating), the lack of maintenance and alternance phenomenon. At the national level, the average yield of olives obtained during the past six years is 1.26 ton/hectare, which is very low compared to the potential of the olive tree and also to yields achieved by other olive producers in the Mediterranean basin (Spain: 3 ton/hectare). The national output in Bour does not exceed 0.5-1.1 ton/hectare and in irrigated areas 1.5-3 ton/hectare. The evolution in the yield during the period 2003/20009 is indicated below (DPV/MADRPM, 2009).

Fig. 3: EVOLUTION IN THE OLIVE TREE YIELD AT THE NATIONAL LEVEL OVER THE PAST 6 YEARS



Yield in tons/hectare
Years

The olive oil sector has suffered from several technical- and organizational-type restrictions and accordingly upstream of the production chain, the sector includes 400,000 operations of which the smallest represent 94% of the total number. These are characterized by small size (0.1 to 5 Ha), by the planting of local varieties and by low productivity (0.5 to 1 ton/hectare). With regard to large and medium-sized operations that represent 6% of the total, these are characterized by a size exceeding 5 hectares, by the use of selected varieties and by a productivity on average between 1.5 and 2.5 tons per hectare. In spite of the efforts deployed, notably at the national level within the context of the National Olive Oil Plan (PNO), production still remains low and irregular, and planted surfaces are well below the objectives of the PNO with a low density of plantations and a lower productivity. It should also be added the sector has been insufficiently exploited due to the quality of production that is mediocre as well as insufficient marketing. With regard to the organization of the profession, olive producers are insufficiently organized and currently do not have an adequate organizational framework that can defend their interests within the supply chain. In the absence of such a national entity, producers are essentially represented by a few associations working at the regional level involving a limited number of olive oil producing areas (Pack info no. 87, April/May 2010).

Olive Transformation Sector in Morocco

The amount of oil produced in the campaign 2008/2009 was 85,000 tons (Pack info no. 87, April/May 2010).

In Morocco, olives are transformed by two sectors:

- One sector that with almost 700 modern and semi-modern units and a transformation capacity of almost 450,000 tons. These units are concentrated in three main regions (Fès, Marrakech and Meknès) that include almost 66% of the total number of units while these regions only contribute 26% of the olive oil production.

- A traditional sector (*maasras*) that includes 16,000 units spread over the entire national territory. The capacity is less than 1 ton/day. These units are characterized by a very low extraction rate that does not exceed 14% and involve an important quantity of olives (more than 30% of the olives at the national level) and are characterized by a low level of technology, noncompliance with safety measures and production for self consumption (DPV/MADRPM, 2009).

C. Rural young people and women: what role do they play in the olive oil program?

In 2004, statistics on the Moroccan population showed 30 million inhabitants, 48.55% of which (i.e., 14.5 million people) live in rural areas (Ministry of Social Welfare and the Plan, 2004).

In Morocco, women and young people constitute the most vulnerable categories due to their positions in the society (home, work, unemployment, illiteracy, etc.), their living conditions and their relationships with others (Kerzali, 2003).

Rural women – on account of their knowledge and the activities they perform at the operational and household level – are the primary and decisive players of development in the rural world. They intervene at levels that vary according to the type of activities but also as a function of the natural, cultural, traditional and economic specificities of each zone. Rural women perform various types of activities that can be observed and classified as follows (INRA, 2005; Nah, 2007):

- Domestic activities
- Activities linked to crops
- Livestock activities
- Other activities such as weaving

In the domain of fruit and olive orchards, the role of women remains a function of their social class, their status and their availability considering the multitude of their other responsibilities, domestic as well as productive. More specifically, the two tasks to which they devote time on gathering olives and harvesting of various fruit. These are hardly menial duties since they represent the most demanding jobs in the workforce (INRA, 2005).

With the economic development, rural women have become increasingly present as heads of household and have become more active. The number of women's cooperatives has gone from 528 in 2006, to 2,738 in 2008 (Office for Development and Cooperation (ODCO), 2008), with 307 cooperatives in the farming sector (ODCO 2008). There are 4 women's cooperatives in the olive oil program (FIPA, 2008). This reflects a clear trend toward the valuation women's work and the recognition by society of their indisputable contribution to the development process.

With regard to young people in rural areas, according to the estimates of the Center of Studies and Demographic Research, 37% of rural youth are under the age of 15, 40% are under the age of 20 and 65.6% are under the age of 30 (Nah, 2007). In rural areas, young people generally work as family aids (74.5%) and employees (17.5%). With regard to the olive oil program, the role of young people involves the two most demanding tasks, meaning the harvesting and collecting of olives. Rural young people work as employees in the *Maasras* (traditional olive processing plants) (INRA, 2005; Nah, 2007).

At the Farmer Association (OPA) level, although data on OPAs created by youth are unavailable, it is definite that the National Development for Human Development (INDH) and the olive oil plan have contributed to create cooperatives and farmer associations. According to one study published by the ODCO in 2002 regarding cooperatives of young people in Morocco, there are 131 cooperatives of young people in the farming sector and there is a single cooperative in the olive oil program (Office for Development and Cooperation (ODCO), 2002).

D. Farmer Associations (OPAs)

Farmer Associations are currently needed for driving local and regional economies as well as defending the interests of the farmers they represent. They constitute a veritable counterweight to the omnipresence of the private sector, for certain strategic programs (Desrues, 2004; Bessaoud, 2005). In Morocco, the public authorities consider organization into farm cooperatives as a tool of agricultural policy, capable of increasing farm production and making the actions of the government profitable by mobilizing the population to participate more in the development process (Office for Development and Cooperation (ODCO), 2002).

Farmer Associations are voluntary groups of individuals or legal entities in view of together achieving common objectives. Based on this definition and according to the objectives we distinguish:

Agricultural chambers

Chambers are public territorial agencies, managed by elected members. Currently there are 16 in operation for the 37 regions (Dahir no. 1-09-21 dated 22 safar 1430, Official Gazette no. 5714 of March 5, 2009).

Farm cooperatives

Agrarian reform cooperatives are distinguished by their special legal status as established by Law no. 1.72.278 of December 29, 1972 related to Agrarian Reform Cooperatives and Non-Agrarian Reform Cooperatives (Dahir no. 1-83-226 dated 9 moharrem 1405 (October 5, 1984)).

Agrarian reform cooperatives are focused on protecting farmland and the shared use of means. There are 738 cooperatives and they farm 325,000 ha (AIAM, 2008). Non- Agrarian Reform Cooperatives are governed by the Dahir dated October 5, 1984 involving enactment of Law no. 24-83 put into effect by Dahir no. 1-93-166 dated September 10, 1993. These organizations are subject to the protection of a public department, i.e., the Office for Development and Cooperation (ODCO).

Cooperatives function according to certain principles or rules that are based on management and personal and mutual responsibility of democracy, equality, fairness and solidarity.

Table: Importance of cooperatives by program

Supply chain	**Number of cooperatives**
Cereals	12
Milk	1254
Bee culture	502
Livestock	966
Supply	413
Farm equipment	183
Vegetable farming	120
Irrigation	103
Argane Oil	144
Olive oil production	**91**
Total	3788

Source: Annual Statistics of Cooperatives in Morocco, edition 2009

Farmer Associations

Farmer Associations are founded on the fundamental right of associating with others. An association is defined as being the agreement whereby one or two or more persons pool together their knowledge or their activities for a purpose other than sharing profits. Associations are regulated by the Dahir of Public Freedoms of 1958 then by the Law of Associations of July 2002.

There are fewer Farmer Associations than there are cooperatives. They cover a number of sectors involving goods and services and actions of local farm development (Office for Development and Cooperation (ODCO), 2008).

The main objectives of farmer associations are the following:

- To guide and help the development of their members and more generally study and understand the collective or private interests of these members.
- To study the main economic, social, technical, financial, legal and fiscal matters.
- To inform their members of steps, decisions, advice or results.

Successful programs and failures

Three programs have experienced in significant development and have significantly contributed to improving the production and living conditions of farmers. These include the milk, cereals and argane oil programs. The performance achieved by the OPAs in these sectors are the result of a sector-wide policy that notably supported the supply chains in question (Office for Development and Cooperation (ODCO), 2008).

There have been failures, however, notably OPAs created within the framework of initiatives or driven by the local or regional administration, above all those that are created by integrated rural development projects. These involve certain cooperatives created for the shared use of farm equipment (CUMA), associations focused on water use for farm purposes (AUEA); herding and vegetable crop cooperatives, etc. It should also be pointed out that the olive and olive oil sector is still insufficiently organized (Office for Development and Cooperation (ODCO), 2008).

Reinforcement of OPAs

The adoption of the new farm development strategy solidified by the Morocco Green Plan reaffirms the farm heritage of Morocco. Agriculture continues to be one of the key sectors of the economy and the policy of the government (Morocco Green Plan, 2009).

The farm development strategy for 2020, presented in July 2000 introduced a fundamental element, that is the recognition of the role of collectives representing farmers as an intermediary player for the application of the new farm policy. In fact, the government has recognized that there needs to be an organized farm sector as a necessary condition to achieve agricultural development.

The Morocco Green Plan announced in April 2008 confirms this trend, specifically by adopting the association as an organizational mechanism for agriculture. Cooperation remains a driving force in the development and promotion of the farming world (Morocco Green Plan, 2009).

Moreover, one of the prior conditions for the intervention of the fruit tree project at the level of the perimeters concerned is the establishment of a Farmer Association (e.g., an association or a cooperative).

E. The intervention

The Fruit Tree Agricultural Project (PAF), a part of which is the intervention in question, is one of the "main intervention models envisioned by Pillar II of the Morocco Green Plan that advocates the collaboration of small farmers" (DP TC-5A, 2009). In fact, the PAF encourages cooperation among all participants in the value chain with the creation of partnerships and contract links.

The intervention that will be implemented for the Rehabilitation and Intensification of Rain-fed Olive Plantations will therefore attempt to address needs of olive producers all while bringing olive processing plants into compliance and reinforcing capacities of Farmer Associations of the olive oil sector. The project will also focus on women and young people. This is detailed in the following section on the intervention logic and the project indicators.

IV. The Intervention: Intervention Logic and Indicators

The Rehabilitation and Intensification of Rain-fed Olive Plantations ("the Intervention") focuses three major populations: the farmers (olive producers), Farmer Associations (OPAs) and the olive processing plants. Of these three populations, the farmers are the most important considering the fact that the net income of farmers is the key indicator of the impact evaluation. Farmers will be trained in the best olive production and harvesting techniques in order to increase yields and the quality of their production. OPAs are an essential component of the intervention since they play a key role in the organization of farmers for a more effective implementation of the intervention. The creation of an OPA is a prior condition for farmers to exploit the perimeters can participate in the intervention. The intervention will reinforce the capacities of the OPAs and will encourage them to expand to give farmers a stronger negotiating power and more important buying power.

An important component of the olive value chain is the production of olive oil. A high quality olive oil can potentially be sold at a higher price. There are therefore plans to bring olive processing plants up to grade and to establish standard quality systems in certain units and higher quality in others to ultimately obtain a better quality oil. OPAs will be encouraged to sign contracts with olive processing plants in a win-win partnership, since the OPAs will have a guaranteed buyer and the olive processing plants will have a supplier of reliable quality. Marketing and promotion will be provided by OPAs and the olive processing plants. Moreover a price system will be set up to ensure the marketing of the production with encouraging prices allowing olive oil farmers to invest in the maintenance of their olive tree plantations.

There are two subpopulations targeted by the intervention. The first is that of young people who receive training to become experts in various aspects of olive production, such as size or plant disease control. The second is that of rural women who will have the opportunity to receive financing for small "pilot" projects related to the production of olives.

The NORC team has created a preliminary model of intervention logic, based on its understanding of the intervention and the way it can lead to an increase in the income for households of participating farmers. This can be modified or adjusted when the TC–5A, the entity responsible for implementing the intervention, the APP and other key players supply (1) more details on the intervention, and (2) their conception of the way the intervention will achieve its objectives. The basic theory is that each activity of the intervention leads to a certain chain of short-, medium- and long-term results that when combined, will contribute to an increase in the farm income. Each major element of the intervention logic is described below and illustrated in Annex C.

Training, organization and technical assistance of farmers: the adoption by farmers of best production techniques of olives and olive oil will allow their improvement in quantity and quality, which will result in increased sales and accordingly an increase in revenue.

Creation and organization of the OPAs: the creation and structuring of the OPAs will lead to the development of partnerships with olive distributors and transformers (the larger the OPA, the more significant buying power it will have), which will lead to a more advantageous sale price for farmers. Moreover, the creation of a price information system will maximize profits by enabling the OPA or the farmer to obtain the best possible prices. This will result in an increase in the value of production and accordingly to an increase in farm revenue.

The upgrade of olive processing plants: the training of olive processing plants will lead to their adoption of a better olive oil pressing system, which will increase the quality of the olive oil that can thus be sold at a higher price. This increased production value will accordingly lead to a higher farm revenue. A better marketing will help also the increase in the production value by increasing the quantity sold.

Young people: the training of young people will allow the development of local experts in important aspects of olive production, which will yet again demonstrate the adoption of best production techniques.

Women: Pilot projects of women (the nature of which is not been made clear and that must be determined) will contribute additional revenue into households,

The intervention logic proposed below will be reviewed in collaboration with key players of the intervention and a final version will be used to evaluate: (a) if planned activities have been implemented, the quality of these activities and their short-term results (performance evaluation) and (b) if each short-term result effectively leads to an increase in farm revenue according to the chain of logic and to what extent this has been observed (impact evaluation). These indicators are described more in detail Annex B.

Activities		Results		Goal
	Local experts for consultation			
Training of young men to be experts	Use of good growing practices			
Training and technical assistance of farmers and OPA	Development of partnerships with transformers and distributers	Better tree health	Increase in olive yield	
Set up of a price information system	Creation of second-level organizations	Better quality of olives	Increase in quantity of oil produced	
Upgrading of olive processing plants	Better information on prices	Better price obtained for sale of olives and oil	Increase in the value of the production/products	Increase in farm income
Training of marketing and promotional operators	Use of standard quality and high quality oil production systems	Better olive oil quality	Increase in sales	
Marketing training	Increase in marketing activities	Increase in clients for olives and olive products		
Pilot women projects				

The two following sections present the approach that will be adopted for the performance and impact evaluations of the intervention.

V. Performance Evaluation

The performance evaluation aims at evaluating the activities performed within the context of the TC–5A contract with regard to the objectives set by said consultant. The performance evaluation will involve the activities performed by the consultant of the contract as well as the rehabilitation action plans (feasibility studies) prepared by the consultant of the TC-1A contract.

More generally, performance evaluation will answer the following questions:

1. What is the validity of the intervention logic, and what are the hidden hypotheses?
2. Has the program reached the initially-defined objectives within the logical context and according to the needs of the target populations? What are its strengths and weaknesses?
3. What are the success factors and barriers to reaching these objectives?
4. What are the threats and opportunities that can help or harm the long-lasting potential of the results reached?

A. Objectives of the performance evaluation

The performance evaluation will aim at evaluating the efficacy of the intervention, i.e. the reality of the actions conducted following the implementation of the activity. It will take care in evaluating the degree of performance of the actions planned for the implementation of the activities as well as adherence to the timeline and the quality of actions undertaken and anticipated in the TC–5A action plan. It will also evaluate the actual level of participation and or/ of the beneficiaries as well as the operation of the OPAs and their capacity to manage and develop the perimeters under their control. The evaluation team will observe and analyze any deviations between the expected results and the actual results of the activities as well as any unforeseen (positive or negative result).

Moreover, the performance evaluation will focus on determining whether the activity meets the needs of the beneficiaries and if the indicators chosen are relevant and measurable. As far as possible, it will evaluate the level of coordination and integration of the resources mobilized to reach the expected objectives as well as the concordance of the activity with the general guidelines of the MAPM and the other activities undertaken by the government and other sectors of the development. Lastly, the long-lasting effects of the activities after completion of the TC–5A contract will also be evaluated in terms of financial and operational viability of the OPAs, steps needed for the management and development of the rehabilitation perimeters and the protection of the environment.

The evaluation of the achievement of the objectives will be based on the following criteria:

♦ **The relevance:**

- Correspondence with needs and demands of the beneficiaries;
- Compliance with the general guidelines of the MAPM (Strategy of the agricultural development, strategy of rural development, sector-wide strategies, etc.);
- Relevance of the indicators chosen and evaluation of the measurability of these indicators.

♦ **Coherence: relations among the various elements that make up the activity.**

- Concordance of the various means and tools mobilized to contribute to the achievement of the objectives (internal coherence);

- Intervention approach of the activity and its level of integration;
- Concordance of the activity with other actions taken by the Government, stakeholders and other development players (external coherence).

♦ **Efficacy: the reality of actions taken after the implementation of the activity.**

- Overall evaluation of the quality of feasibility studies and rehabilitation action plans prepared by the consultant of the TC–5A contract;
- Degree of performance of the actions planned for the implementation of the activity (TC–5A contract);
- Adherence to the timeline;
- Quality of the actions taken;
- Quality of managing unforeseen events;
- The application of the participatory approach and actual level of participation of the beneficiaries;
- Management steps taken by taken to promote adhesion by the farmers and other sectors of the olive oil program to the activity, and adoption of best practices with regard to technical performance, marketing and exploitation of olive trees;
- Level of organization of the beneficiaries;
- Operation of the OPAs and the ability to manage and develop the perimeters under their control;

♦ **Long-lasting effect:**

- Financial and operational viability of the OPAs;
- Actual management and development of the rehabilitation areas;
- Cooperative measures;
- Possibility of reproducing or disseminating the activity.

Moreover, the evaluation team will evaluate the level of achievement of the objectives of the activity as well as any unexpected (positive or negative) effects. By referring as far as possible to the indicators of the logic framework of the Fruit Tree Productivity Project (PAF), the evaluation team will specifically address the following:

- Expected results and actual results of the program;
- See to what extent the actual results contribute to the achievement of the specific objectives of the project and/or they could contribute in the future,
- Unforeseen results (negative or positive, windfall effects);
- Analysis of deviations observed (by indicating the success factors and any barrier factors).
- What supplemental steps could the Principal of the activity take, in addition to those already taken?

B. Methodology

The performance evaluation will be based on an examination of the following data and documents:

1. Intervention logic
2. Logical framework
3. Feasibility Studies of TC-1A
4. Follow-up reports of TC-5A (Indicator Tracking Table)
5. Quarterly report of TC-5A (if applicable)

6. Follow-up reports of the UGP
7. Qualitative data from NORC surveys
8. Qualitative data from conversations with resource persons
9. Observations of activities if possible

As the impact evaluation is an integral part of the performance evaluation in its broadest sense, the evaluation and performance mission will use the quantitative data to be collected for the impact evaluation. To supplement this quantitative information, the performance evaluation will also use qualitative conversations (semi structured and open conversations).

The target populations of the qualitative conversations are the following:

- Personnel participating in the implementation of the TC-5A program
- Farmers
- Representatives of Farmer Associations
- Representatives of olive processing plants
- Fruit Tree Agricultural Project management unit
- Work Centers
- The Ministry of Agriculture

The evaluation team anticipates working with the TC–5A team after the start of the project to include planned activities and to finalize the intervention logic of these activities. This will therefore consist of qualitative conversations with key personnel of TC–5A at the beginning of the intervention. Moreover, the evaluation team will collaborate with the TC–5A team to establish tracking forms that will address all the described indicators in the logic framework to ensure an exhaustive tracking of the activities and their performance. These qualitative conversations will also be conducted at the end of each year to find out whether there were any changes from the design and intervention standpoint as well as the reasons for these changes.

The qualitative conversations involving other target populations will take place once a year during Phases II and III of the project (2011, 2012, and 2013). The exact number of these conversations remains to be determined. More specifically, the qualitative conversations with farmers, OPA representatives and representatives of olive processing plants will focus on evaluating the "quality" of the activities implemented by TC–5A, the relevance of activities with regard to their needs as well as evaluate whether these activities were implemented in a timely manner. Qualitative conversations with the Work Centers and the Ministry of Agriculture aim to clarify the coherence of activities of TC–5A with the other activities actions undertaken by the government, stakeholders and other development players (external coherence).

Lastly, these talks will also be crucial for evaluating the long-lasting effect of the activity considering that the impact evaluation, ending in 2013, will not have sufficient hindsight to evaluate this criterion definitively. For the last series of qualitative conversations in 2013, the evaluation team will therefore take care to include matters concerning projects and activities in the future of beneficiary populations to evaluate the long-lasting effect of the activities without the assistance of TC–5A and the way that MAPM anticipates continuing some of the activities that could still be necessary.

A cross-check between the TC-5A and UGP documents (follow-up reports) with the data gathered from quantitative surveys and qualitative conversations of the evaluation team will be conducted. Deviations observed among the various sources of data as well as deviations between the initial objectives and the objectives achieved from the activities ("achievements of activities" in the logic framework) could accordingly provide information on the content of the semi-structured questionnaire of the subsequent qualitative conversations.

C. Results

Based on conclusions drawn, experts will make realistic recommendations. The recommendations will help improve performance of the program. The evaluation team will seek to find out to what extent the difficulties or strengths or weaknesses of the intervention could be resolved and how the advantages could be reinforced. It will seek to draw lessons that might be useful to clarify the design and the formulation of similar activities in the future.

A follow-up of the performance evaluation will be included in the annual reports that NORC will deliver to the APP. A final evaluation report of the performance will be delivered to the APP at the end of Phase III.

VI. Impact Evaluation

A. Context

The rehabilitation of the olive tree plantations is planned for 45,000 ha of olive trees in 17 provinces. This zone is divided into groups of neighboring parcels that are called "perimeters" in various rainy regions in Morocco. A perimeter represents a geographic sector measuring between 200 and 250 ha of olive trees on average. From an analytical standpoint, these groups of parcels are perceived as blocks. The block system on one hand makes it possible to organize the logistics of the intervention where the treatment is applied to a large block following the example of the small units/parcels and on the other hand to try to control the factors of variability by grouping together parcels according to homogeneous groups and by eliminating effects from certain potential variables (covariables).

The rehabilitation of these perimeters is scaled over two (2) years. The perimeters concerned by this evaluation of impact should be made up of 70 perimeters of Segment 2 of the TC–5A contract that will be covered by the intervention toward the end of 2010 by the TC–5A consultants[3]. They have been identified as a function of the historic files of farmland covered by other farm programs and government interventions. They have also been studied for their eligibility in the intervention according to a set of criteria. These criteria include: geographic location in the unirrigated farm zones adapted to the olive species; reasonable distance to markets; accessibility; existence of highways and useable trails; existence of technical personnel; proximity of technical services; existence of infrastructure of supply of materials and maintenance; average rainfall greater than 350 mm; topographic characteristics (slope from 5% to 50%); appropriate soil; ownership (*melk*) or similar of the parcels; predominance of small and medium midsize operations (≤ 5 ha); ages of plantation over 10 years; current density of at least 50 trees per hectare; and agreement of potential beneficiaries to rehabilitation actions. The phases followed during the perimeter selection process have let us to take the following factors into account:

1. These perimeters are comparable to each other insofar they share a set of characteristics.
2. These perimeters do not represent a random sampling of olive tree plantations in the rain zones in Morocco or the zone of total intervention of the Fruit Tree Productivity Project (PAF).

This means that no inference of the results on the rest of the population is possible in the current circumstances. Thus, any generalization of the results from this study is therefore to be done with precaution.

B. Randomization

The question that we try to resolve in the impact evaluation of this intervention is the question concerning the construction of an alternative scenario, i.e. to identify what would happen if the

[3] The exact number of treatment perimeters changed for different reasons. Please consult the sample plan below. The final status of each perimeter is described in detail in the reference situation report that will be delivered end of 2010.

intervention had not taken place. To answer this question, we must gather information from perimeters that have received the intervention (the treatment perimeters) and those with similar characteristics that did not benefit from the intervention (control/comparison perimeters) representing the alternative scenario and then compare any differences between the two groups. The control/comparison group must be identical on average to the treatment group at the reference situation in order for a fair comparison to be made.

The best approach to identify a group within the framework of this objective consists of equalizing the perimeters as a function of their characteristics and to randomly assign a perimeter of each pair to the treatment or control group[4]. This randomization is also an ethical and fair way to determine what group has the priority will benefit from actions of the rehabilitation whereas the other group will not benefit from the intervention during the evaluation. The control perimeters will act as reference points to evaluate the progress of the intervention. The randomization phase will therefore make it possible to reduce the subjectivity of the choice of the perimeters involved in the intervention.

The randomization also supplies a basis for future statistical analyses that can demonstrate the direct impacts of the interventions of the project (cause/effect ratio) by analyzing not only the changes over time (before and after) but also by comparing them to a control group that has not yet benefited from the intervention. The result is an analysis based on the method of double differences that can be reflected by the following equation: $(T_F\text{-}T_B) – (C_F\text{-}C_B)$ where **T** represents the treatment perimeters and **C** refers to the control perimeters whereas **B** refers to the baseline situation and **F** the final situation.

C. Suggestions for improvements

Ideally, the randomization of the perimeters should have been made within each level covering an interval of scores anticipated of propensity for allocation in the control group; an improved matched pair between the treatment and control groups can then be obtained which makes it possible to reduce the experimental error. The subsequent stratification and the propensity score matching are scheduled in view of reducing deviations and improving the comparison between the treatment and control perimeters.

A more detailed examination of the matched pair process and observations on site raised some issues that have concerned the NORC work team. These involve certain critical variables (e.g., legal status) that had not been included during the pairing process. It is possible that the cause is the lack of a wide range of relevant information available at the beginning of the project. To address this issue, we propose adding a subsequent matched pair phase (with regard to the perimeter and that of the farmers) for analysis purposes, when additional information on the perimeters in the farmers involving these perimeters are available to increase the precision of the matched pair.

[4] The randomization phase was assigned to an independent consultant. For more detail, please consult the randomization report.

In order to optimize the representativity and to study some of the issues raised above related to the methods of selection and certain problems of pairing, the NORC team proposes the following:

1. Strengthen the quantitative analysis by qualitative analyses regarding the reasons of refusal and non-eligibility as well as the demography of farmers within these perimeters.
2. Add some comparison perimeters to the sampling (e.g., perimeters not eligible, eligible perimeters that are not part of Segment 2). The comparison perimeters are perimeters excluded from the program during the same period as the perimeters of Segment 2 that are not part of the randomization.

These additional comparison perimeters are perimeters that could be affected by effects of the dissemination of the intervention or perimeters that could have been considered for the intervention but were considered ineligible due to the absence of one or more minor criteria of eligibility (e.g., the surface area and the slope). The addition of these comparison cases will mitigate the possible increase in control perimeters with the treatment group and will make it possible to measure the effective dissemination, by using a quasi-experimental approach as an alternative strategy to the need to support the overall evaluation of the program. The *quasi* experimental approach is based on the comparison between treatment and comparison cases and makes it possible to relate associations between the intervention and results observed (Rosenbaum, 2002).

It is necessary to note that the increase in the number of perimeters makes it possible to improve the direct strength of the evaluation versus any unforeseen change. As a consequence, the use of a comparison group would also be used as an alternative plan in the following cases:

1. Control groups become treatment perimeters
2. Farmers refuse to be covered by the intervention,
3. An appropriate match pair for treatment cases cannot be found in the pre-existing control group due to distinct characteristics including the difference of which is significant (in the case of subsequent matched pair),
3. The experimental plan includes problems that require the use of a quasi-experimental analysis plan instead of an experimental analysis.

D. Impact Variables

For the evaluation of impact of the intervention, several economic data must be entered including any change that occurred with regard to farm income, employment, yield of olive trees and the quality of the olive oil among other indicators. Although we are aware of the importance linked to the net farm income as key variable from an impact evaluation standpoint, the response rates of the pilot test will help us determine whether or not it is possible to gather adequate data to calculate these variables. The team, that will then be aware of any challenges associated with gathering of data related to this variable, can be better prepared and can intensify its efforts at two levels (training, groundwork, follow-ups, etc.) to obtain precise and reliable data. Moreover, alternative plans allow the research team to gather information on other variables that are correlated to the farm income, such as production, if difficulties associated with gathering data concerning the net farm income cannot be overcome.

E. Impact Evaluation Analysis

Differences measured in the change of net farm income or any other production indicator, estimated between the treatment perimeters and control/comparison perimeters will be used to evaluate the impact of the program. Since the randomization was used for the allocation of perimeters to the treatment and control groups, the hope is that this difference will be equal to zero under the null hypothesis of lack of differences on average. We express it as follows:

$$E\{(T_F-T_B) - (C_F-C_B)\} = 0$$

In view of reducing any experimental error, this difference between the treatment perimeters and control/comparison perimeters must be conditioned, by known differences between the two groups, such as ethno-linguistic regions or geographic zones (Khuel, 1994). We will use here an adjustment model with covariables and the more this method succeeds, the better the experiment's inferences will be, and in other words we all need fewer perimeters to obtain the same level of precision.

The number of treatment sites is generally equal to the number of control/comparison sites, which makes it possible to use an equilibrated plan. Moreover, certain specialists recommend that the number of control sites be preferably greater than that of treatment sites (Rossi, Lipsey, and Freeman, 2004). However, in view of the limitations imposed by the structure of the intervention by segment, the solution setup is complex and represented a decisive factor for adding comparison perimeters.

These are the phases for preparing the analysis:

1. Conduct a reference survey before proceeding to the intervention in the treatment and control/comparison perimeters.
2. When the intervention has begun, conduct three follow-up surveys (one survey per year). They are to be easily comparable to the reference survey, with regard to the questionnaire used as well as the sample units.
3. Calculate the average difference between the key indicator values before and after the intervention for treatment and control/comparison groups.
4. Calculate the deviation between these two average differences to obtain an estimate of the impact of the program and to test its significance after inclusion of the covariables.

It is important to note that neither parcels nor farmers have been directly allocated to the treatment or control groups. These are the perimeters to be rehabilitated that constitute units that are allocated to the treatment and control groups. However, surveys and data are gathered at the farmer level. The intensity of the treatment depends on the level of active participation of the farmers at the time of the training. It is possible to evaluate the effect of the intervention without having to be concerned with variations within the same perimeter or variations between perimeters. The variation between perimeters is due to differences between farmers. The variation between perimeters can be explained by the main factor (intervention or lack of intervention) and by other covariables that will be taken into account in the analysis. When covariables and differences between farmers are taken into account, it will then be possible to evaluate the effect of the intervention.

As a consequence, and based on the analytic perspective, this plan can be treated as an experimental plan in randomized equilibrated blocks (Randomized Complete Blocks RCC) in which the intervention group represents the main factor. However, there are other covariables.

These covariables are variables that can affect the measured result but that are not part of the central interest. The NORC team will identify these covariables during the implementation of the intervention to take their affect into account on the econometric model. Some of these covariables have already been checked and verified by the block system. The basic concept consisted of creating homogeneous blocks in which covariables are maintained constant while the variables of interest can change.

Several physical characteristics (rain, type of soil, slopes), layers (tenure method, size of the operation) farming (age of the olive trees, density, etc.) and social factors (receptiveness and commitment of the beneficiaries, availability of labor) etc. can be taken into account in the RCC model as covariable if they are not used as block identification factors.

Three additional questions must be included in the analysis:

1. Seasonal and cyclical variability: by taking the seasonal fluctuation into account of the olive harvest, the cyclical fluctuation of alternating production years, the execution of rehabilitation plans at the perimeter level should starts just after harvest of the year in question (October – December). The survey must be conducted each year over four years at the end of each season and cover all activities of the complete season.
2. Spatial correlation and dissemination effect: these two issues must be addressed jointly insofar the spatial proximity of the perimeters represents a factor that affects the inter-and intra-dissemination of knowledge.
3. Interaction effects with other programs or other factors.

F. Sample plan and sample size

The study is made up of three different populations of interest:

1. Farmers (olive producers) who work the parcels in the selected perimeters,
2. Farmer Associations (OPA) associated with the selected perimeters,
3. Olive processing plants servicing the farmers who work the selected perimeters.

Surveys will be respectively conducted with members of the household of the olive producers, representative members of the OPAs and representatives (e.g., director) of olive processing plants. The analysis will therefore be conducted at several levels, for example at the perimeter level, farmer level, farmer association level, and the olive processing plant level.

Parcels:

The initial proposal of NORC consisted of sampling parcels included in the perimeters and then identifying the farmers that operate these parcels to gather data on the parcels as well as their as on the households associated with these parcels and follow them for four (4) years (reference/pre-intervention situation, two years of intervention and post-intervention). However, due to logistics and information availability factors, an alternative approach was adopted for the creation of a baseline survey of farmers working the land in these perimeters.

This baseline will be designed according to the list exercise executed by MAPM and will be reviewed and cleaned up by NORC in preparation of the sampling.

Farmers represent the primary focus of interest of the study since they supply the information on the net farm income as well as on other indicators used for the analysis of the impact evaluation. The baseline will be carefully examined after it is compiled. Generally, the most important problem at this phase, other than typical base errors (e.g., incomplete or obsolete basis) is that of duplication of farmers in the baseline survey. These double cases involve farmers working parcels in different perimeters of Segment 2. To avoid wearing out respondents and to better manage the logistics in the field, "double" farmers will be identified in the baseline (as far as possible) and assigned to a single perimeter.

Associations and olive processing plants:
Analysis units for impact and performance evaluations of the interventions specifically target these groups remain unchanged. A triangulation is integrated in the plan to enter interactions between the Farmer Associations, the olive processing plants and the farmers.

Sample
The sample size for analysis and power is partially reflected by the number of perimeters. These figures are lower than the figures at the beginning of the program due to the merger of certain perimeters and the refusal of one perimeter to participate in the study.

The sampling will be conducted according to a random stratified sampling in which the strata are the perimeters. Due to errors or missing data in the baseline survey, a certain number of perimeters should be excluded from the base. More particularly, 2 treatment perimeters, 1 control perimeter and 3 comparison perimeters will be excluded due to refusal to be listed in the baseline, missing data regarding the province, perimeters that had slightly different names but are in fact the same perimeter. A certain number of perimeters will be excluded due to the complete absence of farmer names; more particularly 1 treatment perimeter, 18 comparison perimeters and 1 perimeter without status identification (this perimeter refuses to participate).

A proportional assignment is used to distribute the total sampling of farmers among the various perimeters. Although the baseline survey includes a number of variables of interest, some abnormalities suggest that some of these data are inaccurate. For example, for some perimeters, the size of the parcels and/or the number of trees per farmer seems to have been calculated or estimated, instead of having been gathered from a reliable and objective source (administrative files or the farmers themselves). The age of the farmers on the other hand seems to have been normally distributed and should therefore be used to sort the farmers within each perimeter, then a systematic sampling will be selected within each perimeter. For perimeters with farmers whose age is under 20, a survey of farmers in these perimeters will be included in the sampling to guarantee not only an adequate coverage of these perimeters, but also sufficient number of respondents for analysis purposes.

The sampling of the farmers in the three types of perimeters (treatment, control and comparison) will include 4,293 farmers in all and should be proportionally distributed among the perimeters in question. By considering a 90% rate of response and a 85% rate of trackability of farmers (due to errors in the baseline survey), the number of complete cases should be approximately 3,284 farmers. It should however be noted that this random sampling of farmers in each perimeter does not make it possible to report the results of the study at the perimeter level. They will therefore be reported at a higher level of aggregates (such as the regional level) by comparing the treatment cases and control cases.

Among the treatment, control and comparison groups as follows:

	Treatment	Control	Comparison	Total
Number of perimeters	67	68	14	149
Number of farmers sampled	2,280	1,701	312	4,293
Anticipation of the number of complete farmer cases	1,744	1,301	239	3,284

The sample will then be distributed between C&O and MAPM so they do not gather information from the same *douar*. C&O will cover 48 treatment perimeters, 49 control perimeters and 11 comparison perimeters. The MAPM will cover 19 treatment perimeters, 19 control perimeters and 3 comparison perimeters. These perimeters will be allocated according to a random sub-sample stratified by clusters in which the sample randomly selected will be allocated to C&O and the remaining part allocated to the MAPM.

The final number of farmers surveyed will be detailed in the report of the reference situation. The same farmers will again be surveyed in phases II and III of the project. Thus, a 100% panel sample will be used to increase the statistical power.

The baseline survey of Farmer Associations will be updated once a year to take into account new organizations that will be created within the framework of the direct impact of the intervention and the intervention objectives of creating new Farmer Associations. A survey of the OPAs will be made for phase I (year I) of the project and this sample could possibly increase the number of OPAs to 150 in the following years.

For olive processing plants, the sampling is always uncertain due to the absence of information on the link existing between the perimeters and the olive processing plants. TC–5A has contacted 100 olive processing plants; these 100 units will be selected from among the olive processing plants used by farmers of the treatment perimeters. The treatment olive processing plants will therefore not necessarily be part of the treatment perimeter but could be outside of these perimeters. The contacted units affected by the intervention will be identified as TC–5A makes progresses in each perimeter, and an exhaustive list of all treatment units will not be available before April 2011 at the earliest. There are thus plans to survey all the treatment units progressively, as they are identified and before any intervention is received. Then, an equal number of olive processing plants of the control group will be surveyed, and the treatment and control olive processing plants will be paired off according to observable variables.

According to our discussions with APP, we note that there is a particular interest regarding the impact of the intervention on women and young people in rural areas within the population of the farmers. We have accordingly incorporated the age in the sampling plan of the farmers and will attempt to subsequently stratify according to gender, if the available information at that moment proves to be sufficient. The baseline survey for all populations of interest will be supplied and updated once a year by APP/MAPM with assistance from local work centers.

Calculation of the power

A detailed description of the calculation of the statistical power for our evaluation using a pre-test/post-test design with control group is presented in Annex B of this report.

Collection of Data

A. Design of survey tools

The main objective of the project is to evaluate the impact with an agricultural intervention.[5] To do this, the breadth of the difference, if applicable, needs to be measured between beneficiary groups of the intervention and the other groups, as well as the breadth of the change, if there is one, between the period before and after the intervention; then to determine whether any difference is attributable to the intervention. An effective means for measuring these differences consists of systematically gathering data from beneficiaries and non-beneficiaries of the program before and after the intervention.

The preparation of the survey document will proceed according to the methodology "Ask the same questions"; Harkness 2008; Harkness, *et al.* 2010), in which a source questionnaire in large part is produced in the same language, then reproduced in local languages. The first section gives a glimpse of the process of preparing questionnaires for samples of farmers, Farmer Associations and olive processing plants and the second section involves the cultural and linguistic points taken into consideration during preparation of the documents.

Questionnaires are measuring documents that are applied to the persons interrogated, but that can be made up of different persons interrogated (Groves, 1989). The most common process for preparing a questionnaire includes elements the validity and reliability of which have been evaluated among the population being studied, as well as the use of research elements specified by sector-wide experts and tested on several occasions on the population, object of the research (Harkness *et al.* 2010). The objective of NORC for the preparation of new questions of the questionnaire is to adapt or create elements with a high degree of validity and to test them several times on sample populations in order to establish their reliability. The process of preparing the questionnaire is planned as follows: (the bulleted points indicate the documents from each phase phase):

First phase: The researchers of the NORC team will examine the objectives of the impact evaluation, the general indicator supplied in the reference terms, as well as the additional data of objectives of the program and indicators supplied during the startup of the project in January 2010, in order to create an initial glimpse of the questionnaires intended for the farmers and the OPAs:

- Questionnaire plan: the NORC team will produce a plan of survey documents tools. The forms will include an information sheet on the sample, a consent sheet as well as forms adapted for evaluation, specifically demographic data, ownership of the land, farm production, farm revenue, household income, social participation, the intervention form of the program in the debriefing of the interviewers.

Second phase: The researchers of the NORC team will examine the existing data gathering tools that were used for similar purposes with the same population or with a similar population and will insert validated elements of the questionnaire in the appropriate sections. The NORC team will examine the elements adapted to the objectives of the research and to the evaluation indicators.

[5] The Impact evaluation is the main research objective, the performance evaluation of the intervention program is an important corollary activity.

- First preliminary version (mainly in French, with a little English): the forms may include, in addition to the initial sections, the production of olives, employment in the farming sector, marketing, and establishment of olive and olive oil prices and other byproducts, the training, OPA (adapted from the "Social Participation" form)

Third phase: the local data gathering partner (C&O Marketing), the local expert consultants, the monitoring and evaluation experts of the APP, as well as local sector experts of the Ministry of Agriculture will examine the preliminary version and suggest revisions.

- Second preliminary version (in French): it will include updated forms and items and reflect the comments of the experts and other stakeholders; a preliminary version in Arabic will be created in order to be used in the pre-test.

Fourth phase: that will include the cognitive and pre-test. The cognitive tests will cover information related to indicators and will be conducted with a sampling of convenience in one or more rural regions; it will preferably include subpopulations such as minority linguistic groups and farmers. The pre-test of the complete questionnaire will also be performed with a sampling of convenience in one or more rural regions; it will preferably include subpopulations such as minority language groups and women farmers.

- Report on the cognitive tests: it will include results from the cognitive tests as well as suggestions for improving terms and orders used.

- Report on the pre-test: it will include pre-test results as well as suggestions for improving terms and orders used, as well as methods of gathering data (e.g. contact strategy, obtaining of cooperation)).

Fifth phase: review of reports on the pre-tests by the NORC team and by its experts. X review of questionnaires for the pilot.

- Pilot questionnaire (version in two languages, French and Arabic).

Sixth phase: pilot test of questionnaires updated and systems. The pilot test will be conducted with a sampling of convenience drawn from the main sampling; it will include the test for making contact, obtaining cooperation and entering data.

- Report on the pilot test: it will include results of the test pilot as well as suggestions for improving the terms and orders used, as well as data-gathering methods. Pilot tests will be entered, produced and subject to examination.

Seventh phase: Final review of questionnaires as a function of the report on the tests on the pilot test, as well as a review of the pilot data by the main team and the experts of NORC

- Final questionnaires.

Lessons drawn from the collection of the reference data and details of the program obtained from persons responsible for the implementation of the intervention make it possible to make adjustments to the survey tools for the collecting of follow-up data during the 2nd, 3rd, and 4th years. The updates of the questionnaires during the follow-up years will include an abbreviated version of the above process.

What does a cognitive test consist of?

The cognitive test is an essential phase of preparing the survey tool (Willis, 2005). The objective of this test is to know how the person interrogated understands the question asked, finds the information and comes up with their answer. It involves a structured conversation during which the interviewer first asks the question to the person being interrogated, obtains an answer then asks him/her several follow-up questions, called "probes".

The purpose is to discover problems related to the understanding of the question, the difficulty of recall or knowledge of the information being requested as well as the willingness to answer correctly. This type of test makes it possible to reveal hidden problems that do not appear during a simple pre-test of the document and that could result in an imprecise gathering of data.

There are different methods of using these probes. The most commonly used method will be made up of structured and spontaneous probes. Structured probes involve questions on possible problems that are anticipated; probes are written beforehand, then asked by all interviewers. Spontaneous probes are those that come up during the conversation as a reaction to answers or to nonverbal signs of the person interrogated.

Here are two examples of a question with structured probes. The probes are the questions in blue:

***For each of the categories of following expenses, can you please tell me what is the approximate amount of your food expenses for the past week*?**

• How did you calculate this amount?:
- Did you evaluate your expenses per day and then multiply them by 7?
- Or did you think about what you spend each time you go shopping and multiply that by the number of times or you went shopping during the week?
• What did you include in your expenses?:
- Did you include beverages?
- Did you include the amount of products that you farm?
- Did you include all of the expenses of your household?
• Was it difficult for you to come up with this number?
• Detect through gestures if the question seems to disturb the person being surveyed

How can you judge your [economic situation] during the past season?

• How have you understood this question?
• Can you say it in another way?
• Based on what indicators do you evaluate your economic situation?
- on the yield of the harvest (quality and quantity)?
- on the sale of cattle?
- on the compensation received by the family members?
- on the margin?
- on your savings?
- on your investment in the land? The intangible assets? …
• To judge your economic situation, do you compare yourself to your neighbors? To members of your (enlarged) family? To your situation from last year? With respect to the situation of the perimeter/of the *douar*?

Language and cultural considerations

The target populations of the evaluation include several cultural/linguistic groups speaking Moroccan Arabic (written language) and different dialects of Amazigh, more commonly called Berber (spoken language). Today, Berber speakers are estimated at approximately 40% of the total population of Morocco (31,000,000 inhabitants).

The *Amazigh* constitutes the oldest language in the country and in Maghreb in general. The Amazigh language is divided into dialect areas; it is especially used in rural or even mountainous regions, it is also increasingly used in the cities, after the rural exodus of the Amazighs and the urbanization of Amazigh speakers.

Amazigh is subdivided into three dialect groups (called tarifit, tamazight and tachelhit; the borders between the dialect areas are not always obvious.

1. *Tarifit* is a set of dialects used in the northeast region in general and in the Rif mountain chain in particular.
2. The *tamazight* variety is used in an area from the Jbel Saghro in the south to the Taza corridor in the north.
3. Tachelhit is used in a limited area in the north by a cartographic line connecting Essaouira to Tanant in the Azilal province, in the east and south along the course of the Dra and in the west near the Atlantic Ocean.

Map of linguistic areas in Morocco

(Source: http://www.tlfq.ulaval.ca/AXL/afrique/maroc-map-lng.htm)

To take these various linguistic groups in Morocco into account, questions will be translated from French into Moroccan Arabic. Moreover, Berber glossaries will be created as much as possible for the technical terms that farm farmers might not know in Moroccan Arabic. The translation method is described more in detail below:

The theory and practice with regard to preparing multilingual surveys have made considerable progress as the number of transnational surveys has increased over the past decades. According to good practices recommended for the translation of surveys and the adaptation methodology, the NORC research team will use a "target translation with two primary versions" ("double draft target translation"; Harkness *et al.*, 2010). This method uses a translation of the source document by independent translators; in the last phase, it can include either an arbitration by a third-party or a translation team. For survey documents, the first questionnaire source in French will be translated into local Arabic by at least two independent translators with experience in survey

preparation; their translations will be checked by a third person. A team discussion will make it possible to prepare the final versions. As the questionnaires are prepared, they will proceed by the pre-test phase, translations will be updated and rechecked. The main bilingual researchers of NORC will conduct an additional check of the last versions of the questionnaires in French and in Arabic. The survey tools will then be handed out in Arabic with a copy in French for coding.

Whenever possible, conceptual and linguistic differences among the four main groups will be considered in the preparation of the survey documents. Considering that it is not possible to supply a complete translation of the questionnaires into Berber dialects, technical glossaries of Berber terms spelled phonetically will be supplied if possible to the interviewers gathering data from persons interrogated in Berber. Interviewers will be trained to determine the fitness of persons interrogated to understand and respond in Moroccan Arabic and if possible bilingual interviewers will be assigned to persons interrogated in Berber in order to verbally translate the questionnaire into Arabic, by using the Berber glossaries for technical terms. Interviewers will be requested to indicate on the questionnaire the language that was used and to indicate when debriefing the type and location in the questionnaire where any problems of understanding or translation as well as the way the difficulty was resolved.

B. Ground work

During phase I of the project, the evaluation team will conduct a cognitive test of the questionnaires, as well as the pre-test and a pilot test. During phases II and III of the project, the pilot will again be conducted to test any changes in the questionnaire and ensure the proper implementation of data gathering each year. The following sections describe each of these phases in detail.

Pre-test and cognitive test

A pre-test as well as cognitive tests will only be provided for phase I of the project (2010). In order to save time, the pre-test will be conducted in parallel with cognitive tests in the space of one week. The pre-test and cognitive tests will be conducted in the same zones as the main data collection, in order to be able to receive answers from a sample that will be as similar as possible to the main sample. They will therefore be conducted in the 4 zones of the rehabilitation program of olive tree plantations and in the regions listed in the table below:

Activity	Population	Zone			
		1	2	3	4
Pre-test	Farmers	Larache Tetouan	Taounate Sefrou	Taza	Azilal Haouz
	OPA	Larache	Taounate	Taza	Azilal
	Olive processing plants	Larache Tetouan	Taounate Sefrou	Taza	Azilal Haouz
Cognitive	Farmers	Chaouen Sidi Kacem	Fes Khenifra	El hoceima	Essaouira
	OPA Olive processing plants	Chaouen Chaouen Sidi Kacem	Fes Fes Khenifra	El hoceima El hoceima	Essaouira Essaouira

For both pre- and cognitive tests, the planned sample is:

Activity	Population	Zone				
		1	2	3	4	Total
Pre-test	Farmers	5	5	5	5	**20**
	OPA	1	1	1	1	**4**
	Olive processing plants	2	2	1	2	**7**
Cognitive	Farmers	5	5	5	5	**20**
	OPA	1	1	1	1	**4**
	Olive processing plants	2	2	1	1	**6**

In order to have answers that are representative of the variability of answers that will be given during the main data collection, the pre-test and the cognitive tests will be conducted with respondents representing a diversified population. In other words, the planned survey will be given to:

1. Farmers and OPA representatives speaking Arabic
2. Farmers and OPA representatives speaking Berber
3. At least one woman farmer per zone (speaking different languages if possible)
4. Young and older farmers
5. Representatives from different types of OPAs (association, cooperative, etc.)
6. OPA representatives with members of the female sex
7. Directors of different types of olive processing plants (traditional, semi-modern, and modern)

Pilot test

The pilot test will be conducted after the pre-test and the cognitive test in order to test not only the questionnaire but also all logistics related to the gathering of data. It will then be a question of testing ground logistics (the setup of ground teams, supervision of interviewers, means of

transport on the ground), lines of communication between C&O, the APP and the Work Centers, obtaining cooperation of the persons being surveyed, transport of paper questionnaires, the establishment of the input mask and data entry. These phases outline the performance of the main data collection on the smallest scale and make it possible to determine any obstacles for resolving them before the start of the official data gathering.

C&O interviewers participating in the pilot will also be supervisors on the ground of the main data collection. Their participation in the pilot therefore makes it possible to establish an in-depth knowledge of the project that will enable them to better answer the questions of the interviewers when the time comes.

The pilot surveys will be distributed over the following provinces:

Activity	Population	Zone			
		1	2	3	4
Pilot	Farmers	Sidi Kacem Tetouan Larache Chaouen	Fes Sefrou Taounate Khenifra	Taza Al-Hoceima	Essaouira
	OPA	Sidi Kacem Tetouan Chaouen	Sefrou Khenifra	Taza	Essaouira
	Olive processing plants	Sidi Kacem Tetouan Larache Chaouen	Fes Sefrou Taounate Khenifra	Taza Al-Hoceima	Essaouira

For the pilot, the planned sample is:

Activity	Population	Zone				
		1	2	3	4	Total
C&O Pilot	Farmers	18	6	6	6	**36**
	OPA	2	1	1	1	**5**
	Olive processing plants	4	4	2	1	**11**
MAPM Pilot	Farmers	9	6	0	6	**15**
	OPA	3	0	0	0	**3**

Primary collection of data

The primary collection of data consists of three questionnaires and will be conducted each year for total of four data gathering cycles. Data will be gathered in the treatment, control and comparison perimeters and will be divided between the MAPM (statistics services) and NORC/C&O, as follows:

Phase I (Year 1): NORC/C&O will gather 70% of the data and will organize the logistics for this portion of the sample. The MAPM will be responsible for gathering 30% of the data and logistics related to this collection.

Phase II (Years 2 and 3): NORC/C&O will collect 30% of the data and the MAPM will collect 70% of the data. Each entity will be responsible for organizing the logistics related to its own collection of data.

Phase III (Year 4): NORC/C&O will collect 70% of the data and the MAPM will collect 30% of the data (similar to Phase I). Each entity will be responsible for organizing the logistics related to its own collection of data.

The division of the data gathering between NORC/C&O and MAPM will supply more reliable and better quality data over the four years[6], by taking into consideration the seasonal effect and the cyclical effect of the yield of olive trees, and of the effect of logistics and quality of the interviewers. Moreover, this will make it possible to reinforce the local capacities of Morocco in the ongoing use of best international practices.

The preparation and performance of the main data gathering will be done in several phases (this section is specific to the collection of data managed by NORC):

Desk Preparation:

Interviewers will be organized into several teams; each team having three interviewers and one supervisor. Each team will be assigned to a certain number of perimeters according to their geographic location so that the transport means are used effectively.

Before the arrival on the ground of the teams, the desk will alert the Work Centers (WCs) of the arrival of interviewers with the assistance of the APP in several ways:

a. A letter will be sent to General Affairs Office of the Ministry of the Interior in Rabat to inform them of the purpose and progress of the survey
b. A letter will be faxed to the DRA and DPA by the APP to communicate the dates of the survey
c. The APP will deliver the list of Work Centers (WCs) with their telephone number to C&O and each Work Center (WC) will be contacted by C&O
d. C&O will check the list of the farmers to be surveyed with each Work Center to find out about any problems

The local authorities will thus be completely informed of the project and the Work Centers may ensure the proper performance of the survey by assisting interviewers of C&O in locating farmers.

[6] This division between NORC/C&O and MAPM will be randomly done so the sub-samples are representative of the population of interest and to avoid any problems during the gathering of data.

Ground Activities:

For the primary collection of data, each supervisor manages a team of 3 interviewers. The duties of the supervisors are the following:

- Contact with the administration: local authorities, Work Centers
- Introduction to the project on the ground
- Preparation of the: numbering of the questionnaires, assignment of the survey participants to the interviewers and conducting interviewers with survey participants
- Quality Control and validation of questionnaires
- Debriefing of the interviewers
- Management of the tracking sheet/unit sheet
- Gathering and secure transporting of the questionnaires
- Link with the desk: alert of any problems, update of progress on the ground.

For each perimeter to be surveyed, here is the flow of activities on the ground:

Day 1: Upon arrival on the perimeter, the supervisor will contact the Work Center. The Work Center will explain the mission will request assistance in locating farmers and if possible, he will obtain their telephone number. He can also contact the local authorities (*caïd, pacha*) and the copy of the mission letter filed with the Ministry of the Interior in Rabat. According to the best dispersion of farmers in the perimeter, he will established a schedule of contacts as well as the routes to be taken. He can also conduct a reconnaissance to set up appointments with farmers for the following day. Then he will find lodging and meals that can be a basis for the perimeter.

Following days: The supervisor will drive the interviewers to each of the farmers or to the agreed meeting places and will make sure that the farmer agrees to the conversation before dropping off the second interviewer. He will go back to pick up the interviewer when he has finished the conversation (after a phone call from the interviewer) and will thank the farmer for his participation. He will pick up the questionnaire to keep it in the trunk of the car and thereby assure the secure and confidential transport of questionnaires at each gathering phase. At the same time, he can take advantage of trips between farmers to locate other farmers scheduled in the perimeter. Towards the end of the day, he will reread all questionnaires and will debrief the interviewers. He will update the tracking sheet and will send the sheet to the desk at the office of C&O. Lastly, he will prepare the next day (phone calls for meetings, outline the trips on the map, assignment of farmers to interviewers, etc.).

Tracking sheet

Each supervisor shall keep up-to-date a ground tracking sheet that will consist of listing the history of contacts made and questionnaires validated. Supervisors are required to contact the back office desk team each day to supply this update. This sheet will contain information enabling the supervisor to perform a daily tracking of the work, i.e.:

- Name of the supervisor
- Name of the perimeter/zone/region
- Name of the terms and Work Center responsible for the perimeter
- Name of the farmers
- History of the contacts
- Interviewer
- Questionnaire no.
- Duration
- Telephone no. of the survey participant
- Observations of the supervisor

Here is an example of tracking sheet:

Superviseur : Périmètre : CT: .....
Zone : Région :

	nom agriculteur	contact 1			contact 2			contact 3			FINAL			enquêteur	N°Qre	Durée			N°Tél	Observations
		date	heure	résultat	date	heure	résultat	date	heure	résultat	date	heure	résultat			début	fin	durée		
1																				
2																				
3																				

Supervisor:
Zone: Region: Perimeter CT (Work Center):
Name of farmer interviewer Questionnaire no. Duration Tel. no.
start end duration

Quality Control

Checking and coding

Toward the end of the conversation, the interviewer will fill out his report and will submit the questionnaire to the supervisor. A debriefing will be conducted with the interviewer to gather his feedback and any problems that the supervisor will note in his tracking sheet.

Each day, the supervisor will check all questionnaires to make sure the questionnaires have been properly filled out. He will check the thoroughness, coherence, compliance of the filters, the clarity, the legibility of the verbatim, and the no response coding. There are five cases that can explain no responses and that might then be codified differently:

- The survey participant refuses to answer (Code '-1')
- The survey participant does not know the answer (Code '-2')
- The survey participant is in a situation where the question is not applicable to him/her (Code '-8')
- The question was not asked due to a filter (no special code on the questionnaire)
- The interviewer forgot to ask the question (illegal skip, Code '-3').

In the latter case, the supervisor should immediately review the missing answers with the interviewer and correct the mistakes while the survey is still fresh in the memory of the interviewer. If it is not possible to correct these mistakes, the illegal skip will then be noted by the supervisor.

Validation

The supervisor will also be responsible for validating questionnaires. The proportion of valid questionnaires to be validated is set at 20% of the questionnaires. The very first questionnaire must be validated and then, the remaining questionnaires to be validated will be chosen at random. For this validation, NORC will develop a validation script that will contain a certain number of questions to be validated (around 5 questions).

If there are major inconsistencies at the level of a certain validated questionnaire, then it may be 100% re-surveyed. Lastly, if it turns out that an interviewer failed his responsibilities (e.g., falsification of surveys) or if the data of his questionnaires have an error rate of more than 5%, then all questionnaires for the interviewer in question will be taken back.

Any problem linked to the quality and any incoherence will be the object of an ex-post transformation of the data, or a modification of the data collection for subsequent cycles. In conjunction with the survey specialists, the team leader responsible for gathering data will prepare a report on the ground work that will describe the agreements and procedures followed on the ground, as well as problems encountered and solutions implemented.

C. Ethical considerations for the ground work

NORC acknowledges - as fundamental institutional value - its responsibility to protect the rights and well-being of humans participating in the research project. Rigorous ethical guidelines involving research on behavioral and social sciences dictate the method of interaction of NORC and its subcontractors with those participating in the studies. The main guideline followed by NORC is the "common rule") i.e. a series of principles codified by the US Code of Federal Regulations[7]. The common rule is the result of a long and serious work that was carried out to establish international standards of ethical research, as they are included in the Nuremberg Code, the Declaration of Helsinki and the Belmont Report.

NORC requires the informed consent for all activities of data gathering, from pre-test activities to surveys on the ground. The protection of persons interrogated within the framework of the performance and impact evaluation project of the rehabilitation activity and intensification of olive tree plantations in rain-fed zones complies with the standards of NORC and with its understanding of the Moroccan laws on the protection of personal data. In particular, after having presented the object of the study and informed the interrogated person that he/she can refuse to answer the questions and end the conversation at any time, the interviewer provides him/her with the chance to ask questions. This is only after an interrogated person has given their consent for the conversation to begin. All paper copies of the questionnaires are handled to guarantee their security; moreover, all personnel in contact with data of interrogated persons must take a training course on the confidentiality and sign the ethical declaration of NORC (see Annex D). NORC also continues to protect the rights of the research subjects after the end of the study by separating the personally identifiable data from response data, by maintaining strict security of all personal data and by destroying the data at the end of the study.

[7] Title 45–Well-being of the public, Part 46–Protection of human subjects, Sub-part A, Base policy of the HHS for the protection of human research subjects

Questionnaires, consent texts, documentation on recruitment and descriptions of research agreements are examined by the ethics committee of NORC before approval to start the research activities is issued. In the event of substantial violation to the agreement, the project director of NORC will inform the ethics committee and the APP; an action plan will then be prepared in conjunction with APP to mitigate any negative effect.

D. Data Management

Management tools

The input mask will be developed with the SPSS software. It will be placed directly on each questionnaire with the necessary filters. A dictionary of data will be created, making it possible to know the exact number of variables and the nature of these variables. This dictionary of data will be an integral part of the meta data concerning the evaluation project.

Moreover, a relational database will be created. In fact, our survey involves three different but inter-relational populations: the olive tree producers, the representatives of Farmer Associations and the representatives of olive processing plants. Since several farmers can be operators in several perimeters, and OPAs are linked to perimeters, and several perimeters can be linked to one or more olive processing plants, NORC and C&O will develop a relational database that will link the three populations and will allow the analysis of data and meta data associated with these cases.

Entry and cleaning of the data

Data entry will start as soon as the first questionnaires arrive at the back office desk of C&O. All questionnaires will be entered two times by different data entry personnel, and when a difference arises, a third person – the data entry supervisor – will be the arbiter and will make the final decision concerning the answer to be kept. Rates of error for each data entry employee will be recorded and delivered to NORC each week within the framework of the quality control process. If the error rate for a data entry employee is greater than 3%, all questionnaires for the employee in question will be reentered. The cleaning of these data will follow the entry and the unification of the data according to predetermined rules.

The gathering of data is performed by two different entities – C&O Marketing and MAPM, and each entity will then be responsible for entering its own data. Thus, during phase I of the project (2010), C&O Marketing will be responsible for gathering, entering 70% of the data while MAPM will be responsible for 30% of the data. During phase II of the project (2011 – 2012) these proportions will be reversed and MAPM will be responsible for gathering and entering 70% of the data while C&O will be responsible for 30% of the data. During phase III of the project (2013) the proportions will go back to those of the phase I.

Merger and control of the data

After each phase of gathering and entering data, MAPM will deliver its database to C&O Marketing. C&O Marketing will then be responsible for merging the two databases in a coherent and analyzable set. This merged database will be delivered to NORC for verification and statistical analysis. Before the analysis, NORC will examine the data to determine if there are aberrant observations and missing data and to judge their effect on the analysis and to take this into account before or during the analysis.

Meta-data

In addition to the data collected from structured questionnaires and qualitative conversations, NORC will attempt to document all phases of the evaluation project and of the survey cycle in a structured manner. These meta data will be an integral part of the quality control process of the project and will assure the proper use of the data by other analysts. Notably, these meta data will be available at several levels:

1. Development of the questionnaires: reports of the pre-test, cognitive test and pilot test.
2. Training: training plan, training manuals, list of interviewers and supervisors.
3. Survey baseline: cleaning rules.
4. Sample: methodology and statistical model used for the sample, rules of managing the sample on the ground.
5. Ground work: tracking sheets, rules for coding the questionnaires
6. Entry and cleaning of the data: weekly entry reports indicates error rates, data cleaning rules.
7. Dictionary of the data

VII. Calculation of the Indicators

In our questionnaire, intended for olive producers, there are variables (questions) involving farm income and other indicators of interest of the evaluation. Among these are the expenses in farm inputs, such as manure, transport and labor, and the revenue from the sales of olives and other farm products. Our objective is to calculate these indicators and their differences between the reference year and the final year and between the treatment and control/comparison groups in order to determine whether or not an improvement was recorded. To this end the following approaches will be adopted:

Approach 1: Univariate processing

This involves comparing the percentages the various indicators between two years and see whether or not it is possible to accept the hypothesis of no change or on the contrary if there was a difference between percentages. In other words, to see whether there is independence between the criterion that is in rows: (the two years) and the criterion that is in columns: the different values or methods of our question. For example, it is possible to see if the production of olives has improved or not.

This analysis should not be underestimated because it makes it possible to see, question by question, if the fact of benefiting from the intervention has a significant effect or not. This will enable us to see the indicators that changed from those that did not change and possibly explain the situation.

Approach 2: Use of grades (or scores) for qualitative variables

This approach consists of transforming each of the qualitative variables into a new ordinal variable with a minimum score and a maximum score, i.e., on a scale of 0 to 10. The type of "in the open" storage can be graded with a zero and the storage on a thin layer in a box with perforated walls with a 10, for example. And so forth for the other variables.

Once each of the variables has been codified (transformed into a score), it can be integrated into a global index. This synthesis index can be summarized in a simple mathematical average if it is judged that the importance of the variables is identical, but a weighted mathematical average can be used by giving more weight to certain variables than to others.

This approach 2 is a little subjective because it can depend on our grades. However, this disadvantage is negligible because what interests us is the relative effect of the evolution between reference year and the final year.

Approach 3: Multivariate Approach

In year 1, a multivariate analysis will be conducted on all variables with a connection to the farm income or other indicators of interest. This analysis is essentially based on qualitative variables and it is of the Multiple Correspondence Analysis (MCA) type. It resembles an analysis in main components since it creates principal non-redundant variables. The analysis of the first axis (by hoping that it will have a strong inertia) will probably make it possible to distinguish operations according to their level of income. We will consider the coefficients connecting this axis 1 (and possibly the second axis) because we will use them in the final year to recalculate the new values taken by this axis 1, but with data from the final year. In this way, it is possible to see if the income (and other variables of interest) of the farmer has changed (according to this axis) in the direction of improvement, in the direction of degradation or has stayed stable. [8]

[8] For the moment, this approach is arbitrary because these results are unknown beforehand since they will depend on the behavior of the axes from the analysis in the first year.

Comparison with other data

Lastly, if possible, our analysis will include a comparison with the pre-existing data. It might be proposed to use objective measurements on yields harvested by MAPM based on parcels worked by farmers in view of validating the accuracy of the estimated yields supplied by our respondents in order to make any necessary adjustments to the estimates.

VIII. Awareness raising and Reinforcement of Capacities

Since the beginning of the project, NORC has tried to contact the local entities such as the high commissioner to the plan, the statistical services of MAPM, the Agronomic and Veterinarian Institute, the TC–1A and TC–5A consultants and local experts to encourage collaboration among these various groups and to adapt the methodology to the context in Morocco all while integrating international standards.

NORC with the APP and the Statistical Services of the MAPM have identified an opportunity to strengthen capacities of the Ministry of Agriculture by statistics and survey methods training sessions, by leading collaborative seminars and by working together on research articles for potential publications in scientific journals. Several topics have been identified for these training sessions, for example quality of the surveys and data, standards of meta data and the analysis of data based on complex random samples.

These training sessions fall within the framework of the NORC mission to advance the discipline of social sciences by means of the exchange of free ideas with the global research community. NORC will develop presentations and workshops to meet the requests of the MAPM. The APP has also mentioned its interest in using the data and evaluation project results for purposes of publishing articles in international journals and conferences. The NORC research team will develop a collaborative plan of research with the APP for this objective.

IX. Bibliography

AIAM (Association des Ingénieurs Agronomes Marocains). "Secteur des organisations professionnelles agricoles (OPA), cas de la COMADER." *Revue HTE* no 139 (Mars-Juin 2008).

Arriagada, Rodrigo A., Erin O. Sills, Subhrendu K. Pattanayak, and Paul J. Ferraro. "Combining Qualitative and Quantative Methods to Evaluate Participation in Costa Rica's Program of Payments for Environmental Services." *Journal of Sustainable Forestry* 28, no. 3 (2009): 343-367.

Bessaoud, O. "Les organisations rurales au Maghreb: un essai d'évaluation de leur role dans le développement agricole et rural." *Communication séminaire SFER*, Novembre 2005.

Biemer, P., and L. Lyberg. "Introduction to survey quality." *Wiley series in survey methodology*, 2002.

Boukous, A. "Dynamique d'une Situation Linguistique: Le Marché Linguistique au Maroc." In *Rapport 50 ans de développement humain et perspectives 2025*. Rabat, 2005.

Bulmer, M., and D. Warwick. *Social research in developing countries.* Wiley & Sons, 1983.

Chaker, Salem. "Berber, a "long-forgotten" language of France." *Paper presented at the Conference on Language and Im(migration) in France, Latin America, and the United States: Sociolinguistic Perspectives, September 25-26, 2003, Austin, Texas.* September 25-26, 2003. http://www.utexas.edu/cola/insts/france-ut/_files/pdf/resources/chaker_english.pdf (accessed May 19, 2010).

Dachmi, A., and E. Henri Riard. "Adolescence et projet de vie chez les jeunes marocains." *Université Mohammed V, Essais et Etudes* no 40 (2004).

Daponte, B. *Evaluation essentials: Methods for conducting sound research.* Wiley & Sons, 2008.

Desrues, Th. "La nouvelle politique agricole et les associations professionnelles agricoles au Maroc." *Fifth Mediterranean Social and Political Research Meeting*, March 2004.

DP TC-5A. "Demande de Propositions: Services de formation, d'appuis techniques et d'encadrement aux beneficiaires, organisations professionnelles et operateurs de la chaine de valeurs des filières d'oliviers, d'amandier et de figuier."

DPV/MADRPM. "Rapport Annuel." 2009.

Fadlaoui, A., N. Bahri, and A. Bentaibi. "L'agrégation agroalimentaire, cas du secteur oléicole." *Pack info* no 87 (2010).

FIPA (Fédération Internationale des Producteurs Agricoles). "Role des cooperatives dans le developpement agricole au Maroc." *38e Congrès*, Mai-Juin 2008.

Groves, Robert M. "Survey Errors and Survey Costs." In *Wiley Series in Probability and Mathematical Statistics*, edited by Barnett et al. New York: John Wiley & Sons, 1989.

Guo, S., and M. Fraser. "Propensity score analysis: Statistical methods and applications." *Advanced quantitative techniques in the social sciences series*, 2010.

Harkness, J., F. Van de Vijver, and P. Mohler. "Cross-cultural survey methods." *Wiley series in survey methodology*, 2002.

Harkness, Janet , et al. "Translation, Adaptation and Design." In *Survey Methods in Multinational, Multiregional, and Multicultural Contexts. Wiley Series in Survey Methodology, Couper et al., (series eds.)*, edited by Harkness et al. Hoboken, New Jersey: John Wiley & Sons, 2010.

Harkness, Janet. "Comparative Survey Research: Goals and Challenges." In *International Handbok of Survey Methodology*, edited by Edith de Leeuw, Joop Hox and Donald A. Dillman. New York: Lawrence Erlbaum/Psychology Press, Taylor & Francis Group, 2008.

Haut Commissariat au Plan. "Pauvreté, développement humain et développement social au Maroc." *Données cartographiques et statistiques*, Septembre 2004.

INRA. "Identification et caractérisation des principales filières et analyse des systemes de production, de valorisation et de commercialisation et proposition de voies d'amélioration et mesures d'accompagnement." *Projet DRI-PMH*, 2005.

Kerzali, M. *Migration et développement au Maroc.* Rabat: Université Mohammed V, 2003.
—. *Migration et Développement au Maroc.* Rabat: Université Mohammed V, 2003.

Khuel, R. *Statistical principles of research design and analysis.* Duxbury Press, 1994.

Leeuw, E., J. Hox, and D. Dillman. *International handbook of survey methodology.* European Association of Methodology, 2007.

Lohr, S. *Sampling: Design and analysis.* Duxbury Press, 1999.

Lynn, P. *Methodology of longitudinal surveys.* Wiley & Sons., 2009.

MCC-Morocco Compact. 2007. http://www.mcc.gov/mcc/bm.doc/compact-morocco.pdf (accessed June 30, 2010).

Ministère de la Prévision et du Plan. "Les statistiques sur la population marocaine." 2004.

Nah, O. "Analyse des représentations et des stratégies d'existence des jeunes ruraux en zones défavorisées de montagne (cas des Douars Ouaourkhsen et Bouhamed)." Mémoire de troisième cycle de l'ENA, Meknès, 2007.

ODCO (Office du Développement de la Coopération). "Annuaire statistique des coopératives au Maroc." 2008.

ODCO (Office du Développement de la Coopération). "Les coopératives de jeunes au Maroc: Etat des lieux." *Etudes coopératives* no 10 (avril 2002).

Plan Maroc Vert. "Pilier II du Plan Maroc Vert. De la stratégie à l'action." *Conseil Général du Développement Agricole*, 2009.

Quitout, M. "L'arabe, le francais, l'amazighe au Maroc: un patrimoine culturel national." *Développement linguistique: enjeux et perspective, Cahier du RIFAL*, 2001: 60-65.

Rosenbaum, P. "Observational studies." *Springer series in statistics*, 2002.

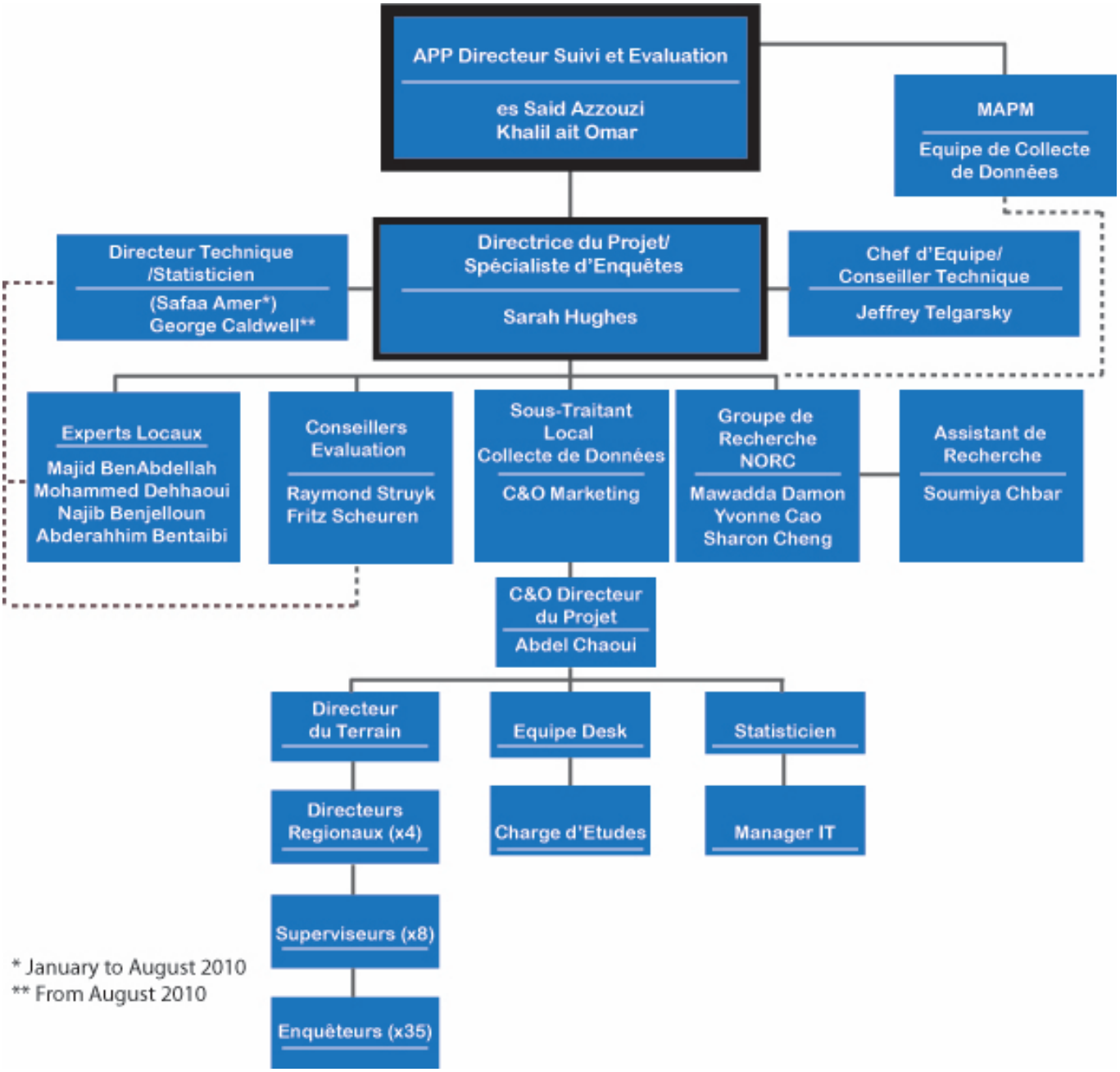
Rossi, P., M. Lipsey, and H. Freeman. *Evaluation: A systematic approach.* Sage Publications, 2004.

Udry, C. "Fieldwork, economic theory, and research on institutions in developing countries." *American Economic Review* 93 (2003): 107-111.

Willis, G. *Cognitive interviewing: A tool for improving questionnaire design.* Sage Publications, 2005.

X. Annexes

A. Annex A: Organizational chart - Composition of the Teams



		APP Follow-up and Evaluation Director		MAPM Data gathering team
Technical Director/Statistician		Project Director Survey Specialists		Team Leader/Technical Consultant
Local experts	Evaluation consultants	Local Data Gathering Subcontractor	NORC Research Group	Research Assistant
		C&O Project Director		
	Director on the Ground	Desk Team	Statistician	
	Regional Directors (x4)	Study supervisor	IT Manager	
	Supervisors (x8)			
	Interviewers (x35)			

B. Annex B: Action Plan

Activities	1st Phase (December 2009 -December 2010)	2nd Phase (January 2011 – December 2012)	3rd Phase (January 2013 August 2013)
Phase 1: Design of the evaluation and survey of the reference situation (December 2009-December 2010)			
Overall and preparation examination: examination of the Compact, of the program, of the general and sector documents, internal organization and preparation of the team, preparation of the first trip to Morocco.			
Preparation and submission of the work plan for the trip inside the country.			
First visit to Morocco (two to three weeks): meeting with the APP team, implementation organizations and S&E teams of MCC, joint with the C&O Marketing to define the terms and characteristics of the data gathering activities.			
Preparation and submission of the initial report of the preliminary design draft.			
Implementation of the randomization strategy, notably the new estimate of the size of the samples, if necessary, and the selection of samples.			
Preparation of the survey documents for reference surveys of farmers and OPA, submission to APP for examination, cognitive test and pre-test of the questionnaires, integration of comments in the final documents, preparation of ground agreements, preparation and submission of the training plan of the interviewers.			
Pilot test/finalization of the questionnaires of the reference survey,			
Preparation and presentation of the definitive design report			
Training of framers and interviewers of MAPM and of C&O Marketing			
Performance of the ground work for the reference survey with farmers.			
Development of a relational database			
Completion of the data entry for the reference survey with farmers, of the cleaning, quality control and of the data documentation,			
Drafting of the survey documents for the reference surveys of the olive processing plants, submission to APP pour examination, pre-test of the questionnaires, inclusion of the comments in the final documents, preparation of the ground agreements, preparation and submission of the training plan of the interviewers.			
Training of the framers and interviewers of MAPM and of C&O Marketing			
Performance of the ground work for the reference survey with the OPAs and the olive processing plants.			
Completion of the data entry, of the cleaning, of the quality control and of the data documentation; delivery of sets of cleaned data to APP			
Prepare and submit to APP the survey report of the reference situation, after an analysis of the data.			
Written approval of the deliverables of the Phase 1 and authorization to continue to phase 2			
Phase 2: Annual follow-up surveys and collection of performance data (January/December 2011 and 2012)			
Drafting of survey tools, adaptation of survey agreements and ground procedures, if necessary			
Training of framers and interviewers of MAPM and of C&O			
Conducting of the surveys on the ground by MAPM and C&O Marketing, entry, cleaning and quality control of the data			
Drafting of the qualitative conversation documents for the performance evaluation of the implementation activities, selection of participants, preparation of conversation agreements.			
Holding of qualitative conversations of resource persons			
Reinforcement of the capacities of the Statistics Service of MAPM; training and workshops			
Drafting of annual follow-up reports			
Phase 3: Final surveys, summary of performances and the impact evaluation report			
Pilot test/completion of the final survey questionnaire			
Adaptation of survey agreements and ground procedures, if necessary			
Training/refresher courses of framers and interviewers			
Performance of the ground work for the final impact evaluation survey and the interview of resource persons			
Completion of the entry, the cleaning, the quality control and the data documentation; overall delivery of cleaned data to APP			
Temporary report include the results from the performance and impact evaluation			
Preparation and submission to APP of the final performance and impact evaluation report			
Verbal presentation of findings from the impact evaluation			

C. Annex C: Logical Framework and Indicators

	Indicator
Overall goal of the Fruit Tree Agricultural Project (PAF)	
Reduce poverty among rural households located in rain-fed zones.	Socio-economic status of households
	Change of activities and jobs of the households
Program goal	
Increase in income of target farmer households	Net income of farmers
	Net income of olive oil activities of farmers
	Average farm income per beneficiary operation of the project
Program objectives	
Increase in yield from olives	Yield of olive orchards in rain-fed zones in tons/hectare
	Production of olives in tons
	Yield per tree
Increase in oil produced	Oil production in liters
	Rate of oil produced per 100 kg of olives
Increase in the production value	Value of the olive production
	Value of the oil production
Increase in sales	Percentage of production sold
Long term outcomes	
Better health of trees	Number of farmers who have had diseases, insects, wounds, parasites on their olive trees
Better quality of olives	Comparison with the previous year
Better quality of olive oil	Percentage of the production that is virgin and extra virgin
	Percentage of the oil certified organic
Better price obtained for sale of olives and oil	Average of prices paid (price per kilo)
Better purchasing price of production inputs	Average of prices paid
Increase in clients for olives and olive products	Number of people/entities to whom the production is sold
Development of small businesses with regard to size, plant disease treatment, working the soil fertilization, harvest	Number of young people employed by small businesses
Immediate Outcomes	
Use of good growing practices	Percentage of farmers who have adopted the new intensification techniques
	Percentage of farmers who use good farming practices
OPAs are operational	Percentage of operational OPAs (with regular meetings and activities)
	Management quality
Development of partnerships with transformers and distributers	Percentage of OPA with partnership or collaboration with olive processing plants, large businesses for olives or large businesses for olive oil
	Percentage of OPAs that organize and help with the shipping cost
Creation of second-level organizations	Percentage of OPAs that are a part of a partnership/collaboration with other OPAs
Betters price information	Frequency information is obtained regarding prices and quality of the information
	Existence of communication of the information with farmers
Use of standard quality and high quality oil production systems	Number of olive processing plants using a standard quality and high quality oil production system
Increase in marketing activities	Number of OPAs that take part in trade shows, international competitions and other marketing opportunities
Olive processing plants adopt a business plan	Percentage of olive processing plants that benefit from training and technical assistance and have adopted a business plan
Local experts available for consultation	Knowledge level of young trainees
Completion of activities (Outputs)	
Farmers are trained	Number of trained farmers
Farmers receive technical support	Number of farmers who received technical support
Farmers adhere to the project	Number of farmers within the project
OPAs are founded	Number of OPAs founded
	Number of women OPAs founded
OPAs are trained	Number of trained OPAs
OPA receive technical support	Number of OPAs that received technical support
OPA adhere to the project	Number of OPAs within the project
Price information system set up	Number of OPAs that stay informed of prices and that communicate them to farmers
Valuation and marketing operators are trained	Number of trained operators
Operators receive technical support	Number of operators that have received a technical support
Operators adhere to the project	Number of operators within the project
Training in marketing	Number of persons trained
Young men are trained as experts	Number of young persons trained
Pilot projects of women are carried out	Number of projects completed

D. Annex D: Professional Ethics Code and Confidentiality of the data

Data confidentiality is an extremely important subject for the National Opinion Research Center. Making sure that the respondent understands that his/her answers will be strictly confidential will help the interviewer obtain answers from the respondent who will be more likely to participate in the survey. Respecting confidentiality also maintains the reputation of NORC without which NORC could no longer conduct its research work.

What the respondent must understand:
Our survey will ask respondents to give us information concerning their income and expenses. This involves delicate information, it is therefore of the utmost importance for the respondents to understand the following:

1. Their participation is voluntary and they can refuse to answer any question they don't want to answer, but it is hoped that will answer as many questions as possible.
2. Information on the identity of the respondent will NEVER be used for purposes other than the conducting of this survey.
3. Information on the identity of the respondent will be electronically separated from the answers of the respondent and will be destroyed at the end of the study.
4. The data will be used solely for statistical purposes and any person compromising the confidentiality of the data of the respondent will be subject to prosecution.

The role of NORC in maintaining confidentiality:
NORC has maintained the confidentiality of its data for over 60 years and is respected for its conduct vis-à-vis study participants. It should be understood that:

1. NORC does not share any data on the identity with persons outside of NORC, this information is destroyed at the end of each study.
2. NORC reviews the data that could identify the respondents (names, addresses and telephone numbers) and replaces them with generic labels that secure the confidentiality of the personal information.
3. The data on the identity are removed before the questionnaires are delivered to the client (here MCA Morocco).
4. NORC's data security system complies with the specifications of the Federal Information Processing Standards (FIPS).

The role of the interviewer in maintaining confidentiality:
The interviewer plays an important role in maintaining the confidentiality of the data. Here is what the interviewer can and cannot do:

1. DO NOT write the names, addresses and telephone numbers in places other than those specified in the questionnaire.
2. DO NOT keep documents that have data on the identity in a place to which other persons might have access.
3. DO NOT authorize any person outside of NORC and this survey to see the data on the identity of the respondent.

4. DO NOT discuss the subject of the respondents with any person other than the NORC personnel of this survey.
5. DO NOT hold conversations in a public place or a place where others could hear the answers of the respondent.
6. DO NOT use your personal email for sending or receiving information related to the identity of the respondents. If you must use your personal email, refer to the respondent in general terms only.
7. USE generic terms such as “respondent”, “wife of the respondent”, etc.
8. KEEP documents that have information on the identity in a secure place to which other persons do not have access.
9. INFORM the respondent of your promise of confidentiality and DESCRIBE the steps taken by NORC to ensure the confidentiality of the data.

Other requirements for the interviewer:

1. DO NOT hold a conversation with a person whom you know personally without prior approval.
2. DO NOT influence the answers of the respondent by criticizing or approving the answers of the respondent.

Responsibility of the subcontractor (C&O):

The subcontractor must also respect all of the above instructions and regulations.

Promise of confidentiality:

Before beginning the survey, any interviewer must sign a promise of confidentiality. A copy is supplied below.

Professional Ethics Code of NORC

Promise of Confidentiality

(Each employee of NORC and each visitor in the activity zones of NORC must read the professional ethics code above and sign this agreement as a condition for being hired or granted access.)

I have read the attached professional ethics code of NORC and understand that these obligations apply to my person and are a condition for employment by NORC and/or access to research documents.

I respect the professional ethics code of NORC described in this declaration. If I do not respect this code, I understand not only the possibility of legal prosecution, but also of dismissal if I am an employee, or of refusal of access and additional cooperation if I am a visitor, subcontractor or researcher.

I understand that this code and this promise, concerning the confidentiality of the respondents of NORC, of the documents and internal procedures, are applicable during and after the duration of this job.

NAME: _____

SIGNATURE: _____

DATE: _____

Please keep a copy of the ethics code and return this signed promise of confidentiality to [Organization].

E. Annex F: Analysis of statistical power of a sample

1. General Discussion

There are two basic approaches for determining the size of a sample. For a descriptive survey, the sample size is determined to supply a desired level of accuracy of the estimates of interest (descriptive statistics, such as average, proportions and total of a population or of subpopulations). For an analytical survey, the sample size is determined to supply a desired power level of hypothesis tests, such as the test of the hypothesis that the impact of a program exceeds a given value (such as zero or a desired minimum level). Since this project involves an analytical survey, the latter approach is appropriate for determining the sample size or evaluating the accuracy of the size of a specific sample.

At the time this report was submitted, the sample size was already determined for the baseline survey of the impact evaluation – it involves approximately 1,600 households in the 67 treatment perimeters and 1,600 households in the 67 control perimeters. The question of accuracy of the sample size must be raised. This section involves this question.

The power of a sample to detect the effects of a specified size depends not only on the sample size but also on the design of the sample. In addition to the evaluation of the correctness of the current sample size for the design of the planned sample for follow-up surveys, this section examines an alternative to this design (more particularly, the question of the substitution of a portion of the panel in each wave of the survey).

The sample size of a study, such as an impact evaluation implies the consideration of several factors, including the following:

- What is the available budget?
- What are the objectives of the analysis (e.g., a specified level of precision for estimates of the double difference or a specified level of power (probability) for hypothesis tests, such as the detection of an effect (impact) of a specified size, using tests of a given level of significance)?
- What is the nature of the variability of the populations of interest (variances or standard deviations of variables of interest and correlation among the principal sample units)?
- What are the sample costs (for first-degree and second-degree sample units)?

In the case of a second-degree sample, there are two sample sizes of interest – the first-degree sample unit and that of the second-degree sample unit (households) within the units of the first degree.

The following paragraphs describe this process in detail.

2. Analysis of statistical power of a baseline sample

The approach used to estimate sample sizes in the case of analytical surveys is called "analysis of statistical power". With this approach, the sample sizes are determined so as the probability or detection power of an impact of a specified size meets or even exceeds a specified level. In this application, we will estimate the sample sizes necessary to obtain levels of power from 90 to 95%.

The sample size required to obtain a specified power or the power associated with a particular sample size depends on the number of suppositions made on the design of the evaluation, the estimate of the impact, the design of the sample survey used to gather data, the statistical test as well as the population that is the object of the study. The following paragraphs address these suppositions, specify the values of the various perimeters of interest, indicate the sizes and powers of samples corresponding to the suppositions and values of the perimeters.

Measurements of results. The evaluation consists of determining the impact of the interventions of the TC–5A program on the economic status of the households in the project zones. There are several results of interest associated with this project, such as income, employment and production but income is the most interesting variable. Since the statistical characteristics of these various measurements of results differ, the sample size required for the evaluation of impact on them or the power associated with the specified sample size will be different. In order to estimate the size with the power of the sample, we will focus on the estimate of the project's impact on income. The reason for this focus is not only justified by the fact that it involves the measurement of the most interesting result but also because a considerable quantity of knowledge is available from previous studies of the impact on income and it is also possible to indicate reasonable informed estimates for various perimeters affecting the size and the power the sample for this variable.

Design of the evaluation and measurements of the impact. The main objective which in fact is the ultimate objective of the survey proposes evaluating the increase caused by the program on the economic well-being (income, production and employment). For the evaluation of the impact, the basic question is "to what extent did income or another measurement of result of interest change (upward or downward) due to the program?". The critical question is however "compared to what?" It is not enough to simply observe that the income has increased by a certain quantity during the program, since this increase could have occurred in the absence of the program. The increase resulting from the program can be the same increase experienced in the entire country. In view of evaluating the impact of the program, it is desirable to compare the measurements of result within the program's framework to those that could have occurred in the absence of the program. This hypothetical construction is qualified as "alternative scenario". The problem that evaluators must overcome is the fact that it is impossible to directly observe an alternative scenario for a physical sample unit, insofar as each sample unit receives the treatment (services of the program) or does not receive it. The effect of the program must in fact be observed by indirect means (such as by means of a regression model).

To obtain a good estimate of the impact of an intervention of the program on treatment units, it is necessary to compare the changes made in the treatment units compared to changes in similar untreated unit. The most worthwhile method of evaluation of the causal effect of a program consists of using an experimental design, whereby the experimental units are selected at random among the target population and the randomization is used to assign the treatment level (in this case, i.e. treatment or not treatment [control]) for each unit selected. By using randomization, the combined distribution of all variables (except treatment) is the same for the treatment and control groups and it is possible to obtain an unbiased estimate of the effect of the program based on these data. A so-called ideal design for the evaluation is the "randomized group pre-test post-test control" design that includes four groups: pre-test treatment group (time 1); post-test treatment group (time 2); time 1 control group and time 2 control group. With this design, the standard measurement of the program's impact is the double difference: the difference between the treatment groups and control groups, of the difference on averages (of the income, employment and other variable of interest) at time 1 (pre-test) and time 2 (post-test). With this design it is useless to estimate the alternative scenario results, since the causal effect is directly evaluated.[9]

In the case of the current application, the perimeters that must benefit from the program's intervention were determined by randomization. The design was that of matched pairs, in which the total proportion of 140 perimeters was divided into 70 pairs and one member per pair was assigned as treatment (the number of pairs was slightly reduced from 70 to 67 for various reasons). The pairs were matched by calculating the *Mahalanobis* distance between each pair and by putting together perimeters that were close to each other. (The *Mahalanobis* distance represents a weighted amount (square form) of the differences in the matched variables where the weighting matrix represents the variant/covariance matrix of the matching variables. The matching procedure is described in the *Echantillonnage randomisé des périmetres de traitement et de contrôle en perspective de l'évaluation de l'impact de la réhabilitation de l'olivier dans les zones pluviales (ME5-B): Report final*, B. Crépon (Crest and J-PAL), L. Behaghel (EEP, Crest and J-PAL), October 2009.) The matching was made to optimize the accuracy of the estimates of the impact of interest that the differences are (between the treatment group and the control group).

Design of the survey. The data required for the construction of the estimate of the double difference of the program's impact are gathered using a sample survey. The data obtained from the survey are used to produce statistical estimates of the variables of interest. The accuracy of these estimates depends on the type of design of the sample, the nature of the population surveyed and the type of estimate. The two components of the accuracy are precision (reproducibility) and bias (systematic error). A standard measurement of precision is the variance of an estimate or the square root of the variance that represents the standard deviation. The standard measurement of precision is the average square error that is equal to the variance plus

[9] The significance of the double difference can be tested using a standard T or Z test for which the standard error of the estimate takes into account the nature of the design. In most cases, the design includes the correspondence over time (by means of a panel survey, in which the same household is interviewed at time 1 and at time 2) and that of the treatment and control units. In other words, the design involves double matched pairs and the deviation of the T or Z test is calculated from *n* paired quadruples where *n* represents the sample size for each of the four groups). (it should be noted that within the framework of the experimental design, it is useless to observe the pre-test. Insofar as the randomization brought about the equivalence of the treatment and control groups, a single difference)of averages between the treatment and control groups is not biased).

the square of the bias. Since the bias is generally unknown and complex to estimate, the focus on the design of the sample is generally placed on the accuracy.

There are two main types of sample survey designs: "descriptive" designs that are used to obtain estimates of overall characteristics of the population, such as population totals, averages and proportion; and the "analytical" designs that are used to estimate the relationship between the dependent variables and the independent variables (explanatory variables) and to support the hypothesis tests. The impact evaluations mainly involve the estimate of relations (of the impact on the explanatory variables) and on the performance of the hypothesis tests, such as detection of an impact of a given size and are implemented using analytical survey designs. The design of descriptive surveys and analytical surveys is totally different. The design of a descriptive survey in principle is performed by building a design ending up with the specified level of accuracy for estimates of the overall characteristics of the population; the design of an analytical survey is generally conducted by building a design ending up with a level of probability or specified power, for tests of hypothesis of interest, such as a test on the estimate of the double difference of a program's impact.

In descriptive and analytical surveys, the characteristics of the design of the surveys such as stratification, sampling at several degrees in selection with variable probabilities are used. These techniques optimize the effectiveness of the design, i.e. the "return" of the precision or of the power with respect to the cost incurred.

The characteristics of the survey's design depend on the nature of the population of interest (e.g., its geographic distribution, such as the concentration of the human population in villages), the sampling costs (e.g., the sampling cost of the village or of a perimeter with respect to the sampling cost of a household within the village) as well as the type of data available on the population of interest before the survey (e.g., the previous survey data, from government statistical systems or vendors of geographic information systems). In the case of our baseline survey, the design of the survey is second degree in which the sampling of the perimeters is selected at the first degree and a sampling of households is selected at the second degree within each perimeter of the sample. The first-degree sample units (perimeters) are called Primary Sample Units (PSUs) and the second-degree sample units (households) are called ultimate sample units. Although we have treated the "sample size" as if it involved a unique concept, in a 2-degreesurvey, it in fact involves two samples of interest: the first-degree sample unit (PSU) and that of the second-degree (household).

Approach to determine the sample size. As we have indicated, the standard approach for determining the sample size for analytical surveys consists of estimating the size required to reach a specified level of power for detecting an impact of a given size. This approach for determining the sample size is called "statistical power analysis". We will estimate the sample size required for detect an effect (double difference) of a given size with a specified probability (power).

Parameters affecting the sample size. By using a two-degree sample design to gather data of interest, the following parameters affect the sample size:

1. Size of minimum effect being detected. (Absolute to relative with respect to the standard deviation.)
2. Variance of the measurement of result of interest. (Absolute or relative with respect to the average.)
3. Power of the statistical test.
4. Level of significance of the statistical test.
5. Correlations between the units in the different design groups.
6. Coefficient of correction among units.
7. Relative cost of the sampling of the first-degree and second-degree sample units.
8. Optimum second-degree sample size.
9. Effects of the sampling plan (stratification, sampling at several degrees [coefficient of correlation among units], selection with variable probabilities).

The following paragraphs specifically describe each of these elements, defining the parameters that measure them and then present a formula for determining the sample size according to the parameters.

Size of minimum effect to be detected; variance of a measurement of result of interest. The planned evaluation will gather data over a large number of dependent and independent variables (variables of result and explanatory variables). The optimal design varies according to dependent variable. In order for this problem to be manageable, the objectives are indicated for measurements of main interest results and a design is prepared making it reasonably it possible to reach these various objectives. After having examined the objectives of the project, the NORC team identified "design objectives" in view of guiding the development of the design of the impact evaluation. We will more particularly choose designs that can detect a change of income of 8% over time between the baseline survey and the final survey.

We will identify designs that can detect the effects of this breadth with powers of 90 and 95%.

After indicating the size of the effect, the most important parameter affecting the sample is the variation in the dependent variable (measurement of result). The standard measurement of variation is the variance or its square root, the standard deviation. Previous studies have supplied a good deal of information on the variation in income. It has particularly been observed that in the rural zones of developing countries, the rate of the standard deviation of income with respect to an average income (i.e., the deviation coefficient) is generally in the range of 0.5 to 2. Within the framework of this analysis, we will suppose the value of 0.5 but other values have also been examined (in a parametric analysis or power sensitivity analysis). This involves an optimized supposition corresponding to the hypothesis that the population being studied is relatively homogeneous.

Power and level of significance of the test A statistical test of hypothesis includes two perimeters: probability α or a Type-I error (stipulating there is a non-null difference when there is none) and probability, β, of a Type-II error (stipulating there is no difference, when there is one). Parameter α is called the level of significance of the test. The "power" of the test is 1 – β (i.e., the probability of deciding there is a difference, when in fact there is one). The standard values for α and β are 0.05 and 0.1. Within the framework of this analysis, we will define the level of significance (α) equal to 0.05 and the probability of Type-II error to 0.05 or 0.1, which corresponds to the values of 0.95 or 0.90 for the power (1- β).

Correlations among units in the different design groups. Within the context of a descriptive survey, one generally tries to have observations and sample subgroups without correlation, such as strata and sample units (such as first-degree sample units [clusters, primary sample units]). This condition is advantageous for estimating overall characteristics of the population, such as totals and averages. Within the context of analytical surveys, that involve estimating differences, it is advantageous in certain cases to introduce correlations in the sample. To do this, there are two methods that consist of using a panel sampling (in which the same sample unit is observed at time 1 and at time 2) and the correspondence of the treatment and control groups (such as the formation of matched pairs of the eligible perimeters and the random selection of a member in each pair for treatment).

In the pre-post-test design, there are six possible correlations from the standpoint of pairs among the four design groups. If we label the groups as 1: treatment before; 2: treatment after; 3: control before; and 4: control after, these six correlations can then be noted as ρ_{12}, ρ_{13}, ρ_{14}, ρ_{23}, ρ_{24}, and ρ_{34}. With the panel sample, we expect to observe a high level of correlation; in this case, the reasonable values for ρ_{12} and ρ_{34} can be 0.5. The breadth of the correlations between the treatment and control samples depends on our success in having the treatment and control units coincide. The algorithm used for the matching of the perimeters had 26 socio-economic and geographic variables which represents a substantial figure. On one hand, the match is made at the perimeter level and not at the household level. This is why the correlation between the treatment and control units will not be very high. Reasonable values for ρ_{13} and ρ_{24} can – let's say - be 0.3. The two other remaining correlations, ρ_{14} and ρ_{23}, will certainly be slightly lower, let's say 0.2. We will suppose these values for this analysis.

The initial plan of the follow-up surveys consists of keeping only 25% of the initial sample as a panel for the duration of the project. The values of the previous perimeters correspond to the supposition that the total base survey is considered as the panel for all of the remaining years. If the panel is uniquely defined at 25% for the baseline survey, the correlations introduced from the panel sample will be considerably reduced. In this case the reasonable values for the correlations can be $\rho_{12} = 0.1$ $\rho_{13} = 0.3$ $\rho_{14} = 0.05$ $\rho_{23} = 0.05$ $\rho_{24} = 0.3$ and $\rho_{34} = 0.1$.

Correlation coefficient between units; sampling costs; optimal second-degree sample size; effect of the sample plan. At this state, we have taken into account the size of the effect; of the variance; of the level of significance and of the power as well as the correlations between the groups. After these parameters, all remaining parameters stipulated above effect the sample size by a composite measurement called "effect of the sample plan". The effect of the sample plan represents the ratio between the variance of an estimate with the specified sample design and the variance with a simple randomization sample. In the current application, we will use a stratification and a sampling at several degrees in these two types will affect the value of the effect of the sample plan.

Of the remaining parameters, the most significant is perhaps the intraclass correlation coefficient or ICC. This measures the correlation between the second-degree units (households) in the same first-degree unit (PSU, perimeter). This involves a measurement of the internal homogeneity of the first-degree unit (a high value indicates that the units in the same community tend to be similar).

It should be noted that the effect of the icc on the precision and the power associated with the estimated differences essentially depends on the matched pair level. If the pairing is done within the PSU (perimeters), the icc improves the precision and the power. In this case, the value of ρ_{13} is essentially the icc and the larger the icc, the more the precision and power are improved. On the other hand, if the pairing is done between the PSUs, as is the case of this study, an increase in the icc increases the variability of the first-degree sample units and the precision and power are reduced.

The value of the icc varies according to the variable and is generally a function of the size of the first-degree units, i.e. it generally diminishes, as the size of the first unit increases. For relatively homogenous PSUs, as perimeters are supposed to be, the value of the icc should be low, e.g., 0-0.2.

In most of the evaluation studies, an optimum value (fixed) of the sample size in the PSUs is determined (by considering the icc and the relative costs of the sampling of the PSUs with respect to the units within the PSUs) and the determination of the sample size involves the number of PSUs, as a function of the number of units sampled for each. The current study is different from most, in that the number of PSUs was fixed (at 67) and the problem of determining the sample size is the definition of the number of units to be selected per PSU.

It should be noted that in the case of an analytical survey, the finished correction of population is irrelevant. Accordingly, the fact that we took samples from all the PSUs of the target population does not reduce the variance of the sample (as would be the case in a descriptive survey). The samples of the 67 treatment PSUs and 67 control PSUs can be perceived as samples of 67 first-degree sample units of a conceptually infinite population and not as a stratification (or block) of a fixed population of 67 strata. The purpose of the impact evaluation is to create inferences concerning the generation process of the units and not the fixed population in question.

Generally, the value of the icc has a certain bearing on the sample size of the units (households) within the PSUs (generally small geographic zones, such as communities or survey zones). If the value of the icc is low, it is more effective to select more households per PSU then when it is high. Another factor affecting the sample size of the household within the PSUs is the relative sampling cost of the PSUs with respect to that of the households within the PSUs. As the cost increases per PSU, this sample of households per PSUs also increases. As indicated, in this application, the sample size of PSU is fixed and accordingly the size of the sample within the PSUs will not be "optimum".

Within the framework of the survey, the second-degree sample size (households falling within the perimeters) is relatively constant. This approach is practical from the administrative standpoint and is generally effective for socio-economic surveys of households in developing countries. Due to the value of the icc and the c_1/c_2 ratio of the first-degree sample cost with respect to the second-degree sample cost, it is possible to calculate an optimum value for the second-degree sample size (household), m_{opt}. This value is approximately m_{opt} = sqrt $((c_1/c_2)icc/(1\text{-}icc))$. For socio-economic surveys in developing countries, where the community represents the first degree sample unit and the household represents the second-degree sample unit, the value of m_{opt} is generally situated in the range of 5 to 30. Within the framework of this survey, the sample size within the PSUs is approximately 25. It should however be noted that this number comes from the fixed number of PSUs and not from considerations of what is optimum.

In the case of an analytical survey, stratification is generally used to optimize the precision of the estimates of ratios and not the precision of the estimates of the overall characteristics of the population (or even on the estimate of the overall double difference). To do this, it is necessary to check the deviation, equilibrium and the orthogonality of the explanatory variables by means of a marginal stratification. Used in this manner, stratification generally has little bearing on the effect of the sample plan. In this case, the effect of the sample plan is mainly determined by the icc and m. If the effect of stratification is ignored, that of the sample plan depends only on the icc and m via the deff formula (effect of the sampling plan) = 1 + (m-1)icc. In most studies, the values of the icc and m are indicated and that of the effect of the sampling plan (deff) is determined from these values. Within the framework of this study, m is not fixed and is – with deff – determined by the specified number of the UPEs and by icc.

Formulas for the power, by using a double difference evaluator with a control-group-pre-post-test design.

The formula of the power of a test of hypothesis on a double difference average is as follows:

$$\Pr\left(\frac{\hat{\mu}_1 - \hat{\mu}_2 - \hat{\mu}_3 + \hat{\mu}_4}{(deff\ \mathrm{var}(\hat{\mu}_1 - \hat{\mu}_2 - \hat{\mu}_3 + \hat{\mu}_4))^{1/2}}\right) > z_{1-\alpha} \left| \mu_1 - \mu_2 - \mu_3 + \mu_4 = D \right) = 1 - \beta$$

where

$$\mathrm{var}(\hat{\mu}_1 - \hat{\mu}_2 - \hat{\mu}_3 + \hat{\mu}_4) = \frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2} + \frac{{\sigma_3}^2}{n_3} + \frac{{\sigma_4}^2}{n_4} - \frac{2\rho_{12}\sigma_1\sigma_2}{\sqrt{n_1 n_2}} - \frac{2\rho_{13}\sigma_1\sigma_3}{\sqrt{n_1 n_3}} + \frac{2\rho_{14}\sigma_1\sigma_4}{\sqrt{n_1 n_4}}$$

$$+\frac{2\rho_{23}\sigma_1\sigma_3}{\sqrt{n_1 n_3}}-\frac{2\rho_{24}\sigma_2\sigma_4}{\sqrt{n_2 n_4}}-\frac{2\rho_{34}\sigma_3\sigma_4}{\sqrt{n_3 n_4}}$$

where

μ_1 = average of group 1 (treatment, time 1)
μ_2 = average of group 2 (treatment, time 2)
μ_3 = average of group 3 (comparison, time 1)
μ_4 = average of group 4 (comparison, time 2)
n_1 = sample size of group 1
n_2 = sample size of group 2
n_3 = sample size of group 3
n_4 = sample size of group 4
σ_1 = standard deviation of group 1
σ_2 = standard deviation of group 2
σ_3 = standard deviation of group 3
σ_4 = standard deviation of group 4
ρ_{12} = correlation between elements of groups 1 and 2
ρ_{13} = correlation between elements of groups 1 and 3
ρ_{14} = correlation between elements of groups 1 and 4
ρ_{23} = correlation between elements of groups 2 and 3
ρ_{24} = correlation between elements of groups 2 and 4
ρ_{34} = correlation between elements of groups 3 and 4
(The correlation matrix must be defined as positive).
α = significance level of the unilateral test of hypothesis of equality of the group averages (the probability of the Type-I error, i.e., the probability of rejection of the hypothesis of equality of the group averages, when in fact it is true) (e.g., 0.05)
β = the probability of making a Type-II error, i.e., the probability of acceptance of the hypothesis of equality of the group average, whereas it is in fact false) (e.g., 0.1).
$1 - \beta$ = power of the test (par ex., 0,9)
$z_{1-\alpha}$ = 1-α percentage point of a normal distribution (e.g.. 1.6449 for α=0,05 or 1,2816 for α=0,1)
deff = effect of the sampling plan (here it represents the ratio between the variance of an estimate for a specified survey design and the variance with a simple randomization method).
D = (true) size of the mean double difference
and a sign of omission (ˆ) over a symbol denotes a sample estimate.

It should be noted that in this application, the value of deff is not indicated as a function of a fixed *m* and *icc*. Instead of this, the number of PSUs per group is specified (at 67) and m and deff are determined as a function of that value. The formulas making it possible to determine the total size of the sample per group were derived from the above-mentioned equation and are as follows, in the event all sizes of groups (n) are equal, let's say, at *n*:

Total size of sample per group = T = $\frac{ESS(1-\rho)}{1-\rho ESS/m}$

where

Actual sample size = $ESS = \frac{(Z_\alpha + Z_\beta)^2 S^2}{D^2}$

and

$$S^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 - 2\rho_{12}\sigma_1\sigma_2 - 2\rho_{13}\sigma_1\sigma_3 + 2\rho_{14}\sigma_1\sigma_4 + 2\rho_{23}\sigma_1\sigma_3 - 2\rho_{24}\sigma_2\sigma_4 - 2\rho_{34}\sigma_3\sigma_4$$

(To express the previous formula in terms of coefficient of variation (CV), simply replace the σ with CV multiplied by the respective mean, e.g., replace σ_1 by $\mu_1 CV_1$.)

Using the previous formula, we can estimate the sample sizes required to obtain a specified level of power corresponding to the parameter values discussed as well as other values. We can also estimate the power associated with specified sample sizes according to different suppositions. We will present several results below. It should be understood that the estimated sizes and sample powers are sensitive to parameter values and that the parameter values differ for each variable of result of interest. A better idea of the parameters values will be available after analyzing the base data. While the results indicate that the baseline sample size of 1,600 per group is reasonable according to particular sets of reasonable suppositions on parameter values, these results should not be interpreted as forecasts of highly precise power but perceived as a range of powers associated with a range of suppositions.

Estimated sizes and power of sample for an estimate of impact of income according to various hypotheses

Case 1. Research to detect an increase of 8% in income with the probability of 95% (i.e., power). The standard deviations of income with respect to the average (i.e., , coefficient of variation) for the four groups are $\sigma_1 = \sigma_3 = \sigma_4 = 0.5$; $\sigma_2 = 0.5\ (1.08) = 0.54$. The value of the icc is supposed to be 0.055. The total size of the sample of households per group (treatment before, treatment after, control before, control after) = 1,588. (The value of the icc was defined to supply a sample size near that which had been used for the baseline survey.)

The initial plan of the follow-up surveys consists of only selecting 25% of the initial survey as a panel for the entire duration of the project. The previous case 1 supposes that this sample panel is made with the entire baseline sample. If the panel sample is only done for 25% of the base sample, the required sample sizes increase considerably or the sample power of 1,600 per group to detect an 8% increase in revenue falls considerably.

The problem here is that an 8% increase is a small increase, even with respect to the relatively small value of 0.5 supposed for the CV. Larger sample sizes are required to detect small differences. In the case of a such a small minimum detectable difference (change of 0.8 compared to the CV of 0.5) it is considered indispensable to select a panel sample for all base samples.

One could almost think that the CV is smaller within the perimeters. This is probably true. The CV used in the previous formula is the relative standard deviation (relative to average) for the entire population. The fact that the CV is certainly smaller within the perimeters is reflected in the value of the icc.

For samples of PSUs of the size of 67 per group, it is impossible to have a greater power for detecting a difference of 0.08, if the icc is medium or large (e.g., 0.1 or 0.2).

Case 2. Identical to case 1, but the power = 0.90. Let's suppose that the icc = 0.081. Total size of the sample of households = 1,583. (As for Case 1, the value of the icc was defined to produce a sample size near that used for the baseline survey)

A question was raised concerning the recombination of 11 pairs of perimeters (out of 67). For three of these pairs, the treatment level was the same (treatment-treatment or control-control) but for the others, the level of treatment was different (i.e., treatment/control). A major characteristic of this study is the fact that a member of each pair was randomly assigned to the treatment. This modification was made to reduce the social tension that resulted from a former division of these perimeters. Although any departure from the initial design can introduce a bias, it does not seem that this recombination could introduce a significant bias. This could even involve estimates of impact that would be a better projection of the impact of the program in normal conditions.

Besides this question of bias, it was wondered to what extent this change can affect the power. The answer was that there would be no significant impact, since it is a relatively small portion out of 67. If the number of perimeters goes from 67 to 63, let's say following this change, the results of case 1 would change as follows:

Case 3. Similar to Case 1, except for the number of perimeters that went from 67 to 63. For the power of 0.95, the total size of the sample required should go from 1,588 to 1,740 if the number of perimeters becomes 63.

Case 4. Similar to Case 1 except for the number of perimeters that went from 67 to 63. By keeping the size of the sample (e.g., 1,588), the power will go from 0.95 to 0.944 if the number of perimeters becomes 63.

F. Annex E: Statistics Glossary

G. Annex G: Questionnaires

[See Separate Sheet]