



Cocoa household income study approach

A sector-wide approach to assessing the living income status of households in the cocoa sector



van der Haar, S., Janssen, V.C.J., Diallo, O., Boza, F.B., Diarra, I., Ingram, V., Kouadio, K.A.S., Laven, A., N'dri, A.N., N'guessan, A., Waarts, Y., 2024. *Cocoa household income study approach; A sector-wide approach to assessing the living income status of households in the cocoa sector*. Wageningen, Wageningen Economic Research, Report 2024-038. 60 pp.; 1 fig.; 4 tab.; 22 ref.

This report is funded by WCF and developed in collaboration with GIZ and SWISSCO



This document outlines a collaborative and sector-wide approach to evaluating the living income status of households in the cocoa sector, addressing challenges in data collection and sharing that have been hindering progress. The Cocoa Household Income Study (CHIS) methodology is introduced to harmonise and provide a comprehensive framework for assessing living incomes. Focused on impacting at scale, the document serves as a step-by-step guideline for narrowing living income gaps in the cocoa sector. It lays the foundation for aligning data collection, sharing, and learning, offering detailed instructions for various purposes, including yearly monitoring, impact evaluations, and targeted studies. The guidelines cover income and intervention data collection, study design, sample selection, and analysis, with specific instructions for enumerators and guidelines for involving cocoa-producing households in the process.

This report can be downloaded for free at <https://doi.org/10.18174/652510> or at www.wur.eu/economic-research (under Wageningen Economic Research publications).

© 2024 Wageningen Economic Research

P.O. Box 29703, 2502 LS The Hague, The Netherlands, T +31 (0)70 335 83 30, E communications.ssg@wur.nl, <http://www.wur.eu/economic-research>. Wageningen Economic Research is part of Wageningen University & Research.

© 2024 Royal Tropical Institute (KIT)

Mauritskade 63, 1092 AD Amsterdam, The Netherlands, T +31 (0)20 568 8711, E communication@kit.nl, <https://www.kit.nl/>.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.

© Wageningen Economic Research, part of Stichting Wageningen Research, 2024

© Royal Tropical Institute (KIT), 2024

The user may reproduce, distribute and share this work and make derivative works from it. Material by third parties which is used in the work and which are subject to intellectual property rights may not be used without prior permission from the relevant third party. The user must attribute the work by stating the name indicated by the author but may not do this in such a way as to create the impression that the author/endorse the use of the work or the work of the user. The user may not use the work for commercial purposes.

Wageningen Economic Research accepts no liability for any damage resulting from the use of the results of this study or the application of the advice contained in it.

Wageningen Economic Research and Royal Tropical Institute (KIT) are ISO 9001:2015 certified.

Cover photo: Shutterstock

Table of contents

Summary	5	4.4	Developing the sample selection strategy	32
1 Background and objectives	7	4.5	Calculating the required sample size	33
1.1 Background	7	4.6	Using qualitative methods	33
1.2 Objectives and scope	8	4.7	Dealing with existing data	34
1.3 Who is this report for?	10	4.8	Additional analyses, cleaning and reporting.....	34
1.4 Reading guide	10	5 Enumerator approach	36	
2 Measuring household income and the living income gap	11	5.1 Selection of enumerators	36	
2.1 Selection of the respondent.....	11	5.2 Enumerator training	36	
2.2 Household size	12	5.3 Testing of data collection tools	37	
2.3 Total household land size.....	13	5.4 Conducting the survey	37	
2.4 Net cocoa income	15	5.5 Confidentiality and data ethics	38	
2.5 Net non-cocoa income.....	20	5.6 Data quality assurance	39	
2.6 Total household net income and living income gap	22	6 Farming household engagement in studies, learning and intervention design	41	
3 Collecting intervention data.....	24	7 Glossary	43	
3.1 Accurate intervention data at the individual level is needed to evaluate the impact of interventions and policies.....	24	References	46	
3.2 Two types of intervention data	25	Appendix 1 Survey	47	
3.3 Data from intervention monitoring systems	25	A1.1 General sections	47	
3.4 The intensity of interventions	26	A1.2 Minimum requirement	48	
3.5 Intervention data from surveys.....	27	A1.3 Recommended approach	51	
3.6 Intervention data analysis.....	27	Appendix 2 Conducting regression analyses	58	
4 Designing living income studies: sample selection and analysis	28	A2.1 Comparing between groups while controlling for other factors... 58		
4.1 Defining the population of interest	28	A2.2 Difference-in-Difference	58	
4.2 Strategies for different objectives of living income studies	28			
4.3 Dealing with challenges on finding comparison groups for impact assessment	32			

List of boxes

Box 1.1	Important definitions: Living Income	7
Box 2.1	Important definitions: cocoa producing households.....	11
Box 2.2	Data collection: Interview respondent(s)	11
Box 2.3	Important definitions: Household members and size	12
Box 2.4	Data collection: Household members and size.....	13
Box 2.5	Important definitions: Household land	13
Box 2.6	Accounting for sharecropping arrangements	14
Box 2.7	Data collection: land size and type	15
Box 2.8	Important definitions: Net cocoa income	15
Box 2.9	Important definitions: Total value of cocoa production	16
Box 2.10	Data collection: Total value of cocoa production	17
Box 2.11	Important definitions: Total costs of cocoa production.....	17
Box 2.12	Data collection: Total costs of cocoa production	19
Box 2.13	Important definitions: Net non-cocoa income.....	20
Box 2.14	Data collection: Net non-cocoa income	21
Box 2.15	Important definitions: Living income gap	22
Box 3.1	Important definitions: Intervention data	24
Box 5.1	Best practices for data collection	40

List of tables

Table 1	Extension on LICoP indicators and definitions	5
Table 2	Main elements of, and differences between the minimum requirements and the recommended approach	6

List of figures

Figure 3.1	Example of an intervention logic of an intervention on farm and income diversification	24
-------------------	--	----

Summary

A sector-wide approach to assessing the living income status of households in the cocoa sector

The cocoa sector has been working to deliver living incomes for cocoa farming households since 2015, but learning at the sector level aimed at achieving living incomes has been hampered by challenges in data collection and sharing. To address these challenges and to inform policy and strategy design and delivery with a focus on impact at scale, a collaborative approach is required. With this goal in mind, and with the intention of harmonising and ensuring the availability of a comprehensive household income evaluation methodology across the cocoa sector, the Cocoa Household Income Study (CHIS) was developed. This document presents this methodological approach.

A step-by-step guideline for assessing living income gaps for intervention and policy design

The objective of this document is to establish the groundwork for further aligning of data collection and the sharing of data and learnings aimed at reducing living income gaps in the cocoa sector. It should be applicable for different types of data collection purposes, ranging from yearly monitoring rounds to regular impact evaluations and specific deep-dives related to incomes. The document offers detailed guidelines for collecting data on incomes and intervention participation, crucial for assessing intervention effectiveness in increasing household incomes. Additionally, it provides comprehensive instructions on designing living income studies in terms of sample selection and analyses, tailored to the study's objectives. Finally, it includes specific instructions for enumerators during data collection and guidelines for involving cocoa-producing households throughout the process.

Adjusting living income assessments to the realities of cocoa-producing households globally

This methodology builds upon the work of the Living Income Community of Practice (LICO-P) (de los Rios and Tyszler, 2020; Romo, 2022), as well as the

authors' extensive experience in assessing living income in the cocoa sector. It integrates recent discussions in the LICO-P Technical Advisory Committee and the Digital Integration of Agricultural Supply Chains Alliance Farmer Income & Costs of Production Working Group. Table 1 summarises the main deviations in this guideline from the LICO-P indicators and definitions, aiming to better align to the realities of cocoa-producing households, to improve comparability of living income data, as well as to improve applicability to socio-cultural settings in different cocoa-producing countries globally.

Table 1 Extension on LICO-P indicators and definitions

Indicator/ definition	Adjustment
Household size (§2.2)	Further specified definition of household members to accommodate complex social realities of rural (cocoa-producing) households to include 'people living in the household for a at least six months per year' and 'dependent children who live elsewhere (e.g. in school) and to exclude 'paid labourers who live with the household.'
Household land (§2.3)	Further specified the definition of household land to accommodate tenure arrangements common in cocoa production and diverse cultural contexts by including leased land and land under sharecropping arrangement (household member sharecropping in someone else's land or the household sharecropping out some of their land).
Cocoa income (§2.4)	More precise assessment of cocoa production volume by accounting for post-harvest losses, agroforestry/intercropping and sharecropping arrangements. Additionally, more details on material and labour costs of cocoa production, defined per production step.
Non-cocoa income (§2.5)	Combination of (non-cocoa) farm income, off-farm income and other income into 'non-cocoa income' to develop a simple estimate of non-cocoa income based on net cocoa income and its share in total net household income. The distinction between income sources is preserved in the, more detailed, recommended approach. Finally, agricultural production that is directly consumed by the household is not measured separately.

A methodology suitable for different types of data collection

The updated operationalisation of the main indicators discussed in Table 1 is detailed in Chapter 2. To accommodate various types and purposes of data collection, this document refers to both the minimum requirements in terms of data collection, as well as a more detailed recommended approach. Table 2 offers a concise overview of the main elements of, and differences between the minimum requirements and the recommended approach for each of the key indicator operationalisations further elaborated upon in Chapter 2.

Table 2 *Main elements of, and differences between the minimum requirements and the recommended approach*

Indicator/topic	Minimum requirements	Recommended approach
Respondent selection (§2.1)	Person most knowledgeable about cocoa production	+ Person most knowledgeable about other household income sources
Household size (§2.2)	Total number and age groups	+ Socio-demographic characteristics (household roster)
Household land size (§2.3)	Land size and tenure arrangement	+ Land use and cocoa tree age
Net cocoa income (§2.4)	Cocoa revenue from own production	+ Accounting for sharecropping arrangements
	Cocoa production costs for main cocoa production activities	More detailed cocoa production activities
Net non-cocoa income (§2.5)	Share of contribution of non-cocoa income activities to total household net income	Actual income data and costs per income source

1 Background and objectives

1.1 Background

The cocoa sector has aimed to achieve living incomes for cocoa farming households since 2015

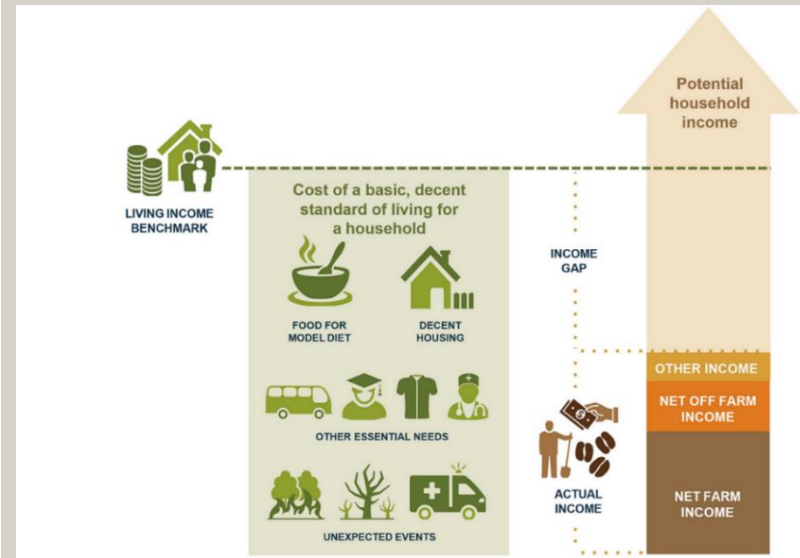
Given persistent and extreme poverty rates among cocoa producers worldwide (Fountain and Hutz Adams 2015, 2022), since 2015 many actors in the cocoa sector have been reflecting on how to ensure that cocoa farmers earn at least a living income ('revenu décent'), which is the amount they need to earn to be able to afford a decent standard of living (Box 1.1). In 2018, leaders from across the cocoa and global chocolate industry came together in Abidjan,¹ Côte d'Ivoire, to provide feedback on the draft Living Income benchmark study for the main cocoa growing regions. In 2019, the vision of the Conseil du Café-Cacao (CCC) was presented.² In 2018, leaders also met in Accra, Ghana,¹ to provide feedback on the draft living income benchmark study for cocoa growing regions of Ghana. Since then many more independent studies on living income have been conducted around the world.³ The international multi-stakeholder Alliance on Living Income in Cocoa (ALICO) was formed during the Berlin Declaration of the World Cocoa Conference in 2018. The Declaration affirmed the importance of a living income for cocoa farmers as part of achieving sustainability in the cocoa sector. The ALICO brings together stakeholders in the cocoa value chain including governments from cocoa consuming countries, civil society, companies and industry representatives, The World Cocoa Foundation (WCF), standard setting organisations, the Living Income Community of Practice (LICO) and national European multi-stakeholder-platforms, the Initiatives for Sustainable Cocoa (ISCOs) (Beyond Chocolate Belgium, Dutch Initiative for Sustainable Cocoa (DISCO), Swiss Platform for Sustainable Cocoa (SWISSCO), German Initiative for Sustainable Cocoa (GISCO)). ALICO aims at contributing to systemic change in the cocoa sector to ensure that cocoa farmers earn a living income.

¹ See: <https://www.living-income.com/single-post/2018/09/26/c%C3%B4te-d-ivoire-validating-the-living-income-benchmark-for-cocoa-growing-regions>

² See: <https://www.snrd-africa.net/fr/conseil-du-cafe-cacao-living-income-en-francais/>

Box 1.1 Important definitions: Living Income

Living income is defined as: 'the net annual income required for a household in a particular place to afford a decent standard of living for all members of that household.' (LICO, 2023). It covers the costs of food, water, housing, education, healthcare, transport, clothing, and a margin for unexpected events (see LICO figure below), and is intended as a stepping stone towards achieving prosperous incomes.



To measure to what extent households reach a living income, the **Living Income Gap** can be calculated, which is the difference between **actual annual household income** and the relevant **Living Income Benchmark** (for a specific geographical area, period and household reference size).

³ For an overview of existing benchmark studies or benchmark values, please see the ALIGN website: <https://align-tool.com/>

Learning at the sector level aimed at achieving living incomes is hampered by issues with data collection and sharing

Different organisations are implementing different types of activities aimed at achieving living incomes for cocoa farming households. As cocoa-producing households in the major cocoa-producing countries still have a large living income gap (Tyszler et al., 2018), identifying the most effective interventions and policies affecting different types of households is crucial. Finding the right interventions and policies requires adequate monitoring of households' incomes and the activities they have been exposed to, as well as effective sharing of key learnings and data between different actors. Currently there are several issues with collected data, including:

- **Restricted accessibility:** data are not available for parties other than the data owners and processors (generally manufacturing or trading companies, certification organisations and NGOs) nor for cocoa-producing households and farmer organisations.
- **Low representativity:** data are not representative of the entire sector. Information from households in *indirect supply chains*, and on women, youth and sharecroppers in *direct supply chain* is missing.
- **Lack of comparability:** data collection methods are often different and key concepts such as cost of cocoa production or income from non-cocoa sources (on-farm and off-farm) are not defined and measured in a harmonised way.
- **Not suitable for its purpose:** the collected data is often not specific enough for evaluations or policy design by public and private actors, NGOs/CSOs or platform organisations, and/or is incomplete and incomparable in terms of information on intervention participation and performance of participants. This refers to, for example, claims on impact being made without having used the right approach to make such claims.

All of the above leads to incomplete information for sector actors to use for evaluations and the design of interventions and policies.

A study that addresses these issues to deliver impact at scale

To tackle these challenges, and to deliver impact at scale by feeding policy and strategy design in the cocoa sector, a collaborative approach was needed. Given this context, ALICO in collaboration with World Cocoa Foundation (WCF), the ISCOs, COCOBOD, Cocoa Research Institute Ghana (CRIG) in Ghana, Conseil du Café Cacao (CCC), the Working Group on Living Income in Côte d'Ivoire, LICoP,

Voice Network, *Deutsche Gesellschaft für Internationale Zusammenarbeit* (GIZ), Sustainable Trade Initiative (IDH), the Research Institute of Organic Agriculture (FiBL), KIT Royal Tropical Institute and Wageningen University & Research (WUR) joined to design a Cocoa Household Income Study (CHIS) aimed at harmonised and complete sector-wide monitoring and impact study results, as well as to encourage open-source data and information while protecting the interests of data owners and households.

An engagement phase on the implementation of a Cocoa Household Income Study showed that the first step would be the development of a sector-wide approach to measuring household incomes

During the engagement phase in 2022 discussions were held with stakeholders in Côte d'Ivoire, Ghana, Cameroon, Nigeria, Indonesia, Peru and Ecuador as well as stakeholders active globally such as companies, NGOs, platform organisations, and consuming country governments by WUR and FiBL with support of GIZ and SWISSCO. During this engagement phase it was discussed how a Cocoa Household Income Study could best be implemented. Results from this engagement phase show that there is high interest in sector-wide data collection for intervention design, and that there is a need for an evidence base for learning and creating impact, but that there are key challenges in terms of methodology and the sharing of data, as well as that there is a need for local ownership and inclusive activities and governance (e.g. with cooperatives). This includes gradually introducing shifts towards increased access to and ownership of data by households and/or farmer organisations. It was decided that the first step within the process of the study would be the development of a sector-wide approach to measuring household incomes. The current document presents this approach, which was developed in partnership with CIRES and EMC from Côte d'Ivoire.

1.2 Objectives and scope

A step-by-step guideline on assessing living income gaps for intervention and policy design

The objective of this document is to lay the foundations for further alignment of data collection and sharing within the initiatives of data and learnings aimed at

decreasing the living income gap in the cocoa sector.⁴ This document serves as a step-by-step guideline to assess the income of cocoa-producing households. It is intended to be a living document, allowing for methodological changes or additions as time passes. The document also provides guidelines on how to design and implement different types of living income studies, ranging from 1. one time assessment of the living income gap to, 2. monitoring of income changes, 3. comparing different groups of households (e.g. male/female headed, certified/uncertified, agroforestry/full-sun production system), and 4. evaluation of impact of interventions. For the latter, this methodology includes guidelines for a clear definition and measurement of interventions aimed at closing the living income gap. As it is important to design interventions that consider the large diversity of cocoa-producing households, this methodology can be used to measure the living income gap in general, and, in addition, pays particular attention to groups of farmers who have been largely 'invisible' in past living income studies. These include women, youth, sharecroppers and households in indirect supply chains.

Questions on additional topics of interest should be added depending on the focus and purpose of a study

This document focuses on assessing household incomes and living income gaps. However, to learn more about the drivers of changes, or any other factors at play, it is important to include additional survey or interview questions covering any topic of interest. These additional topics of interest may include for example, good agricultural practices, pest and disease, household (financial) decision making or total household labour availability. As the additional topics of interest strongly vary from study to study, these questions are not included in this approach, but they can easily be added to the questionnaire developed here. It is important to determine at the start of a study which additional questions may be important to add, as policy or intervention design can only be done correctly if the drivers of change are well-covered and understood. The questionnaire included in this approach (see Appendix 1) covers the most important characteristics needed for the presentation of disaggregated results or

farmer segmentation (e.g. gender of the household head), but also these may be further expanded depending on the purpose of a specific study.

Accommodating for different data collection objectives: minimum requirements and recommended approach

This approach is intended to cover a wide range of different data collection objectives. On the one hand it should be applicable by researchers who want to measure household incomes in detail, but many organisations in the cocoa sector already collect data on all farmers in their direct supply chain on a yearly basis for monitoring purposes. The approach is also thus intended to be used for such data collection purposes. To make the most efficient use of both financial resources as well as of the time of cocoa farmers, we suggest keeping the number of variables collected during these monitoring moments as limited as possible. Throughout the document as well as the survey in Appendix 1, we therefore refer to the 'minimum requirements' as well as the 'recommended approach'. The first term indicates the minimum requirements that should be met when measuring income levels and living income gaps, and can for example be used for yearly monitoring rounds. It is important to make sure that the indicators and definitions used are in line with at least the minimum requirements as stated in Chapter 2 and in Appendix 1. This specifically pertains to the definitions of household size, land size, farm income and off-farm income as differences in these indicators can lead to incomparable living income and productivity measures. If the minimum measurement criteria are not met in existing data, comparison with this data may yield incorrect results. Also, the measurement units used should be double-checked to ensure consistency. The guidelines specified under the 'recommended approach' are those recommended as best practice, as they allow for a more accurate assessment of incomes and the calculation of additional indicators (e.g. yield, labour allocation), but do require some additional investments in terms of time and/or money.

⁴ To further facilitate the sharing of data, we have also developed a draft data sharing protocol, which is a draft contract to be signed by different organisations with the aim of sharing data through a neutral, independent party to solve practical issues around data sensitivities. Such a collaboration could be used to anonymise data for learning at the sector level rather than at the

individual company, cooperative or community level, while organisations remain owners of the data proved by them. The draft contract is based on an existing similar initiative and can be requested through the authors of this document.

1.3 Who is this report for?

This approach is aimed at all stakeholders in the cocoa sector with an interest in measuring and evaluating the living incomes of cocoa-producing households and/or implementing interventions or policies aimed at increasing incomes, ranging from cocoa buying and trading companies to governments, cooperatives, NGOs, researchers and consultants. On the one hand, the methodology proposed can support the stakeholders in adjusting their existing tools for data collection to ensure adequate monitoring of activities as well as accurate impact assessment of living income interventions. On the other hand, it can also be used for primary data collection to build a representative baseline of cocoa farming households, both on the regular value chain and the more 'invisible' groups that are usually overlooked by the living income studies.

1.4 Reading guide

This report is structured as follows. Chapter 2 provides a step-by-step guidance on measuring household incomes, chapter 3 on collecting intervention data, Chapter 4 on designing living income studies with guidelines on sampling and analyses, depending on the objective of a study. Finally, Chapter 5 provides guidance for enumerators' selection and training and Chapter 6 provides guidelines on involving cocoa-producing households in the entire process of a study. A glossary of key terms is found in Chapter 8. Survey questions are listed in Appendix 1 and a further guidance on conducting regression analyses is presented in Appendix 2.

The chapters contain several different types of boxes. Grey boxes provide key definitions, green-edged boxes contain a summary of questions for data collection and calculations of key indicators. Blue-edged boxes provide other types of relevant information.

2 Measuring household income and the living income gap

To be able to calculate total household income and the living income gap, a detailed household survey is used. This chapter provides step-by-step guidelines for measuring all relevant aspects of cocoa-producing households' income, including definitions and calculations. The final survey instrument, combining all the subsections, for the minimum and the recommended approach separately, can be found in Appendix 1.

Living income is a household concept: household net income is all the net income earned by all household members together in one year. It covers net farm income (e.g. income from cash crops, food crops, livestock), net off-farm income (e.g. employment, business) and all other income (e.g. pensions and remittances). Because these guidelines are for measuring the living income of cocoa-producing households, we distinguish between the measurement of the net cocoa income (Chapter 2.4) and the net non-cocoa income (Chapter 2.3). But first we explain how to select respondents (Chapter 2.1), define household size (Chapter 2.2) and farm size (Chapter 2.3).

2.1 Selection of the respondent

Box 2.1 Important definitions: cocoa producing households

Cocoa-producing households: Household of which at least one of the members is a cocoa farmer.

To assess the living income of all cocoa-producing households an inclusive definition is used (Box 2.1) based on the following definition of a cocoa farmer:

'Everyone who cultivates cocoa is recognised as a farmer, irrespective of gender or landholding status'⁵

In a traditional approach, usually the registered farmer (often the 'household head') is invited to participate in a survey. This is known to create a bias towards interviewing elderly men who are registered as farm owner/manager, potentially overlooking other household members and labourers who also work on the cocoa farm. Our definition allows a better representation of sharecroppers, youth and female spouses of registered cocoa farmers in data collection. Recognising the contribution of women to cocoa production and incomes, and getting their input will also help to get higher quality data on children (e.g. age, gender, education status), food and non-cocoa income.

Box 2.2 Data collection: Interview respondent(s)

Minimum requirements: Person most knowledgeable about cocoa production

- Name of the person
- Gender of the person
- Age of the person
- Position within the household (household head, spouse, son, daughter, etc.)
- Other personal/household characteristics of interest not covered in the minimum approach

Recommended approach: 1. Person most knowledgeable about cocoa production and; 2. Person most knowledgeable about the other household income sources. For both persons we would collect at least the same information as covered under the minimum requirements.

⁵ Mars Cocoa for Generation Progress Report, 2022.

As living income is a household concept, it is important to capture the income sources of all household members. This has an implication on the data collection approach, as ideally, all household members that contribute to a household's income are interviewed. However, this would significantly increase data collection costs and complicate data cleaning and analysis. Therefore, we recommend two options:

1. **Minimum requirement:** interview the person in the household most knowledgeable about cocoa production. This person should be able to report on both the income earned and the costs incurred during cocoa production.
2. **Recommended approach:** interview two persons per household: 1. The most knowledgeable person about cocoa production and 2. The most knowledgeable person about the other household income sources.

The second option is preferred because accurate information about other farm and off-farm income sources is still a major knowledge gap in living income studies globally.

At the start of the household survey, some **screening questions** should be added to make sure that the respondent represents the population of interest. Depending on the objective of the living income study, the population of interest can be defined based on farmer characteristics (female farmers, youth, sharecropper, indirect supply chains) or interventions (beneficiary, control group, certified/uncertified). This is explained in more detail in Chapter 4. Examples of screening questions are included in Appendix 1.

Box 2.2 summarises the main data points to collect on respondents. Because the **minimum requirements** are less detailed in terms of household characteristics (Chapter 2.2), we recommend adding any specific household/respondent information that is needed to do a disaggregated analysis based on the objectives of the study. For instance, to compare certified and non-certified cocoa-producing households, certification status should be added and a sufficient number of certified and non-certified households should be selected to significantly represent both groups. See Chapter 4 for more information about sample selection. For the **recommended approach** respondent characteristics should be obtained for both respondents.

2.2 Household size

Box 2.3 Important definitions: Household members and size

Household members are defined as a group of adults and children, regardless of their relationship, who contribute to or are dependent on a shared economic pot. This includes people living in the household for at least six months per year and dependent children who live elsewhere (e.g. in school) and excludes paid labourers who live with the household.

Total household size (number) = SUM(household members)

What constitutes a living income for a particular household depends on the household size. The more household members, the higher the household's total annual net income needs to be to provide a decent living for all of them. Measuring household sizes can be complex, especially in contexts where people live together with extended family, household members live outside of the household for a period of time and polygamous families are common. Therefore, a uniform definition of household members is key for comparing living income assessment results.

Box 2.3 defines **household members and household size**. What is important here is that household members contribute to or are dependent on a shared economic 'pot'. Household members don't have to be related by blood or marriage. This definition expands upon the LICoP in two ways. First, short-term guests are excluded by defining a minimum of six months stay in the household except for children who live elsewhere but are dependent on the household's economic pot. Second, paid labourers who temporarily live with the household are excluded.

In the context of West-Africa, it is common that children live with another family members or in boarding schools. Here the dependence on the household's economic pot is of particular importance. Biological children who live in another family or in boarding school but who are primarily supported by the household are still counted as household members because the household will need to produce a higher income to cater for these children as well. If they, however, are primarily supported by their host family, they are excluded, because the

household income does not need to provide for them. Likewise, (adult) children who have started their own household with their own economic pot are excluded. When measuring the household size, it is critical to explain this definition very carefully. This minimises the risk of double-counting of children who live in another household.

To accurately assess households' living income gaps, the **minimum data requirement** is to capture the total number of household members and the number of adults (age 18 years and older), children aged 14-17 years and children aged 0-13 years. This decomposition allows to adjust the Living Income Benchmark to the household composition according to the OECD adult-equivalent scale as explained in Chapter 2.6. In addition, adults should be gender disaggregated to allow for gender analyses.

Box 2.4 Data collection: Household members and size

Minimum requirements: total number and age groups

Number of household members:

- total
- adults age 18 and older – men
- adults age 18 and older – women
- children aged 14 to 17
- children up to the age of 13

Recommended approach: household roster

Ask the respondent to list all household members by name, starting with themselves. Then for each household member relevant demographic information is collected. This is at minimum the gender and age, but can also include education level, marital status, active contribution to the household income (adults), current enrollment in school (children), or other questions depending on any additional topics of interest of a study.

The **recommended approach** is to use a household roster. This method provides a structured framework to capture all household members one-by-one and add relevant socio-demographic characteristics. Which socio-demographic characteristics are relevant is dependent on the objective of the living income

study. As a minimum, age and gender should be included to capture the household composition for the living income benchmark and to do gender-disaggregated analyses. In addition, to investigate possible child labour risk and ways different household members earn an income, we can also ask the number of adults active in certain income generating activities and the number of school-age children enrolled in school.

2.3 Total household land size

Box 2.5 Important definitions: Household land

Total household land size (ha): 'the size of the total land area that ANY household member either (i) owns (with or without ownership title), (ii) has rights to use (possession, assigned communal land, land reform titles, etc), (iii) has any land-use arrangement with third parties (loans, rentals, lease), (iv) has under a sharecropping arrangement (out or in).'

It therefore is: all land owned + all land with user rights + all land with temporary rights + all land sharecropped in by the household – all land given out to a sharecropper.

Total cultivated land (ha) = total household land size – uncultivated household land

Total cultivated cocoa land (ha) = cultivated land dedicated to cocoa production

Total cultivated cocoa land sharecropped = total cultivated cocoa land under sharecropping arrangement

As an intermediate step to capture household cocoa and farm income, the household farm should be delineated consistently. Box 2.5 provides a standardised definition of total household land size. It covers all land that household members have the right to use, either through ownership, user rights, third-party land use arrangements (loan, rent, lease) or sharecropping arrangements. This definition largely follows the one used by LICoP (2020) but includes leased land and adds more detail on sharecropping arrangements to better accommodate the range of land tenure arrangements in different geographical contexts.

Box 2.6 explains the difference between land **sharecropped out and in** with examples from sharecropping arrangements common in West-Africa. To correctly account for sharecropping in household income assessment the details of the sharecropping arrangement need to be captured: a. the share of the volumes of cocoa (or revenue or profit) or the share of trees (or land) that goes to the household that is registered as owner of the farm and to the tenant and; b. who carries (which part of) the costs of production, including labour, inputs and equipment. This is explained in Section 2.4 on net cocoa income. Likewise, for land under a third party land use arrangement, the costs of renting or leasing the land should be included in the total costs of production.

While it is not strictly necessary to know the total farm size to be able to assess total household income, it is useful to validate production volumes and as a starting point to assess farm income sources. For data collection, the **minimum requirement** is to capture the total household land size, the number of separate plots, the area that is cultivated (not left fallow) and the cultivated area that is planted with cocoa (Box 2.7). Moreover, for cocoa land the tenure arrangements need to be distinguished to adjust the cocoa production volume and costs accordingly (see Chapter 2.4).

All land sizes should be captured in relevant local units (e.g. acres, poles) and converted to hectares. Respondents should be able to select a land unit of their preference. A list of local land units and the equivalent in hectares should be made before the start of the survey so enumerators don't have to do the conversion during the interview at the risk of miscalculation. For increased accuracy, it is recommended to use GPS measured land size (through polygon mapping) rather than self-reported information. Studies have shown that self-reported land sizes deviate from GPS measured land sizes by an average of 23% in Cote d'Ivoire (Ingram, Waarts et al., 2014, 2018). In the absence of GPS mapped measures of land size, self-reported data on land size can be used. The number of different plots (also called farms in Ghana) can be used to cross-validate the information about land size and tenure arrangements.

Box 2.6 Accounting for sharecropping arrangements

In main cocoa producing countries, such as Côte d'Ivoire and Ghana, sharecropping land is quite common. In these countries two main **sharecropping arrangements** in cocoa production can be distinguished:

1. Share contracts between owner and tenant, whereby the tenant manages already established cocoa farms, and in return get a share of the cocoa volume, revenue or profit (usually one-third, depending on the relation between tenant and owner). This arrangement is known as Abusa and Nnewsoo. Usually the tenants pay (at least) part of the costs of hiring additional labourers and other production costs.
2. Share contracts between owner and tenant, whereby the tenants establish new farms and gets a share of the cocoa volume, revenue or profit, the trees or the land (usually in two equal parts). This arrangement is known as Partager-Travailler (Work-and-Share) or Planter-Partager in Côte d'Ivoire and Abunu or Yemayenkye in Ghana.

To account for sharecropping arrangements, it needs to be clear whether or not the respondent belongs to a household that is involved in a sharecropping arrangement as **owner** or as **tenant**. If a household owns the land, but has a tenant managing their farm, we refer to that land as being 'sharecropped out'. If one of the household members works as tenant on a plot of land owned by another household, we refer to that land as being 'sharecropped in'. It is possible that within one household both sharecropping in and sharecropping out occur. So,

- **Sharecropped out** refers to the land that is owned by the respondents' household members but cultivated by another household under a sharecropping arrangement;
- **Sharecropped in** refers to the land that is owned by another household but cultivated by the respondents' household under a sharecropping arrangement.

The **recommended approach** expands on the minimum approach by capturing the other land uses beyond cocoa (other cash crops, food crops and other agricultural practices including livestock) and the age of the cocoa trees. The latter helps to distinguish the productive from unproductive cocoa lands and accurately assess yield levels. Please note that especially in the first years after

planting, cocoa is typically intercropped with food crops such as plantain or cassava. Hence, cocoa land is not restricted to the production of cocoa. The income from the intercropped crops or agroforestry is captured in Section 2.5.

Box 2.7 Data collection: land size and type

Minimum requirements: land size and tenure arrangement

- Land unit preferred by respondent (hectares, acres, pole, other)
- Total household land size (land unit)
- Number of plots (integer)
- Total household land size (land unit) under cultivation (not left fallow)
- Total household land size (land unit) under cocoa cultivation
 - Total household land size (land unit) under cocoa cultivation that the household *owns or has user rights over* (possession, assigned communal land, land reform titles)
 - Total household land size (land unit) under cocoa cultivation that the household *loans, rents or leases* from others
 - Total household land size (land unit) under cocoa cultivation that is *sharecropped* out to someone else (land unit)
 - Total household land size (land unit) under cocoa cultivation that is *sharecropped* in by the household (land unit)
- Whether the cocoa farm has been GPS (polygon) measured.
 - If yes, total land size (land unit) measured.
 - If yes, measured by whom.

Recommended approach: include land use and cocoa tree age

In addition to the minimum requirements:

- Size of cultivated land used to produce: (land unit)
 - cocoa
 - other cash crops
 - food crops
 - other farm activities (e.g. livestock, timber trees)
- Size of cocoa land with trees aged:
 - < 5 (young trees not yet productive)
 - 5-25 (mature and fully productive trees)
 - > 25 (older trees with declining productivity)

2.4 Net cocoa income

Box 2.8 Important definitions: Net cocoa income

Net cocoa income = total value of cocoa production – total costs of cocoa production

Accurately assessing households' annual income from cocoa production, requires detailed information on the total value of cocoa production and the total costs of cocoa production. The total value of production is also called the gross cocoa income. To get the net cocoa income the total costs of production are deducted from the gross cocoa income (Box 2.8).

The most accurate way to capture farm production and sales volumes, cost of production and time spent on activities is to record this information at the time such costs are incurred or volumes are sold. This can be done using, for example, a **Farmer Field Book approach** as developed by Agri-Logic and implemented in the coffee sector since 2001 and in the cocoa sector since 2014. In this approach farmers record detailed information at the end of every day they (or others) spent on cocoa farming and sales activities. This information is collected every two weeks; the person collecting the information performs a completeness and consistency check in discussion with the farmer. The advantage of such an approach is that recall bias is minimised. If the approach is implemented well by e.g. farmers receiving adequate guidance and adequate forms that can be filled out easily, farmers do not spend much time on such record keeping. In that case, it takes a few minutes per day worked to keep such records (an estimated 1% additional time spent per workday), and about 15 minutes every two weeks when the information is collected. Especially in contexts with low literacy levels, frequent checking of records is needed to secure data quality and using an app for recording the information is also often not possible.

Because of the intensity of implementing the Farmer Field Book approach (frequent farmer visits), it is often found to be time consuming and costly to implement this method at scale. Therefore, in this guide we allow for the use of recall data. While this could lead to less accurate data, when using a large enough sample size, it does provide a good proxy for actual cocoa revenues and

costs and can be used to monitor and evaluate changes over time. It is recommended to implement the Farmer Field Book approach whenever possible, to conduct in-depth assessments of especially (changes) in cost of cocoa production including household labour deployment and farm management activities and their relationship to cocoa yield and income. Such information can also be used to triangulate survey results. See Chapter 6 for more information on sharing data and information back with respondents.

The survey instrument allows for both recall data and farmer records. When farmer records are available, we recommend using them. Otherwise, recall data can be used, possibly supplemented with farmer records for a subsample of the population to validate the data (See Chapter 4). The survey includes questions to capture the data source. In addition, to minimise recall bias, cocoa production volumes and prices are asked for the main and light cocoa production season separately (Box 2.110).

2.4.1 Total value of cocoa production

Box 2.9 Important definitions: Total value of cocoa production

Total value of cocoa production = (total volume of cocoa produced (kg) * cocoa price) + total premium received

Total volume of cocoa produced (kg) = (Volume of cocoa from land cultivated by the household (kg) + Volume of cocoa from land sharecropped out (kg) + Volume of cocoa from land sharecropped in (kg)) – post-harvest loss (kg)

Volume of cocoa from own land (not under sharecropping arrangement) (kg) = gross volume of cocoa produced (kg) – post-harvest loss (kg)

Volume of cocoa from land sharecropped out (kg) = household share of production (e.g. 2/3) * total volume produced under this arrangement (kg)

Volume of cocoa from land sharecropped in (kg) = household share of production (e.g. 1/3) * total volume produced under this arrangement (kg)

Cocoa price = most frequently received price per unit of cocoa by the farm household

Total premium received = volume of production sold as certified/high quality (kg) * premium/kg

Box 2.9 demonstrates how to assess the **total value of cocoa production**. It is the total annual volume of cocoa produced by the household, accounting for any post-harvest losses, times the cocoa price plus any additional premiums received.

- The **total volume of cocoa produced** is all the cocoa produced on the household's cultivated cocoa land. To report on these volumes more easily, the respondent should be allowed to select the most convenient unit (e.g. kgs, tons, bags). To convert all volumes to kgs, we then need to capture how many kgs are contained in each unit chosen by the respondent (e.g. most of the cocoa bags in Ivory Coast are between 65 to 80 kgs). To account for post-harvest losses, not only the volume produced but also the volume sold is captured.
- **Price for cocoa** received by the farmer could be different from official farm-gate prices set by governments for various reasons (e.g. quality of beans, distance from buyer, differences between buyers). Therefore, it is crucial to capture the amount declared by the respondent and not rely on the official price. If different portions of the production were sold for different prices, it is difficult for farmers to estimate the average price received in the season. Therefore, instead we recommend recording the most frequently received price.
- **Premiums** are usually paid to compensate for the extra efforts required to comply with certain standards and/or practices as part of certification programs or for higher quality cocoa. Just as for prices, an official premium is announced by certification bodies, but cocoa farmers rarely get this premium for the total volume as demand for certified cocoa is lower than supply. Thus, ask the percentage of the cocoa production sold as certified or with quality premiums and the monetary value of the premium, whether it's an aggregate value or per certain unit of production (kgs, bags, tons, etc).

As a **minimum requirement** for data collection, **the** total volume of cocoa produced by the household as well as the received cocoa price and premiums should be captured irrespective of the land tenure arrangement (Box 2.10). The **recommended approach** is to account for sharecropping arrangements. This provides not only more accurate estimates of total value of cocoa production but also the opportunity to do additional analyses on the sharecropping arrangements. For accuracy, the land under sharecropping arrangement recorded in Chapter 2.3 is used as starting point. For land that is sharecropped out and for land that is sharecropped in by the household, we separately capture the (total) amount of cocoa produced, the (total) amount of cocoa sold and the sharecropping arrangement. The details of the sharecropping arrangement are

asked to be able to isolate the share of the production sold that adds to the income of the household of interest. If the arrangement is not a share of the production but a share of the revenue, profit, land or trees, the volume of cocoa from land sharecropped out should be adjusted accordingly.

For land that is sharecropped out by the household of interest, the share of the production that the sharecropper can keep should be deducted from the household's production volume. For land that is sharecropped in by the household of interest, only the share of the cocoa income earned from that plot should be included in the total net cocoa income of the household of interest. The distribution of production costs is captured in the next chapter.

Box 2.10 Data collection: Total value of cocoa production

Minimum requirements: *own production*

- Cocoa production unit preferred by respondent (bags, kg's, tonnes other)
- Number of kg's per unit (for bags or other)
- Amount of cocoa produced (for light and main season) (production unit)
- Amount of cocoa sold (for light and main season) (production unit)
- Most frequently received price per unit of cocoa (light and main season)
- Amount of cocoa sold as certified or high quality *for a premium* (for light and main season) (production unit)
- Most frequently received premium per unit of cocoa (light and main season)

Recommended approach: *account for sharecropping arrangements*

In addition to the minimum requirements:

For land that your household *sharecrops out* to another household:

- Amount of cocoa produced (for light and main season) (production unit)
- Amount of cocoa sold (for light and main season) (production unit)
- Compensation of the sharecropper by the household (Categorical: share of cocoa production, revenue, profit, trees or land)

For land that your household *sharecrops in* from another household:

- Amount of cocoa produced (for light and main season) (production unit)
- Amount of cocoa sold (for light and main season) (production unit)
- Compensation of the household by the land owner (categorical: share of cocoa production, revenue, profit, trees or land)

2.4.2 Total costs of cocoa production

To evaluate the net income a household derived from cocoa production, we need to deduct the costs of cocoa production from the gross income derived from cocoa sales. Total costs of cocoa production are the sum of material costs and labour costs.

Box 2.11 Important definitions: Total costs of cocoa production

Total costs of cocoa production $SUM(\text{labour costs} + \text{material costs})$

Total labour costs = $SUM(\text{annual costs of: permanent labourers, temporary labourers, labour groups, communal labour})$ on all land under cocoa cultivation, including sharecropped in and sharecropped out

Total material costs = $SUM(\text{tools/equipment, inputs, transport, land, debt/insurance/tax})$ on all land under cocoa cultivation, including sharecropped in and sharecropped out

Material costs cover the costs of tools/equipment, inputs, transport, land and loans, insurance and taxes incurred in the cocoa production process.

- **Tools and equipment** are for example machetes for cutting branches or weeding, equipment for pruning, spraying or irrigation and protective gear. It also covers the costs of centralised drying.
- **Inputs** include planting materials (e.g. cocoa seedlings) as well as phyto-sanitary products (fertilisers, pesticides, herbicides, fungicides).
- **Transport** covers any expense incurred for renting a vehicle, including the maintenance or fuel for transporting the cocoa from the farm to the fermenting/drying/selling location and transporting inputs (e.g. seedlings) to the farm.
- **Land** costs are the costs of renting and leasing of land.
- **Loans, insurance and tax** costs should be included if relevant in the production area. For loans only the costs of taking out the loan, the interest rate, is included.
- Larger investments such as buying vehicles or land require a different accounting as these can often serve multiple purposes (outside of cocoa production) and are used over several years. Moreover, as very few

households make these types of investments and the calculation of the annual cost related to their purchase is quite complex, we recommend not necessarily including them in the estimation of the total costs of cocoa production.

Nonetheless, if it's really necessary to isolate these costs for a given study, we recommend capturing the cost and year of purchase, expected lifetime of the investment and proportion of time used for cocoa related purposes. One could then calculate the average annual cost of the investment by dividing the total cost of purchase by the expected lifetime. The value obtained is valid for the year of the purchase only, for subsequent years, a depreciation rate (depending on the type of asset considered and its frequency of use) needs to be applied to adjust the real cost. Furthermore, the annual cost also needs to be adjusted to reflect the proportion of the investment that is specific to cocoa production.

Labour costs cover four different types of labour:

- **Permanent labourers' costs** cover remuneration paid to long-term labourers such as caretakers or farm managers typically for a fixed period (6, 8 or 12 months). These permanent labourers are usually involved in all the production steps but in contrast to sharecroppers, they are paid in cash instead of a share of the harvest or the land. Sometimes, permanent labourers can also benefit from some in-kind compensation such as housing or food.
- **Temporary labourers' costs** are more difficult to capture as they are often linked to the hiring of seasonal labourers for labour intensive tasks such as harvesting or weeding. Therefore temporary labour costs should be captured per task.
- **Labour group costs** are for specific production activities. They are also called labour gangs, labour brigades or service groups. In different cocoa-producing regions, spraying or pruning groups are common. These are sometimes fully or partially subsidised by cooperatives, traders or governments. Therefore it is key to only capture the costs incurred by the household. Any in-kind costs of providing food or other services should also be captured for this group.
- **Communal labourers' costs** should also be included. Communal labour is also called 'labour exchange groups' and is typically based on reciprocity: one day household A works on household B's cocoa farm and the next day it is the other way around. While there are no direct remuneration costs involved, the receiving party often provides meals to communal labourers. These in-kind costs should be included as production costs.

- We suggest not adding the opportunity costs of household labour to the labour costs, as these costs are generally used for a different purpose: to make decisions on the division of labour at the household or farm level. This is not the purpose of living income assessments. Furthermore, adding the opportunity costs of labour would lead to the complex challenge of estimating the very case-specific additional income that could be earned if household labour would be spent on other activities. This also assumes that there would be sufficient alternative employment opportunities available that would lead to higher incomes than those earned through cocoa production, which may not be the case in many rural cocoa-producing areas. Mapping alternative employment or income generation opportunities in these areas, and learning why households do or do not engage in them would be interesting additional topics for further research. Questions on this topic could be added to household income surveys or explored through qualitative studies.

To get an accurate measure of the cocoa production costs, for all costs categories, we need to consider only the share that is covered by the household. All the costs that are paid by cooperatives or other parties (e.g. from working gangs, transport, centralised drying) must be excluded.

In the case of agroforestry or intercropping, inputs such as fertiliser and herbicides can be applied to multiple crops (or trees) on the same land. In that case, only the costs that can be attributed to cocoa should be included. The costs and revenue of other crops are covered in Chapter 2.5.

To enhance accuracy, cost information should be collected per production activity, as summarised in Box 2.12. As a **minimum requirement**, labour and material costs should be assessed separately by for four main categories of production activities: 1. Land preparation; 2. Maintenance; 3. Harvesting and; 4. Post-harvest activities. For each of these activities, ask the type of labour that is used, the number of temporary labourers and their remuneration, the number of labour gangs and their remuneration and the material costs used. Material costs should be specified per activity. Maintenance, for instance, includes the costs of tools and equipment for weeding, pruning and pest and disease management (e.g. cutlasses, protective gear, spraying equipment), inputs (phytosanitary products) and transport of these inputs.

Costs of permanent labourers and renting/leasing land are not determined per activity. Therefore, these are recorded as an annual amount both in the minimal as well as the recommended approach.

For a detailed account of the cocoa production costs the **recommended approach** is to capture the costs of each individual production activity. Box 2.12 identifies thirteen activities. This provides more precise estimates of material costs. The input costs of separate maintenance tasks, for instance, are now captured separately (fertilisers, herbicides, fungicides, insecticides). The recommended approach also provides more details on labour use. It expands on the minimum requirements by capturing the number of labourers and the time spent per labourer for each labour group and it includes the in-kind costs of labour use. The recommended approach significantly improves the accuracy of the labour data and facilitates additional analyses on labour allocation and labour productivity.

Box 2.12 Data collection: Total costs of cocoa production

Minimum requirements: main cocoa production activities

1. *Costs by (higher-level) cocoa production activity:*
For each of the 4 activities: 1. Preparation (Land clearing and planting); 2. Maintenance (weeding, pruning, applying phytosanitary products); 3. Harvesting; 4. Post-harvest activities (pod breaking, fermentation, drying, transportation) ask:
 - Which type of labour was used (household members, sharecroppers, permanent labourers, temporary labourers, labour groups, communal labour)
 - The number of temporary labourers and their remuneration
 - The number of (times) labour groups engaged and their remuneration
 - Material costs
2. *Annual costs:*
 - The number of permanent labourers and their annual remuneration
 - Costs of buying/renting/leasing land

Recommended approach: detailed cocoa production activities

1. *Costs by (detailed) cocoa production activity:*
For each of the 13 activities (1. Land clearing; 2. Planting; 3. Fertilizer application (granular, liquid and organic); 4. Herbicide application; 5. Pesticide/insecticide application; 6. Fungicide application; 7. Weeding (multiple rounds); 8. Pruning; 9. Harvesting; 10. Pod breaking; 11. Fermentation; 12. Drying; 13. Transportation (to drying location, to selling location)) ask:
 - Which type of labour was used (household members, sharecroppers, permanent labourers, temporary labourers, labour groups, communal labour)
 - The number of labourers per type of labour (household members, sharecroppers, permanent labourers, temporary labourers, labour groups, communal labour)
 - The time each labourer spent on the production activity
 - The remuneration of temporary labourers
 - The remuneration of labour groups
 - The in-kind costs (mostly food) of the different type of labourers
 - The Material costs
2. *Annual costs:* (same as for minimum requirement)
 - The number of permanent labourers and their annual remuneration
 - Costs of buying/renting/leasing land

2.5 Net non-cocoa income

Box 2.13 Important definitions: Net non-cocoa income

Net non-cocoa income = SUM(Net income from: 1. Other farm IGAs; 2. Off-farm IGAs; 3. Other income)

Other farm income = SUM(net income from cash crops, food crops, timber from cocoa farms and livestock products)

Off-farm income = SUM(net income from all other IGAs)

Other income = SUM(net income from pensions, remittances and cash transfers (recurrent))

Total production costs of non-cocoa IGA = SUM(costs of rent/lease of land/buildings (e.g. shop), equipment, raw materials and inputs (incl fuel), labour and transport)

Net non-cocoa income includes all income from (non-cocoa) income-generating activities (IGAs) performed by all household members and any passive income received by any household member (see Box 2.9). In line with the LICOP guidelines, non-cocoa income sources are divided into three main categories:

- **Other farm IGAs** comprised of other cash crops (e.g. rubber, coffee, oil palm), food crops (e.g. plantains, maize, cassava, rice, horticulture), livestock and trees (timber or other tree products).
- **Off-farm IGAs** covers trade/shop, processing of agrifood (including cocoa by-products such as juice and biobased materials of pods) and other small businesses (e.g. retailing, hairdressing, money transfer). It also includes wage labour and employment in the public (e.g. teaching, administration) or private (company) sector. Sale of fish, forest and bushmeat products, and rent/lease of land are also included in this component of the income.
- **Other income** includes all passive income received by any of the households members (e.g. pensions, remittances and cash transfers) on a recurrent basis. Recurring transfers to cover for children of other family members living in the household are covered here, as well as cash-transfers from NGOs or other public or private programs. Loans are not included as income, because they have to be paid back and often on a short/medium term (max of 12 months) in the context of rural developing countries.

Non-cocoa income is often difficult to measure with high accuracy because multiple income sources are managed by different household members. Moreover, contrary to cocoa or other cash crops for which farmers recall the volume sold and the price obtained, these other activities are performed on a more ad-hoc basis with fluctuating production, investment, revenue, etc. Therefore, getting a precise estimate of the income derived from these various sources would require an in-depth survey of multiple respondents per household coupled with a high frequency data collection to minimise the recall bias. This methodology will be costly and difficult to replicate at a large scale, not to mention the survey fatigue induced among respondents. Therefore, we propose a leaner approach to measuring non-cocoa income by adding a fairly limited number of questions within the living income survey. Providing we have a large enough sample size, the estimates provided by these questions will constitute a good proxy for the non-cocoa income earned by households through their different IGAs. For validation, more detailed data collection can be done for a small sample of cocoa-producing households, but this is beyond the scope of this guideline.

The **minimum requirements** mimic the approach used for measuring the net cocoa income in Chapter 2.4 to estimate the net non-cocoa income (Box 2.14). It starts from a list of all household income sources including cocoa, other farm IGAs, non-farm IGAs, and other income from which the respondent will select all the household's income sources. At this stage, the respondent must be reminded to consider all income sources that contribute to the household pool of resources regardless of whether the respondent is involved in the activity him or herself. Next, for each selected income source, the contribution (in the form of a percentage) to total household income is determined, starting with cocoa. In this approach, the net cocoa income will be used to estimate net non-cocoa income as demonstrated by the equations in Box 2.14. If income from cocoa is USD 1,000 and cocoa contributes for 70% to total household income, then total household income is USD 1,429 ($1,000/0.7$).

While this is a relatively simple method, it requires clear guidance from enumerators to help the respondents estimate the percentages step-by-step (e.g. if your annual income was xxx currency, how much would come from cocoa/other crops/business/trade/employment, etc.). It is recommended to use

a visual aid (e.g. beads, rocks, seeds) to visualise the percentages and to make sure that they add up to 100.

The **recommended approach** captures more detailed revenue and cost information per income source. For this section, a **second respondent** who is the household member most knowledgeable about the other income sources, is interviewed. The approach also starts from a list of income sources, but this time excluding cocoa as it was already extensively covered by the first respondent. This approach unpacks other agricultural income further by listing the other crops that are sold. The respondent is asked to identify the two crops that contribute to income most, these are added to the list of income sources and all other crops are added as 'other crops'. Specific attention needs to be paid to food crops as often these are also cultivated for self-consumption. Therefore, we must first assess which crops are sold before asking about the income earned from these sales.

Next, for all selected income sources gross income and costs are captured. Moreover, the household member responsible for every listed income source is indicated. This step serves as a verification of the list of IGAs as the respondent can then add/delete an income source depending on whether they forgot an item or confused an IGA that is managed by another household member. For each IGA listed, the approach then capture 1. gross income and 2. total production cost per day/week/month/year depending on the frequency identified as the most convenient by the respondent. To account for seasonality of certain activities, we also ask the period (how many months, weeks, days) in which the activity is performed to get an annual estimate of the income earned from each IGA. Production costs include rent/lease of land/buildings (e.g. shop), purchase or rental of equipment, raw materials and inputs (incl fuel), labour and transport costs, loan interests, insurance and taxes. Enumerators should explain this definition and ask the respondents to estimate the total costs per IGA. Moreover, they should be careful not to double count on-farm IGA costs that were already accounted for under cocoa production costs.

Box 2.14 Data collection: Net non-cocoa income

Minimum requirements: % contribution to total household net income

Respondent 1

1. List all household income sources (including cocoa)
2. Ask for each income source the percentage **contribution to total household net income**. Use visualization (e.g. beads) to aid respondents in this exercise and make sure the total is 100%.

In this case:

Net household income = (Net cocoa income * (contribution of cocoa to total household income (%) / 100))

Net non-cocoa income = (SUM of contribution of non-cocoa income sources to total household income (%)/100) * Net household income) + average value of food produced for home consumption

Recommended approach: detailed income and costs per income source

Respondent 2

1. List all household income sources (excluding cocoa)
2. Identify all the crops that are sold and isolate the two most important for income, outside of cocoa. Then, group all the remaining crops into the category 'Other crops'.
3. For each of the IGA selected including the three crops under point 2 ask:
 - The household member responsible
 - Gross income per day/week/month
 - Total production costs per day/week/month
 - The number of days/week/month per year in which the income was obtained
4. For each of the other income sources selected under point 2 ask:
 - Total amount received (per year)
 - Frequency of the payment (times per year)

In this case:

Net income per income source = gross annual income – total annual production costs

Net non-cocoa income = SUM of net incomes of all non-cocoa income sources + average value of food produced for home consumption

Other income sources should only be included if they are received on a structural basis. This is to avoid artificially increasing total household income with a one-off payment. Therefore ask for each other income source of the household the total amount received and the frequency of the payment.

Lastly, to match the calculation of a living income benchmark, non-cocoa income should theoretically also include the value of **food produced by the household for their own consumption**, as the living income benchmark assumes all food within a model diet is purchased from the market. However, collecting such information is very time consuming. We therefore suggest to instead use average values for food produced at home for own consumption to adjust the value of the living income benchmark. We strongly suggest country-specific deep dives to get more detailed insights into the average value of production of nutritious food items for home consumption of cocoa-producing households. Such a deep-dive would then need to compare the nutrient contents to the items included in the model diets used in the construction of the living income benchmark. As a pragmatic approach to dealing with food produced at home without asking respondents long lists of questions on food produced for home consumption, we recommend adding these average values to the net non-cocoa income. Even though this income has not actually been earned, as the items have not been sold, the same holds for the food items in the model diet, which have been added as costs even though they might not have been bought. An alternative option to adding this value to net non-cocoa income would be to lower the benchmark, but to keep clear, non-negotiable living income benchmarks, we advice to not alter the benchmarks. In the absence of country-specific deep-dives, information from the living income benchmarks can be used to make a very rough estimation. These benchmarks usually include information on the food items included in the model diets and the food items most commonly produced at home for own consumption. Upon overlapping these items, we estimate that the monetary value of food produced at home for own consumption accounts for at least 4% of the overall living income benchmark in Indonesia, and about 8% in both Ghana as well as Côte d'Ivoire. These figures would be higher if items that are not mentioned in the model diet, but with similar nutritious values, would also be included.

⁶ <https://data.imf.org/regular.aspx?key=61545849>

2.6 Total household net income and living income gap

Box 2.15 Important definitions: Living income gap

Household living income gap = adjusted living income benchmark – net actual household income

Net actual household income = net cocoa income + net non-cocoa income

Adjusted living income benchmark = raw living income benchmark adjusted for:

- Inflation (CPI current period/CPI of the period when the living income benchmark was computed)
- Household size and composition (use of the OECD modified equivalence scale to determine the number of adult equivalent household members)

The final step of this methodology is to calculate the current household net income and compare it to the (household size adjusted) living income benchmark to obtain the value of the living income gap. Based on previous calculations, **total household net income** is obtained by combining the net cocoa income from all the plots managed by the household with the non-cocoa income earned by all the household members from their income-generating activities (whether on or off-farm) (Box 2.15).

Regarding the living income benchmark, it is always computed for a given geographical area and reference household size (x number of adults and y number of children). This value called the **raw living income benchmark** needs to undergo two types of adjustment to account for inflation and differences in household size. A living income benchmark refers to the costs of a decent living at a given period which means that it require regular updating using for example the quarterly consumer price index (CPI) published by the IMF.⁶ The following formula is used to obtain the **inflation-adjusted living income benchmark**:

Inflation adjusted LI benchmark

= Raw LI benchmark

$$* \left(\frac{\text{CPI of the current quarter}}{\text{CPI of the quarter on which the LI benchmark was computed}} \right)$$

The second adjustment concerns the household size and composition as the surveyed households are not always equivalent to the reference household used to compute the 'raw' living income benchmark. Indeed, both the total number of household members and its composition might deviate from the reference, which means the value of the living income benchmark will not be applicable to each household without an adjustment for their size and composition. There are various approaches for adjusting the living income benchmark, but we recommend using the method of the OECD Equivalence Scales⁷ as these take into account the economies of scales realised by pooling the resources of the different household members into the common pot.

In the example of Cote d'Ivoire, the raw living income benchmark was computed for a reference family of 6 members of which 2 adults and 4 children. According to the OECD equivalence scale, this reference family is composed of 2.7 adult equivalent members. This number is obtained by assigning the value 1 to the first adult of the household (usually the head), a value 0.5 to all other adults and children from 14 years old, and a value of 0.3 to all children of 13 and below.⁸

For each surveyed household, the **household size adjusted living income benchmark** is then obtained by dividing the raw living income benchmark by 2.7 (in the case of Cote d'Ivoire) and multiplying the result with the number of adult equivalent members of the household. The following formula can be used:

HH size adjusted LI benchmark

$$= \left(\frac{\text{Inflation adjusted LI benchmark}}{2.7} \right) * [1 + 0.5 \\ * (\text{number of adults} - 1 + \text{number of children from 14 years old}) + 0.3 \\ * (\text{number of children up to 13 years old})]$$

Once all the different adjustments to the living income benchmark have been made, we obtain the **household living income gap** by subtracting the total household net income from the household size adjusted living income benchmark.

For the purposes of reporting, in addition to the average living income gap in a population and its distribution (standard deviation, range, quintiles), it is also useful to indicate the median size of the living income gap to grasp how far below the benchmark the households are. Moreover, we can report the percentage of households with income levels above the adjusted benchmark. This latter indicator is particularly important for studies aiming at measuring the impact of interventions designed to close the living income gap.

⁷ [Adjusting household incomes: equivalence scales \(oecd.org\)](https://www.oecd.org/equivalence-scales/)

⁸ This approach considers the economies of scale obtained by pulling together the resources and expenses of all household members. Furthermore, children from 14 are assumed to have

similar costs of living than adults members while children up to 13 years old are considered to cost less.

3 Collecting intervention data

3.1 Accurate intervention data at the individual level is needed to evaluate the impact of interventions and policies

To be able to measure the effect of interventions and policies, **intervention data** needs to be collected or combined with data on impact indicators. Many organisations report on the outputs that they have achieved within a year, but whether these outputs lead to the intended impacts, can only be measured by combining the data on outcome and impact level indicators with intervention data. Furthermore, there currently is no aligned guidance in the cocoa sector yet on how to collect and use different types of intervention data, which is what this chapter aims to provide.

Figure 3.1 shows an example of an intervention logic (also known as a theory of change) concerning an intervention on farm and income diversification, which is a tool to understand and visualise the rationale behind a programme, the causal relationships between a programme’s activities and its intended outcomes. The logic should be enriched with additional and revised indicators, pathways, outcomes and external influences. Maintaining this method allows comparability between impacts in different places and over time. As shown in the example in Figure 3.1, an intervention logic diagram starts from the actions of the programme and leads to changes in a farmer’s situation. Measured impacts may also be caused by external factors. Since the external factors are not explicitly part of the rationale behind the intervention logic, they are not displayed in the figure, but should be considered in a study.

Box 3.1 Important definitions: Intervention data

Intervention data: data collected either by the intervening party and/or from cocoa farming household members, on a household’s participation in projects, programmes, trainings etc. that aim directly or indirectly to impact a farming household’s living income. Data includes the frequency, degree of participation, quantity of resources provided, intensity and reach of the intervention.

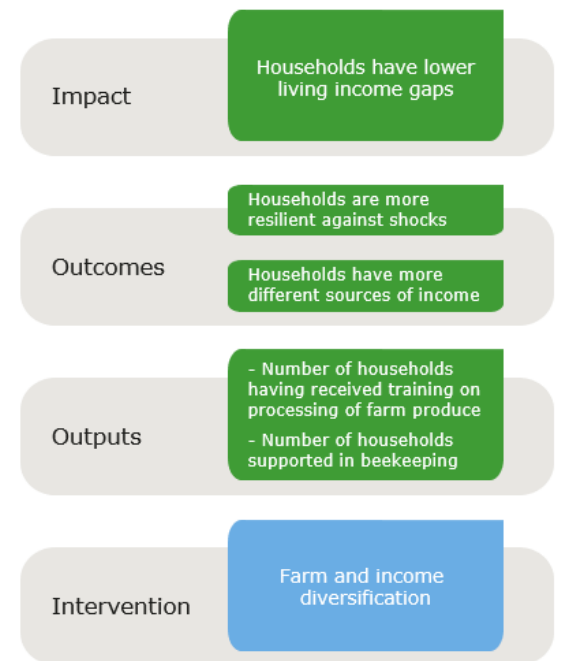


Figure 3.1 Example of an intervention logic of an intervention on farm and income diversification

Information on participation in interventions should only be used for impact assessment when collected at the individual cocoa farmer or household level, as it can then be used to evaluate whether a household that participated in a specific intervention also shows an improvement in outcome and impact level indicators. Outputs reported at higher levels of aggregation rather than the individual level, for example at a community level, are not suitable to be directly linked to impact indicators, unless data on impact indicators is available for enough farmers from a large number of different communities (e.g. at least 100-200 different communities) to be able to find statistically significant differences in key indicators at a community level. This can be checked using statistical power calculations.

3.2 Two types of intervention data

It is recommended to have two different sources of intervention data:

1. Intervention data from the organisations' own monitoring, evaluation and learning (MEL) systems tracking, on who received which intervention and details (e.g. intensity and scale). This includes data from procurement systems on transactions, such as prices, premiums and timing of payment, as well as records of investments amounts per intervention.
2. Intervention data as perceived by the recipients through a survey.

The findings of sources one and two should be compared, as differences between the two serve as important starting points to improve interventions. Farmers who did not perceive to have participated in or benefited from an intervention even though the intervention monitoring system claims that they did may for example indicate issues in the implementation of the intervention such as farmers not having received any intervention at all, or the intervention not being impactful enough for the farmers to remember the intervention. On the other hand, farmers claiming to have benefited from an intervention even though the organisations' intervention monitoring system states that they did not, may indicate spillover effects, or their participation in (comparable) interventions from other organisations. Interventions by other organisations are

⁹ We do not focus on indirect interventions in this document, as the evaluation of these would require different approaches.

important to take into account when measuring impact, as potential improvements in impact indicators may have also been caused by (different, multiple and overlapping) interventions from other organisations. From experience, we know that farmers are often unaware of which organisation(s) have implemented a specific intervention but also that there are often large differences between the intervention rates measured through both different types of intervention data. Triangulating recipient survey data and organisations' MEL data with impact evaluation reports gives a further data verification and validation.

3.3 Data from intervention monitoring systems

Different organisations often have similar interventions or policies under different names. To understand the effect of interventions at the level of the entire cocoa sector, it is important to use consistent and standard categories for data collection on the types of interventions households have participated in. Focusing on direct interventions only,⁹ we identify the following categories with the different types of interventions under each:

- Prices and payments
 - Farm gate prices¹⁰
 - Premiums and bonuses, including the Living Income Reference Price
 - Advance purchase payments
 - Cash transfers
- Financial products
 - Credit, including inputs on credit
 - VSLAs
 - Insurance
 - Savings
 - Banking
- Cocoa production and yield interventions including:
 - Provision of equipment
 - Farm development plans
 - Supply of fertiliser and/or pesticides

¹⁰ These may differ within countries and can be asked during data collection

- Supply of planting material, including tree nurseries
- GAP training, coaching, Farmer Field Schools, demonstration plots
- Organised hired labour, including spraying gangs
- Farm management support/training
- Farm and income diversification
 - Processing of farm produce, e.g. cocoa juice
 - Production and yield interventions on other crops produced by the household
 - Farm diversification (e.g. food production, livestock, agroforestry)
 - Payments for Environmental Services, including carbon credits
 - Off-farm employment - own businesses
 - Off-farm employment – employment by others
- Other
 - Education adults (beyond farm management/ business skills)
 - Community development, including child labour remediation

For the activities under farm and income diversification it is important to also ask about the type of support that was received. This may include: farm/business management support/training, other training, provision of inputs/material, market access support, technical support. Please see the survey question in Appendix 1.

3.4 The intensity of interventions

For each of the interventions implemented or participated in, it is also important to record the different types of intensities. Analyses on impact could thereby not only test whether recipients of certain interventions have higher incomes compared to non-recipients, but they could also test whether those receiving higher intensity interventions have higher incomes compared to those receiving the same intervention at a lower intensity. The concept of 'intensity' manifests in four distinct dimensions, each influencing the impact and effectiveness of the intervention. These intensities can be classified into 5 different categories for each type: very low intensity, low intensity, medium intensity, high intensity and very high intensity. The four different types of intensities are:

1. **The frequency of engagement**, encapsulating the regularity with which the farmer is involved in the intervention, ranging from infrequent annual

interactions to more frequent weekly engagements over the period of the intervention. We identify the following categories for temporal intensity:

1. Very low: 1-2 times over the last year
 2. Low: 2-5 times over the last year
 3. Medium: 5-12 times (e.g. once per month) over the last year
 4. High: more than 12 times but less than 50 times over the last year (e.g. biweekly)
 5. Very high: more than 50 times over the last year (e.g. weekly)
2. **The degree of participation** signifies the number of beneficiaries engaged simultaneously during the activity, varying from one-on-one coaching sessions to training initiatives given to entire villages at once. Note that a low intensity here means a high number of participants at once. We identify the following categories for participation intensity:
 1. Very low: with a very large group of people, over 40
 2. Low: with a large group of people (max 40)
 3. Medium: with a medium group of people (max 15)
 4. High: with a small group of people (max 5)
 5. Very high: one-one-one
 3. In the case of monetary or material interventions, the 'amount' dimension pertains to the **quantity of resources provided**, such as the monetary value dispensed during a cash transfer. We link the monetary intensity to the increase in household income as a result of the intervention, as a share of the average household income of a typical cocoa farming household. This number depends on the country, and may differ over time. In 2018, the household income for a typical family in Côte d'Ivoire was USD 2,346 (Tyszler et al., 2018). We identify the following categories for monetary intensity:
 1. Very low: less than 1% of typical household income (e.g. less than USD 23 in Côte d'Ivoire)
 2. Low: 2-10% (e.g. USD 23-235 in Côte d'Ivoire)
 3. Medium: 10-25% (e.g. USD 235-587 in Côte d'Ivoire)
 4. High: 25-50% (e.g. USD 587-1,173 in Côte d'Ivoire)
 5. Very high: more than 50% of typical household income (e.g. more than 1173 USD in Côte d'Ivoire)
 4. **The reach of the intervention** refers to the share of beneficiaries reached in relation to the total number of potential beneficiaries (e.g. the percentage of farmers reached in a traders' sustainability program). This is not asked to

farmers, but only applied to data from monitoring systems. We identify the following categories for reach:

1. Very low: less than 5% of potential beneficiaries
2. Low: 5-25% of potential beneficiaries
3. Medium: 25-50% of potential beneficiaries
4. High: 50-75% of potential beneficiaries
5. Very high: more than 75% of potential beneficiaries

3.5 Intervention data from surveys

When collecting intervention data from surveys, the enumerator should start with an open-ended question, asking farmers in which interventions they remember having participated in over the last year, and match the answer to the different categories mentioned in Section 3.4. For each intervention, respondents are then asked whether they remember from whom they received the intervention), the frequency, participation intensity and monetary intensity and finally preferably also a subjective evaluation question, aimed at understanding how helpful they perceive the intervention to be in terms of increasing their household income over the past year or in the future. The survey questions can be found in the Appendix 1. As mentioned before, the main goal of these survey questions is not to show the intervention impact of an intervention of an individual organization, but rather the impact of similar interventions offered by any or several organisations. To show the impact of an intervention of a specific organisation the survey data needs to be combined with monitoring data of that organisation on participation in interventions, as well as with an appropriate sampling and data collection strategy for an impact evaluation (including at least two moments of data collection).

3.6 Intervention data analysis

Regardless of whether it concerns intervention data from monitoring systems or from surveys, this data can be used for different types of analyses and reporting. Chapter 4 explains the different options in more detail. The variables on intensities can be interacted with the intervention variables when running regression analyses to show the effect of higher intensities of specific interventions. See Appendix 2 for more information on conducting regression analyses. Moreover, to understand more about the farmers' perception on drivers of change, we also added questions on this topic to the recommended approach questionnaire in Appendix 1. During the analysis phase, it is valuable to compare the results of the quantitative analysis with the perceptions of farmers on drivers of income change.

4 Designing living income studies: sample selection and analysis

Collecting and reporting on data on household incomes and the living income gap requires an appropriate sampling strategy and analysis plan. Both depend on the objective of a study. This chapter provides insights into different sampling and analysis approaches depending on the study's objective.

4.1 Defining the population of interest

Every study on income is related to a given population of interest. That population should be defined at the very beginning of every data collection based on different aspects: geography (e.g. all cocoa farmers in a country or landscape), value chain (e.g. cooperative members, indirect supply chain, households supplying to a certain manufacturer or retailer), a programme (e.g. beneficiaries of an intensification programme), personal characteristics (e.g. youth, women, sharecroppers) or a combination of these (e.g. youth in a specific landscape). When looking at women and youth, it is particularly important to distinguish whether they are usual respondents of surveys (cocoa farmers) and/or household heads as well. In the past, living income studies have predominantly focused on individuals who are registered as cocoa farmers by cooperatives or traders (usually farm owners and/or managers) leaving a large group of farmers 'invisible' in the data. Often, these invisible groups are not recognised for their contribution to cocoa production, and/or they do not directly have access to or benefit from living income interventions. Examples of invisible groups are female spouses of male cocoa farmers, adult children, sharecroppers, and farmers in indirect supply chains. Studies suggest that among these groups you will find farmers who are more entrepreneurial (but limited in assets) and farmers who belong to the poorest segment of farmers. Capturing income information about invisible groups requires a different sampling strategy as explained below.

4.2 Strategies for different objectives of living income studies

Before conducting a living income study, the first step consists of clarifying the objective(s) of the study, as this will directly affect the sampling approach as well as the appropriate analyses to be conducted. This section explores the different potential objectives and, for each of them, describes the implications for the sampling approach and the analyses. Table 4.2 provides an overview of the different research questions, data requirements, and claims that can be made for each objective.

4.2.1 Assessing household incomes and living income gaps

If the objective of a study is to assess household income and the associated living income gap for a specific target population (e.g. country, landscape, participants in a specific intervention, supply chain, etc.) at one specific moment in time, the main requirement is that the sample is representative of the population of interest. A one-time assessment of household incomes may, for example, be used to determine which type of intervention or mix of interventions has most potential in a certain context, but it is generally advised to monitor incomes more frequently and to adjust interventions accordingly. We realise that there is also an increasing number of organizations estimating yields, costs of production, and incomes using basic farmer characteristics. This can be a cost-efficient method of estimation used for intervention design, but as they are based on models, they should not be used to make any claims about (changes in) actual incomes of households or (potential) programme's impact.

When **reporting** on the magnitude of the living income gap, it is also important to report on the key underlying variables used to calculate the living income gap. For each variable that is reported on, the minimum requirements in terms

of reporting are the number of observations, the mean, and the standard deviation for the entire sample as well as for each specific group of interest. Additionally, including the median, the minimum value as well as the maximum value is ideal. Reporting on the median as well as on different quantiles (e.g. top 20%) is important to help show the distribution of farmers within the data, as the means can be heavily skewed by a small set of richer farmers. The list of variables to describe includes, but is not limited to:

- Total cocoa volume produced (in kg and kg per hectare)
- (Cocoa) farm size
- Costs of cocoa production, split into material input costs and labour costs, as well as costs per hectare and costs per kg of cocoa produced
- Net cocoa income, net cocoa income per hectare and net cocoa income per kg of cocoa produced
- Cocoa dependency: share of total household net income derived from cocoa sales
- Net income and costs of production for other income generating activities
- Total net household income
- Number of children and adults in the household
- Living income benchmark (raw and adjusted)
- Living income gap (mean, median, percentage of households above the benchmark and household income as percentage of the living income benchmark)

4.2.2 Monitoring the changes in the living income gap over time

If the aim of the study is to track the changes in income and the living income gap of a specific target population, it is important to make sure that the sample is representative of the population of interest, and that the data that is being compared covers the same time span, i.e. one year, so that household incomes over two different years can be compared. Data can either be collected on the same sample of households, or a different sample of households within the same target population (e.g. country, landscape, participants in a specific intervention, supply chain, etc.), as long as the sample is representative of the target population in both cases.

The simplest way to check whether there is a statistically significant difference between two moments in time in a variable, for the same sample, is through a

paired t-test. Statistical software, such as R or Stata comes with functions to do all this automatically. When **reporting** the results from a t-test, the following should be presented, to allow interpretation:

- The sample size for each group.
- Mean and standard deviation for each group.
- The t-statistic and associated degrees of freedom (df).
- The p-value (probability of having a significant coefficient).

4.2.3 Comparing different groups of households

It is often insightful to compare groups of people within one study. For example, to compare how the living income gap differs between male-headed households and female-headed households or between visible and invisible groups of cocoa-producing households. Statistical tests are not only used to show the differences in indicators over time, but also to show differences in indicators between groups. When reporting on differences between groups it is important to use unpaired t-tests rather than paired t-tests, which is also easy to do using statistical software. In terms of reporting, the same guidelines apply.

While simple, the results of t-tests can be misleading. For example, if we find that female headed households have a higher living income gap than male-headed households, this difference may not be driven by the gender of the household-head. There may be variables that are missing from our analysis that drive the difference. For example, female headed households may be smaller than male-headed ones, and it is this difference in size that drives the change. To control for these other variables, a regression approach is a good fit (see Appendix 2 for more information). All common statistical software can perform regressions. The variables to be controlled for can differ depending on the purpose of the analysis and the type of variable, but some frequently used control variables include: the gender of the household head or respondent, age, level of education, (cocoa) farm size, age of the cocoa farm, whether the farm is owned by the respondent or not, as well as regional indicators. Regional indicators such as rainfall patterns, micro-climate and prevalence of cocoa diseases like the Cocoa Swollen Shoot Virus Disease can significantly influence household incomes through their effect on cocoa yields. For meaningful comparison between regions, these geographical differences should be controlled for.

4.2.4 Assessing the impact of interventions on income levels and the living income gap

To claim any impact on the income levels and the living income gap of a program’s participants or intervention beneficiaries, it is not sufficient to only present the changes over time from two different groups. The impact of an intervention is the difference between what happened when the intervention was made, and what would have happened if the intervention had not taken place. If we would implement an intervention aimed at increasing farmers’ income from cocoa, and we find that their incomes are USD 100 higher after the project than they were before, this is not enough to know the impact, as the increase may have also been caused by something that occurred outside the scope of the project. To find the impact, we need to find out what would have happened if we never implemented the intervention at all (establish a **counterfactual situation**). It is important to note that the measurement of an intervention’s impact should be done preferably by looking at the same sample of farmers over time (panel data). We recommend collecting the data on these groups at the same time to ensure comparability over time. Ideally, a minimum of one year between two rounds of data collection should be observed to allow the measurement of annual income levels.

To create a counterfactual situation, we must identify a group of households that is as similar as possible to the beneficiaries of the program/intervention. The impact is then measured as the difference between the beneficiaries (or treatment group) and the non-beneficiaries group (comparison). One strategy for identifying the comparison group is **randomisation**: splitting a large group of people in two randomly assigned groups yields two groups that are expected to be the same. If randomization is not possible (or desirable) it may be possible to create a comparison group using **statistical matching** procedures. Yet, such procedures require to collect more data to ensure a good match between treatment and comparison groups. Statistical matching procedures, in particular propensity score matching (PSM), can also be used to correct for underlying differences between an intervention and a comparison group when there is no panel data. PSM allows for the creation of an artificial control group by matching each observation from the treatment group with an observation from the control group with similar characteristics (e.g. farm size, gender of the household head).

Once a good counterfactual has been established, we can then assume that in the absence of the programme/interventions, the treatment and comparison groups would have followed the same trends in the main variables of interest for the study. Therefore, by comparing them before and after the intervention (**Difference-in-Difference** approach, noted DiD), we can isolate the impact of the intervention as the difference in trend between the two groups. Table 4.1 below demonstrates the importance of identifying a good counterfactual. It provides income data for two groups of cocoa farmers: one group that benefitted from a project between 2020 and 2025, and one group that did not. The beneficiaries and non-beneficiaries are not the same before the project: the beneficiaries have higher income. We therefore can’t use the difference in incomes in 2025 (USD 500) as our measure for impact. But if we assume that in the absence of our project, the beneficiaries would have seen the same increase in income as the non-beneficiaries (USD 100), we could then get a measure of impact by consider the difference of these two differences. This is the principle behind the DiD approach. For more guidance on how to implement a DiD approach, see Appendix 2.

Table 4.1 Example of changes in income over time, USD

	Income 2020	Income 2025	Difference
Beneficiaries	1,000	1,500	500
Non-beneficiaries	900	1,000	100
Difference	100	500	400

As it might be undesirable to only engage with specific groups of farmers for a certain intervention, another option is to use a **quasi-experimental pipeline design**. This implies that different farmers would be participating in the project gradually in different years, allowing us to compare farmers based on their different stages in the project support. This so-called pipeline design can be used to (better) control for unobserved differences, in the absence of experimental designs and a comparison group (Khandker 2010, Stern 2012, DCED 2013). The main assumption is that support is indeed implemented gradually among the targeted beneficiaries. We can then compare the average status of farmers who have already received support with the baseline status of the farmers who have not received it yet. Let’s however note that a good comparison group must be similar to the treatment group in terms of key outcomes at baseline. Any deviation(s) should be corrected in the analysis of the collected data.

Table 4.2 Overview of the potential objectives, examples of associated research questions, data requirements and claims that can be made

Objective	Examples of related research questions	Data requirements	Claims that can be made
1. Assessing household incomes and living income gaps at one moment in time	<p>What is the current income level and living income gap of a group of cocoa farmers?</p> <p>What are opportunities to increase revenues and/or reduce costs?</p>	One measurement on a representative sample of the target population	<p>This allows for reporting on the average household within a specific population of interest, e.g.:</p> <p><i>'In season 2023-2024, the living income gap of female farmers in our supply chain was 3000 USD, while the income gap of male farmers was 2000 USD'.</i></p>
2. Monitoring changes in incomes and living income gap over time	<p>What is the change over time in income for an average farming household in a given supply chain/programme /landscape/country?</p>	Minimum of two measurements of a representative sample of the population of interest (possible break down by groups of interest – different program components, gender, age, etc.)	<p>This allows for claims on the change in the situation of households within a population of interest, e.g.:</p> <p><i>'Farming households participating in intervention A have increased their income with 5% while farming households participating in intervention B increased their income with 7%'.</i></p> <p>If the sample has changed over time (different households in the second measurement than in the first) this should be indicated when making the claim, as it may explain potential differences over time.</p> <p>This is not sufficient to make any kinds of claims on impact of interventions.</p>
3. Comparing between different groups of farmers	Do beneficiaries of an intervention have higher incomes compared to beneficiaries of intervention B?	Minimum of one measurement of a representative sample of at least two different populations of interest, out of which one can be a control group.	<p>This allows for claims on the differences between the households in two different populations of interest, e.g.:</p> <p><i>'Farming households participating in intervention A have higher incomes in the season 2023-2024 compared to households that did not participate'.</i></p> <p>This is not sufficient to make any kinds of claims on impact of interventions.</p>
4. Evaluating impact of living income interventions	<p>What is the impact of an intervention on beneficiaries' income levels and the living income gap?</p> <p>If data is collected on multiple interventions: What interventions have the greatest (combined) impact on income levels and the living income gap?</p>	<p>If random implementation of interventions: representative sample of beneficiaries at 2 (or more) periods of time</p> <p>Otherwise: representative sample of treatment and comparison groups with baseline and follow-up surveys</p>	<p>This allows for claims on the impact of interventions on incomes or the living income gap, e.g.:</p> <p><i>'Participation on intervention A is associated with a positive significant relation to (or: 'impact on') household'.</i></p>

4.3 Dealing with challenges on finding comparison groups for impact assessment

In reality finding a comparison group that has not had any previous interventions yet may be challenging given the context of the cocoa sector. Many households have already been engaged in various different interventions by different organisations. When collecting data from households that, according to an organisations' own monitoring system should have not participated in the intervention (or combination of different interventions) that is being assessed, it is recommended to ask the envisioned respondent before the start of the interview whether they, or anybody else in their household, have or have not participated in any comparable interventions, regardless of the provider of the intervention, in the last couple of years. If they did not, the interview can then proceed as usual. If they did, the enumerator would have to move to the next person on the list of non-recipients of the intervention that is being assessed. Note that it is not an issue if households have participated in other interventions than the one being assessed, but it would be good to collect information in their participation in other previous interventions, so that this could be accounted for in the analysis. In the case where it is not possible to find enough households who have not participated in a specific intervention to serve as a good comparison group, differentiation could be made in terms of the intensity of the intervention participation (i.e. comparing farmers who have received high-intensity training with those who received low intensity training, see chapter 3), or alternatively qualitative approaches can be used instead to learn about both the effectiveness of the interventions as well as the potential points of improvement of the intervention.

4.4 Developing the sample selection strategy

To accurately represent the population of interest, random sampling of respondents is of vital importance. The easiest way to select cocoa-producing households is by using farmer lists. Cocoa traders and cooperatives, for

example, have lists of their members/producers. To randomly select farmers, simply number all the farmers on the list and use a random number generator to identify the corresponding farmer. How to best access farmer lists is dependent on the organisation of the cocoa supply chain in a country and region. In Côte d'Ivoire, for example, cocoa farmers are registered with and sell to cooperative who then sell to traders. Here, the cooperatives are the best entry point. In Ghana, most cocoa farmers are registered with Licensed Buying Companies who sell to the Ghana Cocoa Board. Farmer data can be obtained from these companies or from the Cocoa and Health Extension Division (CHED) of the Ghana Cocoa Board. If there are other lists available, these can be used as well.

The major disadvantage of sample selection based on farmer lists is that these lists are not exhaustive. Many farmers, especially the more invisible groups, are not always represented on the lists. If we want to reach these farmers or if there are no (comprehensive) farmer lists available, spatial sampling can be a good alternative. Spatial sampling can be used to find cocoa farms that are not on farmer lists. Once a person has been identified that is associated with the selected farm, they could then be asked to contact invisible groups of interest such as women or sharecroppers to participate in the interview. One simple way of spatial sampling is transect walk sampling. In this method, a central point is chosen (e.g. centre of town) as starting point. Then, in a structured way, researchers walk in multiple directions (North, South, East, West) and select every Nth household. While in this way the sample is not restricted to registered farmers, the final sample is dependent on the chosen starting point and more remote farms may not be reached.

For more comprehensive spatial sampling, GIS techniques can be used. Through technical advancements in remote sensing in the past decade, high quality land use maps have become available that distinguish (agroforestry) cocoa farms from (secondary) forest and other land uses at high resolution (10x10m).¹¹ Moreover, as part of deforestation monitoring efforts, in some landscapes all cocoa farms have been identified individually by different value chain actors. To randomly select cocoa-producing households using land use maps, allocate random points (using GIS software) to the cocoa areas identified on the map.

¹¹ See for example: <https://nk.users.earthengine.app/view/cocoa-map> and <https://ghana-national-landuse.knust.ourecosystem.com/interface/>

Then the farm corresponding to the GPS coordinates of the GPS point should be visited to identify the corresponding household. Hence, in this case the cocoa plot is the starting point instead of the farming household. An advantage of this method is that all cocoa-producing households can be selected, including the ones that are not covered by farmer lists. A disadvantage is that the costs of visiting the cocoa farm and finding the household are high. One way to lower these costs is by using a stratified approach. In this case, first villages can be selected before selecting random points around these villages. This way the data points are concentrated around the villages instead of spread throughout the whole landscape and travel and searching costs are decreased.

4.5 Calculating the required sample size

When calculating a sample size, we recommend adopting the standard margin of error of 5% to obtain estimates that fall within the 95% confidence interval. In other words, when the selected sample size is sufficient, the computed values of our indicators have 95% of chances to be representative of the true values within the population of interest. In addition, as the distribution of income levels is usually quite widespread (some very low values and often a few large values), we recommend considering a ratio mean to standard deviation of one-to-one when determining the sample size. This hypothesis will ensure that we capture the minimum detectable effect (MDE) when conducting comparison between different groups and/or impact evaluation of interventions/programs. For the latter category of studies, the sample size should also include a margin to account for possible attrition of surveyed participants during any follow-up surveys (e.g. due to not being able to reach them again, no consent, death, travel, replacement, etc.).

Up to a finite population of 10,000 individuals,¹² the sample size is calculated as a fraction of the population given the retained margin of error and associated confidence interval. For example, considering all our previous assumptions about income distribution and confidence interval, the sample size to obtain a precise estimate of income levels for a finite population of 10,000 individuals

¹² The threshold of 10,000 individuals is considered as larger populations affects only marginally the sample size required to ensure representativeness.

would be 1,332, which would still need to be adjusted for attrition depending on the scope of the study (i.e. if the type of study requires panel data, the attrition rate would increase depending on the number of different data collection moments and/or the time between the data collection moments). If we want to measure the impact of the intervention, we need to define the minimum effect we want to detect knowing that the smaller the effect, the larger the sample size required to capture it. For example, if we expect the intervention to induce a 20% increase in income levels of beneficiaries, we then need a sample size of 1,338 individuals/households (669 for treatment and 669 for comparison). A smaller sample size will be unable to detect a significant change of 20% or lower.

4.6 Using qualitative methods

As the examples in this chapter may show, quantitative analyses can show the differences in key variables over time, and they can also show which treatments are associated with, for example, decreasing living income gaps. However, these analyses rely heavily on the quality and accuracy of the variables used, and do not give any more detailed insights into the context in which the cocoa farmers live and into 'why' and 'how' change happened and how the intervention influenced it. To assess that further qualitative research methods are needed.

Qualitative data collection and analyses serves three important purposes (Schneider, 2021):

1. **Contribution:** It bring in evidence on how change happened and how the intervention influenced it.
2. **Voice and inclusion:** It can present participants' different views on the effect of the interventions, their view on the relevance of interventions and potential challenges in adoption.
3. **Triangulation:** Using mixed-methods strengthens conclusions about the intervention's effects.

To facilitate learning and continuous improvement, it is recommended to use mixed methods and complement the survey with in-depth interviews with key local stakeholders (such as purchasing clerks in Ghana and cooperative leaders in Côte d'Ivoire) and focus group discussions with small groups of farmers, as well as farmers who belong to more 'invisible groups' to learn more about their (different) ambitions, which (different) challenges they face (in general or related to specific interventions), and what they need to tackle these challenges.

4.7 Dealing with existing data

There already are many studies that have been conducted in the cocoa sector, and data collection efforts are only expected to further increase as a result of legislative developments. To enable collective learning on the effectiveness of interventions as a sector, it is important that key results and at least aggregated data points are shared. The sharing of key results and comparison with other studies on the same topic allows for determination of the validity of results. If the results are different and the sample is comparable and representative in all studies analysed, or there is no more comparable sample, it needs to be examined what might be the cause of the discrepancies. The reason why it is important to always report on the number of observations, the mean and the standard deviation for any variable, is that it allows for comparison between studies without having access to the entire dataset of another study. There are plenty of online (unpaired) t-test calculators that only require these three data points of two different samples to conclude whether they are significantly different from each other. To compare between studies, it is of course important to first make sure the variables being compared are the same, and that there are no large differences in the wording of the questions or the build-up of the indicators, which can be avoided by always using the questions as developed in Chapter 2.

When comparing study results to other results, it is important to (self-)assess whether and when existing data is sufficient, comparable, robust and of high enough quality to accurately compare against. Whether existing data meet these requirements depends on the indicators used, on the sampling approach and on the intervention data available. Evaluating existing data always starts by understanding the origin, timing, purpose and scope of the data. If the data was

collected in relation to a specific program this may have implications for the sampling.

- **Origin:** who collected the data? Can we trust neutrality and quality of data?
- **Timing:** when was it collected during the cocoa season? What was the recall period for questions on income? Longer recall periods lead to higher recall bias compared to shorter recall periods. Questions can cover multiple different timeframes but for comparability reasons, the final reporting period of household incomes should always be one year.
- **Purpose:** be aware of for what reason the data was collected. This could for example include monitoring purposes, the evaluation of a programme, or another reason. The purpose may have had an influence on the sampling strategy, and may therefore make existing data less comparable or representative.
- **Scope:** the data collection scope refers to the types of respondents (e.g. are the respondents assumed to be knowledgeable about the questions asked) as well as the number of respondents per region/type (e.g. how many women were questioned). The scope may also have had an influence on the sampling strategy, and may therefore make existing data less comparable.

4.8 Additional analyses, cleaning and reporting

Before conducting any analyses or reporting, it is important to thoroughly clean the collected data. Outliers should be determined and dealt with at the level of each collected survey question, not only at the final indicator level such as household income. The cleaning of outliers happens on a case-specific basis, taking into account the distribution of the question of interest. Outliers should generally not be replaced by mean or median values, unless nearly all respondents have the same response, or if information from external sources is available (e.g. on cocoa prices). Missing values in underlying indicators, whether directly from the survey or through the cleaning of outliers, will therefore usually lead to missing values on the final indicator. To account for this, it is highly suggested to take this into account during the sampling stage by oversampling.

When reporting on the living income gap it is important to calculate the living income gap at the household level, and to report (e.g. mean, median, s.d.) on the aggregated living income gaps for a specific population of interest. Reporting of results should always be gender disaggregated, and we also strongly suggest reporting on any underlying indicators (e.g. costs of production, income sources). For guidance on visualising the living income gap, please refer to the 'Guidance manual on calculating and visualising the living income gap' (Tyszler and De Los Rios, 2020).

As mentioned in Chapter 1, data collection on incomes and living income gaps is often combined with other topics of interest, which may help in explaining the (potential) drivers of change. Topics such as pest and disease or household (financial) decision making will require context-specific additional questions to be added to the questionnaire. Using the questionnaire of the recommended approach, as presented in Appendix 1, the following additional analyses on drivers of living income gaps can be conducted:

- **Land size:** To assess the relationship between land size and cocoa and/or household income or the living income gap, we recommend including all land owned by and worked on by the household, thus also the land sharecropped out. The reason for including land sharecropped out is that the household also earns an income from such a sharecropping arrangement. These analyses can be done for both land used for cocoa production, and total farm size.
- **Cocoa volumes:** To assess the relationship between cocoa volumes produced and cocoa and/or household income or the living income gap, we recommend to include all cocoa produced and owned by the household. This also includes the cocoa received by the household from a sharecropper as they contribute to household income.

Cocoa volume for these analyses = SUM (total volume produced (kg) on own land, land sharecropped IN and cocoa volumes received from sharecroppers for land sharecropped OUT)

- **Price and premiums:** The relationship between prices or premiums and cocoa or household income as well as the living income gap can be analysed to assess the direct implications of price or premium increases on living income gaps.
- **Cost of production and household labour deployment:** Regarding the cost of production, it can be assessed what the relationship is between the cost of production (in total or for material and labour cost separately) per hectare and cocoa yield or between the total cost of production and total volumes produced. In the recommended approach detailed information is collected on the use of labour (household, permanent labourers, temporary labourers, labour groups and community labour). This information can be used for additional analyses on household labour deployment, for example to assess the return on household labour investment in terms of the net income from cocoa earned per day spent by the household on specific cocoa production activities.
- **The relationship between cocoa yield per hectare and net cocoa income per hectare:** To assess the relationship between cocoa yield and net cocoa income per hectare we recommend making sure the analyses are conducted based on land used for cocoa production by the household. When a household sharecrops out land and receives cocoa as a form of payment, such volumes are not included in this calculation (e.g. in the form of net cocoa income per hectare) as this cocoa is produced on land not worked on by the household. If the households sharecrops in land and pays the landowner through delivering cocoa, it should be ascertained that the net cocoa income per hectare (of own land and land sharecropped IN combined) is based on the same calculation as the yield per hectare (own land and land sharecropped IN).

Cocoa yield (kg/ha) for these analyses = SUM (total volume produced (kg) on own land, land sharecropped IN) / SUM (land size (ha) own land, and land sharecropped IN)

5 Enumerator approach

An enumerator is a person who collects data, guided by a protocol that describes the conditions to be met and the steps to be taken to ensure good data collection. The manual is structured to equip the enumerator with the knowledge and skills needed to administer the survey and collect good data and/or information. Large surveys usually involve a team of enumerators. In composing a team, ensuring diversity of team members (in terms of gender, ethnicity and age) is essential.

5.1 Selection of enumerators

The selection of enumerators is based on a set of criteria. To be selected, the enumerator must meet the following criteria:

- Have a minimum level of education, depending on the country.
- Adapt well to working in a group
- Teamwork in rural and urban environments
- Good communication skills
- Good listening skills
- Good command of the local language and dialects.

An important aspect to consider for enumerator selection is that who the enumerator is, or which organisation employs the enumerator should not influence the answers given by respondents. Some companies or program implementers ask their technical staff, who also coach or train farmers, to collect information from or conduct surveys from households. Also, enumerators can be employed who might know the farmers because they are from the same area. This can bias the answers of the respondent because there might be a conflict of interest ('If I give this answer, I might get more help'). In addition, the enumerators might already know much about the farm and farmer which could influence the way questions are asked or how information is recorded. It could also lead to a discussion during the interview in which the enumerator gives or is asked to give

advice. In cases where company or implementer staff conduct the interviews, we recommend that they interview farmers in another area than where they work, such that they interview farmers they do not know. In very specific cases, it could help if a farmer knows the enumerator; for instance, if they do not trust other persons and would not be interviewed if they do not know the enumerator. Or if they would give different answers because they do not know the enumerator. If a company or implementer wants to have enumerators conduct the interviews with farmers they know, which is not recommended, they should assess possible bias, how to mitigate the bias risk, and document this such that the data analysts know data could be biased. It is important for the enumerators to know the setting and the farms, farming and household situation to conduct the survey well. This is addressed during the enumerator training.

5.2 Enumerator training

The training consists of reviewing all data collection tools and becoming familiar with the survey questions. Training involves several stages.

- Study presentation: the presentation of the study consists of outlining the stakes involved in the work to be carried out. This presentation focuses on the objectives, scope and methodology of the survey. It also includes educating the enumerators with topics related to cocoa production and other income generating activities, so that the enumerator can support respondents appropriately when conducting the survey.
- Study objectives: the objective of this data collection project is described, including any specific objectives defined.
- Geographical scope: the geographical scope corresponds to the area covered by the study. The localities concerned must be clearly geo-referenced to the data collectors.
- Survey methodology: the survey methodology aims to indicate the criteria for inclusion of target populations, the data collection technique (random, quota,

purposive, sampling, etc.), the sample size (number of individuals to be interviewed).

- Survey translation: usually the survey is conducted in the language that is best managed by the respondent. To avoid things getting lost in translation, it is good practice to discuss the underlying meaning of questions and harmonise the way questions are phrased in local languages.
- Discussion on how the enumerators and their supervisors should work in preparing the interviews (making appointments up front), remuneration of respondents for the interview, how enumerators should and should not be dressed and what can and cannot be discussed with respondents (e.g. no politics, religion and enumerators also should not give advice to the respondents on e.g. farm management).
- Training of enumerators using the survey: enumerators need to get acquainted with the survey by testing the survey on each other at least twice. This also includes role play to practice dealing with unwilling or unpolite respondents, and should include making sure that the setting is appropriate for conducting the survey (e.g. not too many distractions).
- Create a protocol for data quality assurance to ensure enumerators and supervisors know what is expected from them in terms of how the filled-out interviews will be checked and what needs to be done in case there are inconsistencies, or answers are not complete.

5.3 Testing of data collection tools

After the training of the enumerators, the data collection tools are tested. During this stage, enumerators are invited to ask all kinds of questions to help them understand and appropriate the tools. Enumerators can also make proposals and suggestions on the content of the tools. Preferably, the tools are tested in the field before the full survey is rolled-out. Field testing should be done using a relevant group of households, to ensure that the tools are well contextualised. There should still be enough time to adjust the tools if the field testing proved that certain questions did not work well or were not well understood by respondents. This means that at least 3 days are needed for the enumerator training and the testing of data collection tools.

5.4 Conducting the survey

Collecting data requires a great deal of tact. Some respondents may show a lack of interest or expect compensation in return for their participation. It is therefore up to the enumerator to develop personal tactics for good cooperation and collaboration, and to manage expectations. This section presents some general guidelines on how to establish good rapport with respondents and conduct a successful survey. It also includes some tips on how to conduct the survey and information on the enumerator's attitude during the interview.

The enumerator is expected to:

- Master the survey using the manual;
- Go to the data collection site;
- Introduce themselves to the respondent (for example with a badge and letter of introduction);
- Verify that they are speaking to the right respondent in line with the sampling approach;
- Clearly present the project, in particular its aims and objectives;
- Insist on the confidential nature of the data to be collected and get consent;
- Manage expectations in terms of time. Avoid putting pressure on the respondent, and if possible make another appointment if for any reason the survey cannot be fully completed at the first interview;
- Make sure that all questions have been asked, and that all sections have been properly completed at the end of the interview;
- Thank the respondent.

During the course of the survey, they will come into contact with people of different needs. This can happen in various ways:

- Some will simply refuse to participate in the survey;
- Others will demand personalised letters of introduction;
- Some will agree to participate without asking any questions.

Therefore the enumerator must be prepared for these kinds of situations and any questions that may arise.

There are several rules of conduct for enumerators when dealing with respondents. Some of the most important elements are:

- Their physical appearance is very important. We recommend that they dress and style themselves humbly, without company logos or clothing;
- The way they speak is also very important. Their language must be adapted to the context of the visit. They need to speak clearly, using simple words and phrases, with politeness and confidence, but without arrogance;
- In general, they need to control their introduction so as not to appear anxious or hesitant during the interview;
- Above all, they should not give the impression that they are reciting a text that they have learned by heart;
- Get straight to the point and avoid lengthy explanations which may make the respondent suspicious;
- To avoid propagandistic speeches and not to create unnecessary discussions;
- Put mobile phones on silent/airplane mode to ensure calls or texts are not disturbing the interview.

Some additional practical tips that can make the survey more successful:

- Try to accommodate female enumerators to interview female respondents.
- Find a quiet (open) space in the shade to conduct the survey, so the respondents can discuss things freely.
- Do not allow any other person to listen into the conversation, including the spouse. If a person joins the discussion with the respondent, kindly ask to only speak to the respondent, as otherwise the other person can influence the responses of the responses. If this is not possible at the time of the interview, try to reschedule to a later moment. If that does not work, contact your supervisor.
- All enumerators should take several printed versions of the survey with them and several pencils, in case the tablet runs out of batteries or stops working. After filling out the survey on paper, it should be digitalised afterwards.
- The enumerator should check the answers that farmers gave for consistency. For instance, if a farmer indicated to have sold 4,000 kg of cocoa while the cocoa farm is 2 hectare, the enumerator should check whether the answers are indeed correct because the yield is really high. The survey is coded in software such as ODK and such software can pose questions on certain answers if they are seen as inconsistent with other questions. In this example,

the app could ask whether the farm size or total volume produced is correct because the yield is 2000kg/ha.

5.5 Confidentiality and data ethics

Confidentiality assures that personal information provided by respondents will be kept in accordance with regional, national and/or international standards and legislation. Personal information should not be disclosed inappropriately, and paper and electronic data are treated with appropriate levels of security. Below we provide some practical tips on how to address confidentiality; for full compliance the national and international standards need to be checked and applied. The authors accept no liability for any damage resulting from the use of the results of this study or the application of the advice contained in it.

Confidentiality standards and legislations have implications for:

1. The preparation of the interviews: obtain approval from the relevant local governmental organisation to collect personal data (In Côte d'Ivoire this is ARTCI.¹³ Ensure to collect as little personal information as possible from as few respondents as possible. Train the enumerators and data analysts (anyone collecting or working with personal data) in the protocol of confidentiality towards respondents.
2. During the interviews: Ask the interviewees for their consent and indicate that they can always withdraw such consent including how they could get in touch with the researchers after the interview. This can be done for instance by giving them a brochure in their language with more information and contact details, but such a leaflet cannot replace a verbal explanation. Stop the interview if they do not give consent, even half-way through the survey. If they stop giving consent remove the interview data and document the number of people not having given consent and having stopped giving consent.
3. Storing the data: make sure as few people as possible can access the personal data. Anonymise the data as soon as possible after giving the respondents a new unique ID for the study. As few people as possible have access to the information that connects the new unique ID to the personal data including the address, to enable revisiting the respondents after some

¹³ See: <https://www.artci.ci>

time. If an organisation collecting the personal data shares personal data with other parties, a data processing agreement should be signed.

4. Reporting: Personal data must be treated anonymously. Information in a report on a specific respondent should not be traceable to that respondent. For example, publishing maps based on polygons could lead to people who have seen such a map visiting the farm. This would become more sensitive if such information on mapped farms can be connected to other personal data.

5.6 Data quality assurance

Once the survey has been completed, enumerators need to ensure that all the required information has been collected and is of good quality. Checking is one of the fundamental phases of their work involves careful attention, a clear understanding of the questions asked, and strict adherence to data collection procedures. A poorly completed or incomplete survey requires them to return to the field, and is a waste of time with major financial implications. So, before transmitting the data, data collectors must make sure that the survey has been correctly completed. To this end, they must review the entire survey and correct any inconsistencies or imperfections. Supervisors perform checks of the completed survey according to the protocol established during interviewer training, on the same day they have been filled out. This is to ensure that any inconsistencies or missing information can be addressed the day after by visiting (or calling) the respondent and ask some questions again.

The protocol should include what will be checked, how, by whom (the supervisor) and when. Besides completeness checks, also consistency checks should be done to assess whether the answers correspond. Data points that should be checked for consistency are: Farm size, yield per hectare (which is calculated based on survey responses), cost of production in total and per hectare, the number of household members and children (e.g. if there are many adults and few children), tree age in relationship with yield per hectare, farm size in relation with cocoa or total household income, the total amount of non-cocoa income earned, the proportion of income from cocoa, the total net household income. Some of these data points need to be calculated; this can be done through creating a standardised analysis protocol that is implemented every day/every few days.

With the development of technology, completed surveys can be transmitted directly to a computer server, although it depends on the location and signal how often this can be done. Supervisors are responsible for ensuring the quality and consistency of the data collected. In the event of inconsistencies or errors, the data collector is called upon to provide correct data or consistent information.

Besides living income specific guidelines, general best practices for data collection apply. These are captured in Box 5.1.

Box 5.1 Best practices for data collection

Adjust to local context

Interviews should as much as possible meet the realities and capabilities of the respondents. This means that they should be undertaken in local language, in simple understandable wording. Enumerators should be well-trained and respectful.

Measurement units and time period

In general, it is best practice to let respondents answer in locally relevant measurement units and to convert these units to standard international units (e.g. ha, kg). Besides the measurement unit, the measurement time period should be made very explicit throughout the survey. All reporting is done over a fixed time period: one year, covering the main and light season of the cocoa production cycle.

Informed consent and privacy

Informed consent should always be obtained before the start of the survey. Participation is voluntary and anonymous and respondents have the right to stop the interview at any time they want. Data processing should be in line with ethical standards: data should be anonymised and stored safely. Analyses should be aggregated to avoid traceability to the participating households. Any specific case study should be explained to the participant and requires a separate consent for using individual/non-anonymised information.

Data collection

To minimise enumeration problems, data is collected with the use of a pre-programmed and tested survey tool on tablets. Software such as ODK can be used and internal checks with the logic and constraints should be programmed into the survey to avoid inconsistencies.

Remuneration of respondents

Without direct benefits to survey participation, respondents should always be remunerated fairly for their time spent as well as for their travel costs.

Limiting the number of interviews per day and time spent per survey

Enumerators should not be collecting more than 3 to 4 surveys per day to prevent researcher fatigue and should have enough time between surveys to check surveys for completeness and to rest. The time per survey ideally should not exceed 1.5 hours.

Remunerate enumerator organisations and enumerators fairly

In the terms of references for hiring an organisation employing enumerators ask that they include information in their proposal on what enumerators earn and ask about their working conditions regarding working times (per day and per week), and what happens if for instance there are a few days without interviews because of logistical challenges. Check whether the proposals follow ILO guidelines and local labour laws and do not hire an enumerator organisation that does not comply.

6 Farming household engagement in studies, learning and intervention design

It is important that farming households and farmer groups (e.g. cooperatives) become more integrated in study implementation so that they can indicate what information is important to them. It is also important that more data and information is shared with households and producer groups, in order for them to make use of the data, study results and insights and can also access data about their household and farm, or cooperative, to allow for better decision making. Finally, it is important that they can take part in learning sessions and that they have a voice in intervention design. Engaging with farming households in such a way enables policies and interventions to be as effective as possible. Please find recommendations below for farmer engagement in different phases of the study. It is also important to engage with other stakeholders during and after the study, but we focus on farming households in this section as they are often considered the most challenging group to adequately engage with.

Involving farming households in the design of the study

To involve farmers in the design of the study, it is recommended to conduct focus group discussions with representative group(s) of farmers, also including women, sharecroppers and youth, on the objectives of the study, and to conduct interviews with producer group leaders. The insights from these discussions should be used to adjust the scope of the study and/or to add questions to ensure the study results will be as valuable as possible for addressing the needs of the households in intervention design. Finally, it is important to discuss in what form the farmers and cooperative leaders would like to receive the results so that information and data sharing can be organised in an adequate way.

Example: A study team prepared for the evaluation of a farm management training program. During focus group discussions with households at the start of the study, it became clear that households faced important production challenges which would not have been addressed through the training program. The study design was expanded to generate insights about such challenges to be used for future intervention design.

Involving farmers during the implementation of the study

The data collection tools should be discussed and tested with representative farming households (including women, sharecroppers and youth) to ensure that the respondents understand the questions well because the appropriate terms and language are used. If respondents are expected to fill out information (e.g. as in the Farmer Field Book approach) the forms should also be comprehensible and easy to use for the households. If the forms need to be stored by the household for later use, it should be discussed how this can best be done. In addition, the timing of the interviews and the location of the interviews should also be based on the availability and possibilities of the households (e.g. including safety considerations for women as well as considering the duration and timing of household and care activities). Finally, before or just after the interview, useful information can be shared with the respondent.

Example: Farmers indicated to a study team that they would like to be interviewed very early on the day because they would go to the farm afterwards. The study team organised this to happen. In another study respondents received information on their farm size assessed through polygon mapping after the interview was completed, to avoid influencing the self-reported farm size.

Sharing back information with stakeholders for sensemaking, learning and intervention design

It is important that stakeholders, such as farmers and producer organization, can learn about the results of the study, and that they can make sense of the results to show what they could mean for them and their organisation. This means that study results should be shared appropriately to facilitate such learning and sensemaking, so stakeholders can use the data and insights for decision making. Examples on the form, content and process of data and information sharing are provided below.

-
- **Form:** Publish reports and leaflets and other outputs in appropriate language, length, and style so that stakeholders can understand the information and make use of it. Stakeholder needs and requirements differ greatly, from short one-pagers with key messages to interactive dashboards with detailed data to be explored.
 - **Content:** It is important to ensure that the content of the result allows for stakeholders to: understand the information (e.g. if income is low, that it is explained why this is the case), and find information which applies to them and their work (for instance a cooperative might need different types of information/data than a farmer or a policy maker). For farming households, receiving information for their specific household and farm, compared to peer farmers (e.g. farmers in their community) has been found to be useful. Examples are profit and loss statements for farmers for each season, and changes over time if data is collected for multiple seasons, as well as comparisons with the situation of, or changes over time of their community members.
 - **Process:** Organise learning and sense-making sessions to discuss the study results and the implications for decision making for households, policy and intervention design. This generally should be done in different ways for different stakeholder groups. Discussions with policy-makers or those designing interventions should also include cocoa farmers or representatives of producer organisations. Additionally, separate discussions should be organised with farmers/farming households to share and discuss the study results. Cooperative or program implementing staff can also integrate the lessons learnt in their activities (e.g. training sessions), or during follow-up data collection moments (e.g. yearly monitoring surveys or Farmer Field Book data collection methods). For learning and sense making purposes, it is best to share data and information as tangibly as possible with farming households (participatory methods) to ensure they can connect with the information and reflect on possible differences between their situation and the situation of other farmers, and what the study results might imply for their situation.

7 Glossary

Adjusted living income benchmark: the relevant geographical living income benchmark adjusted for inflation, household size and composition.

Cocoa farmer: Everyone who cultivates cocoa is recognised as a farmer by the cocoa sector, irrespective of gender or landholding status (Mars Cocoa for Generation Report, 2022).

Cocoa price: most frequently received price per unit of cocoa by the farm household

Cocoa farming or cocoa-producing households: Household of which at least one of the members is a cocoa farmer.

Communal labourers: labour exchange based on reciprocity

Direct supply chain: Cocoa that is produced by individual farmers, producer cooperatives, and organizations and purchased directly by a company at the first purchase point, and the producer(s) identity, farm location(s) and volume is/are known. In cases where intermediaries are involved in purchasing, the above criteria will apply to the intermediary (e.g., traitants, pisteurs, LBC's). A producer(s) identity, farm location, and associated volume must be known and shared with the company who buys the cocoa to be considered 'direct'.

Farm size: (see total household land size)

Household members: a group of adults and children, regardless of their relationship, who contribute to or are dependent on a shared economic pot. This includes people living in the household for at least six months per year and dependent children who live elsewhere (e.g. in school) and excludes paid labourers who live with the household.

Household living income gap: Is the adjusted living income benchmark – net actual household income.

Household size: Total number of household members. (see definition household member)

Indirect supply chain: Cocoa purchased by a company from an intermediary where the producer(s) identity, farm location(s), and volume are unknown / not shared with the purchasing company.

Intervention data: Data collected either by the intervening party and/or from cocoa farming household members, on a household's participation in projects, programmes, trainings etc. that aim directly or indirectly to impact a farming household's living income. Data includes the frequency, degree of participation, quantity of resources provided, intensity and reach of the intervention.

Labour groups: groups of trained professionals that rent out their services (also called, labour gangs, labour brigades and service groups).

Living income: the net annual income required for a household in a particular place to afford a decent standard of living for all members of that household. It covers the costs of food, water, housing, education, healthcare, transport, clothing, and a margin for unexpected events.

Living income benchmark: an estimate of the cost of a basic and decent standard of living for a household at a specified period of time. It answers the question: 'how much does a typical household in a particular place need to earn, from all income sources, to live a decent standard of living?'

Living income gap: the difference between the living income benchmark and current income of a typical farming household.

Net actual household income: all the net income earned by all household members together. It covers net farm income (e.g. income from cash crops, food crops, livestock), net off-farm income (e.g. employment, business) and all other income (e.g. pensions and remittances).

Net cocoa income: total value of cocoa production minus the total costs of cocoa production

Net non-cocoa income: all income from (non-cocoa) income-generating activities (IGAs) performed by all household members regardless if they are done on the farm or outside of the farm.

Other farm income: Sum of net income from cocoa, cash crops, food crops, timber from cocoa farms and livestock products). Off-farm income is the sum of net income from all other IGAs. Other income is the sum of net income from pensions, remittances and cash transfers (recurrent).

Permanent labourers: long-term labourers such as caretakers or farm managers that are paid in cash for their services (in contrast to sharecroppers who receive a share of the production)

Plot: individual piece (parcel) of land, in some contexts called a farm

Sharecropper: A person participating in a sharecropping arrangement, also referred to as tenant.

Sharecropping arrangement: land use arrangement in which a tenant (sharecropper) can cultivate a land owner's land in exchange for a share of the production (or revenue or profit) or a share of the trees (or land).

Sharecropped out: land that is owned by the respondents' household members but cultivated by another household under a sharecropping arrangement

Sharecropped in: land that is owned by another household but cultivated by the respondents' household under a sharecropping arrangement.

Temporary labourers: short-term labourers who are hired for specific tasks or seasons.

Total costs of cocoa production: all labour and material costs for all cocoa-producing activities combined

Total cultivated land (ha): total household land size – uncultivated household land

Total labour costs: SUM(annual costs of: permanent labourers, temporary labourers, labour groups, communal labour) on all land under cocoa cultivation, including sharecropped in and sharecropped out)

Total material costs: SUM(tools/equipment, inputs, transport, land, debt/insurance/tax) on all land under cocoa cultivation, including sharecropped in and sharecropped out

Total production costs of non-cocoa IGA: Sum of costs of rent/lease of land/buildings (e.g. shop), equipment, raw materials and inputs (incl fuel), labour, transport, loan interests, insurance and taxes

Total cultivated land: total area of total household land size that is cultivated (not left fallow)

Total cultivated cocoa land (ha): cultivated land dedicated to cocoa production

Total cultivated cocoa land sharecropped: total area of total cultivated cocoa land that is under a sharecropping arrangement (out or in)

Total household land size: the size of the total land area that ANY household member either (i) owns (with or without ownership title), (ii) has rights to use (possession, assigned communal land, land reform titles, etc), (iii) has any land-use arrangement with third parties (loans, rentals, lease), (iv) has under a sharecropping arrangement (out or in).

Total household land size (ha): all land owned + all land with user rights + all land with temporary rights + all land sharecropped in by the household – all land given out to a sharecropper.

Total labour costs: annual costs of all labour types

Total material costs: annual costs of tools, equipment, inputs, transport and land

Total premium received: volume of production sold as certified/high quality (kg) * premium/kg

Total value of cocoa production: total volume of cocoa produced by the household, accounting for any post-harvest losses, times the price received plus any additional premiums received.

Total volume of cocoa produced: all the cocoa produced on the household's total cultivated cocoa land accounting for post-harvest losses and sharecropping arrangements

Total volume of cocoa produced (kg): (Volume of cocoa from land cultivated by the household (kg) + Volume of cocoa from land sharecropped out (kg) + Volume of cocoa from land sharecropped in (kg)) – post-harvest loss (kg)

Volume of cocoa from own land (not under sharecropping arrangement) (kg): gross volume of cocoa produced (kg) – post-harvest loss (kg)

References

- Anker, R., and Anker, M. (2017). Living wages around the world: Manual for measurement. Edward Elgar Publishing.
- Bymolt, R., Laven, A. and Tyszler, M. (2018) Demystifying the Cocoa Sector in Ghana and Côte d'Ivoire. KIT Royal Tropical Institute. <https://kit2018.wpenGINEpowered.com/wpcontent/uploads/2020/05/Demystifying-complete-file.pdf>
- de los Rios, C. and Tyszler, M. (2020) Guidance on calculating household income. The Committee on Sustainability Assessment (COSA) and KIT Royal Tropical Institute, Prepared for the Living Income Community of Practice.
- Diarra, I., and N'Guessan, Y. (2018). Rural Côte d'Ivoire, Cocoa growing areas of Sud Comoe, Indenie, Djublin, Me, AgnebyTiassa, Goh, Loh Djiboua, Nawa, San Pedro, and Tonkpi. Living Income Community of Practice, Series 2.
- Fountain, A., and Hütz-Adams, F. (2015). 2015 Cocoa Barometer (USA edition).
- Fountain, A. and F. Hütz-Adams (2022). 2022 Cocoa Barometer The Voice Network.
- Ingram, V., Waarts, Y, Ge, L., van Vugt, S., Wegner, L., Puister-Jansen, L., Ruf, F., and Tanoh, R. (2014). 'Impact of UTZ certification of cocoa in Ivory Coast.' Assessment framework and baseline.
- Ingram, V., van Rijn, F., Waarts, Y., and Gilhuis, H. (2018). The Impacts of Cocoa Sustainability Initiatives in West Africa. Sustainability 10(11): 4249.
- IMF (2023) Prices, Production and Labor Selected Indicators. <https://data.imf.org/regular.aspx?key=61545849>
- Khandker, S.R., Koolwal, G.B., and Samad, H.A. 2010. Handbook on impact evaluation: quantitative methods and practices. World Bank Publications, 2010.
- LICOP (2023). The Concept, <https://www.living-income.com/the-concept>
- Stern, E., Stame, N., Mayne, J., et al., 2012a. Broadening the range of designs and methods for impact evaluations. Working Paper 38. London: Department for International Development
- DCED. 2013 The DCED standard for measuring achievements in private sector development control points and compliance criteria. Donor Committee Enterprise Development 19.
- Living Income Community of Practice (2020) Living Income Benchmark, August 2020 Update, Côte d'Ivoire, Rural cocoa growing areas. Anker Research Institute. https://c69aa8ac-6965-42b2-abb7-0f0b86c23d2e.filesusr.com/ugd/0c5ab3_a3830a35e23c4e77adc642fba454a93e.pdf
- Living Income Community of Practice (2022) Living Income Benchmark, June 2022 Update, Côte d'Ivoire, Rural cocoa growing areas. Anker Research Institute. https://www.living-income.com/files/ugd/0c5ab3_9aef39b2ef654ab6a8f7bc4dd2bdb026.pdf
- Mars Cocoa for Generations Report, 2022, https://www.mars.com/sites/g/files/dfsbuf106/files/2023-10/Cocoa%20for%20Generations%20-%202022%20Annual%20Report_0.pdf
- OECD (n.d.) Adjusting Household Incomes: equivalence scales. OECD. <https://www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf>
- Romo (2022) Income measurement practitioner's guide. A practitioner's guide and framework for selecting an appropriate approach to income measurement, ISEAL Alliance, Prepared for the Living Income Community of Practice.
- Schneider, K. (2021) Evidence in Development Policymaking and Practice. Policy Brief. Nadel. Center for Development and Cooperation. ETH Zürich.
- Tyszler, M., Bymolt, R., and Laven, A. (2018). Analysis of the income gap of cocoa producing households in Côte d'Ivoire. KIT Royal Tropical Institute, Amsterdam, The Netherlands.
- Tyszler, M., and de los Rios, C. (2020) Guidance Manual on Calculating and Visualizing Income Gap to a Living Income Benchmark. Prepared for the Living Income Community of Practice.
- van Vliet, J.A., Slingerland, M.A., Waarts, Y.R., and Giller, K.E. (2021). A Living Income for Cocoa Producers in Côte d'Ivoire and Ghana? Front. Sustainable Food Systems 5:732831. doi: 10.3389/fsufs.2021.732831

Appendix 1 Survey

This section presents the structure of the living income survey aimed at capturing the necessary data for constructing the indicators listed in the proposed methodology. Throughout the section, we highlight in *italic* the instructions given to enumerators for asking the questions, explaining the concepts and/or probing the questions to the respondent(s). When a question refers to a previously captured variable, we indicate the latter with the following sign `{variable_name}`. No constraints on variable's values or built-in checks are mentioned as these will be coded in the full survey developed digitally in the subsequent phases of the Cocoa Household Income Study.

A1.1 General sections

The following sections should be used for both minimum requirements and the recommended approach.

A1.1.1 Respondent selection

Add screening question based on the study objective and population of interest.

Some examples:

1. Is the household registered at cooperative X?
2. Has the household participated in intervention Y?
3. Does the household sharecrop in land?
4. Is there a female cocoa farmer in the household?
5. Is this person available for an interview?
6. Who in the household is most knowledgeable about the cocoa production process?
7. Who in the household is most knowledgeable about the other income generating activities?
8. Is this person available for an interview?

A1.1.2 General information

1. Date
2. Region
3. Town
4. Enumerator name
5. Gps point

A1.1.3 Respondent characteristics

1. Name of respondent
2. Gender of respondent
3. Position in the household (household head, spouse, son, daughter, etc.)
4. Other personal/household characteristics of interest not covered in the minimum approach

A1.1.4 Informed consent

Hello. My name is XXX and I am working with XXX and XXX.

We are conducting a survey on behalf of XXX, which will be and looking into activities of people in this area working in the cocoa sector. We hope that you will participate in this survey since your views are important in informing relevant policies and interventions to enhance smallholder farmer livelihoods.

We will report all information in an anonymous way, and this information cannot be traced back to individual people. The survey takes about 1,5 hours to complete.

You can choose whether or not to participate in the survey and stop at any time. Your personal data will only be used by the researchers to fill a dataset used for analyses, your responses will be stored anonymously, and reporting will only be done by grouping the responses from many different respondents like you, so

that nobody will be able to trace back your responses to you. We need your personal data since we want to interview you again at a later stage of the project to see how you are doing, and want to compare outcomes of the first and following surveys. You have the right to withdraw the given consent at any time. The personal data you provide will only be used for the purpose for which you provided it and will not be shared with others.

By giving your consent, you declare that you have provided this personal data voluntarily. If we should come to any question you do not want to answer, just let me know and I will go to the next question. You can also stop the interview at any time.

We hope that you will participate in this survey. At this time, have any questions to me about the survey? `

May I begin the interview now?

A1.2 Minimum requirement

A1.2.1 Household size

Enumerator: We would like to ask you a few questions about the people in your household. Household members are the group of adults and children, regardless of their relationship, who contribute to or are dependent on a shared economic pot. Please only include people who live in the household for at least six months per year and dependent children who live elsewhere (e.g. in boarding school) and exclude paid labourers who live with the household.

1. What is the total number of household members in your household? (save $\${hh_size}$)
2. What is the total number of adults members of your household? (18 years and older), including yourself (save $\${hh_adults}$)
3. How many of these $\${hh_adults}$ are men?
4. How many of these $\${hh_adults}$ are women?
5. What is the total number of children aged 14 to 17 in your household?
6. What is the total number of children aged 13 or under in your household?

A1.2.2 Land size

Enumerator: 'We would like to ask you some questions about your land. Total household land size is the total land area that ANY household member either (i) owns (with or without ownership title), (ii) has rights to use (possession, assigned communal land, land reform titles, etc), (iii) has any land-use arrangement with third parties (loans, rentals, lease), (iv) has under a sharecropping arrangement (out or in).'

7. Which area unit do you know best? (1. hectare(s); 2. acre(s); 3. square(s) (0.25ha); 4. Pole; 5. Other) (save as $\${land_unit}$)
 - a. If the respondent indicates Other, please specify the name of the unit (text) and how much of this unit is contained in 1 hectare (integer).
8. During the last production cycle, how many plots (farms) did your households' land consist of? (integer)
9. During the last production cycle, how many $\${land_unit}$ in total did your household have access to? (irrespective of land use) (save $\${land_total}$)
10. During the last production cycle, on how many $\${land_unit}$ did your household cultivated any crop? (save $\${land_cultivated}$)
11. During the last production cycle, on how many $\${land_unit}$ did your household cultivated cocoa? (save $\${land_cocoa}$)
12. During the last production cycle, how much of the $\${land_cocoa}$ did your household own or have user rights over (e.g. possession, assigned communal land, land reform titles, etc)?
13. During the last production cycle, how much of the $\${land_cocoa}$ did you household loan, rent of lease from third parties?
14. During the last production cycle, how much of the $\${land_cocoa}$ did your household sharecrop out? (*Sharecropped out refers to the land that is owned by the respondents' household members but cultivated by another household under a sharecropping arrangement*)
15. During the last production cycle, how much of the $\${land_cocoa}$ did your household sharecrop in? (*Sharecropped in refers to the land that is owned by another household but cultivated by the respondents' household under a sharecropping arrangement*)
16. Has your cocoa land been GPS measured by someone? (Y/N)
 - a. If yes, how many $\${land_unit}$ is your measured cocoa land?
 - b. If yes, which organization measured your households' cocoa land?

A1.2.3 Cocoa production volume

Enumerator: 'We would like to ask you some questions about your households' cocoa production during the last production cycle. The production cycle covers both the main and the light season.' Specify the period (months) of the main and light season.

17. Do you keep records (such as Farmer Field Book or passbook) of for instance the volume produced and sold, and costs of your cocoa production? (Y/N)
18. If yes, how often did you record the information ?

Enumerator: If there are such records, ask the farmer if they can get it and capture:

19. Records used to answer questions in this section (Y/N)
20. Which unit do you use to measure cocoa production? (1. bag(s); 2. kg(s); 3. tonne(s); 4. other) (save $\${prod_unit}$)
 - a. If the respondent indicates Other, please specify the name of the unit (text). For both Bag(s) and Other Units, indicate how many kg(s) are contained in each unit (integer).
21. During the last main season, how much $\${prod_unit}$ of cocoa did your household produced? (save $\${prod_main}$)
22. How much of the $\${prod_main}$ was lost post-harvest (e.g. due to poor quality, rot, theft, buyer's rejection)?
23. During the last light season, how much $\${prod_unit}$ of cocoa did your household produced? ($\${prod_light}$)
24. How much of the $\${prod_light}$ was lost post-harvest (e.g. due to poor quality, rot, theft, buyer's rejection)?
25. During the last main season, what was the most frequently obtained price for MOST of your cocoa excluding premiums? (local currency)
26. Per which unit did you receive that price? (1. bag; 2. kg; 3. tonne; 4. total; 5.other)
27. During the last main season, for how much of the $\${prod_main}$ did you receive a premium?
28. During the last main season, what was the most frequently obtained premium for MOST of your cocoa? (local currency)

29. Per which unit did you receive that premium? (1. bag; 2. kg; 3. tonne; 4. total; 5.other)
30. During the last light season, what was the most frequently obtained price for MOST of your cocoa excluding premiums? (local currency)
31. Per which unit did you receive that price? (1. bag; 2. kg; 3. tonne; 4. total; 5.other)
32. During the last light season, for how much of the $\${prod_light}$ did you receive a premium?
33. During the last light season, what was the most frequently obtained premium for MOST of your cocoa? (local currency)
34. Per which unit did you receive that premium? (1. bag; 2. kg; 3. tonne; 4. total; 5.other)

A1.2.4 Cocoa production costs

Enumerator: 'We would now like to ask you about your household's cocoa production costs for labour as well as materials.'

- a. General questions
 35. During the last production cycle, how much $\${local\ currency}$ did your household spend on loans/rent/lease of land? – (*Only for if the household loans/rents/leases land*)
 36. During the last production cycle, who worked on your cocoa farm (for any cocoa production activities)? (1. household members; 2. sharecroppers; 3. permanent labourers; 4. temporary labourers; 5. communal labourers; 6. labour groups) – save $\${labourer_type}$
 37. During the last production cycle, how many permanent labourers have worked on your cocoa land? (*only permanent labourers*)
 38. How much $\${local\ currency}$ have you paid these permanent labourers in TOTAL? (decimal) - (*only permanent labourers*)
 39. Per which time period? (1. Week ; 2. Month ; 3. Year)
- b. Questions per production activity
Ask separately for each (higher order) cocoa production activity: 1. Preparation (Land clearing and planting); 2. Maintenance (weeding, pruning, applying phytosanitary products); 3. Harvesting; 4. Post-harvest activities (pod breaking, fermentation, drying, transportation):

40. During the last production cycle, was activity carried out by your household during the cocoa production? (Y/N) – (save $\${activity}$)
41. If yes, which $\${labourer_type}$ carried out the selected $\${activity}$? (save $\${laborer}$)
42. How many temporary labourers worked on $\${activity}$ during the last production cycle? (*Only for temporary labourers*)
43. How much in $\${local\ currency}$ were temporary labourers paid for $\${activity}$ during the last production cycle? (decimal) - (*Only for temporary labourers*)
44. Per which time period? (1. Time(s); 2. Hour(s); 3. Days(s); 4. Week(s); 5. Month(s)) - (*Only for temporary labourers*)
45. How many labour groups worked on $\${activity}$ during the last production cycle? (*Only for labour groups*)
46. How much in $\${local\ currency}$ were labour groups paid for $\${activity}$ during the last production cycle? (decimal) - (*Only for labour groups*)
47. Per which time period? (1. Time(s); 2. Hour(s); 3. Days(s); 4. Week(s); 5. Month(s)) - (*Only for labour groups*)
48. How much $\${local\ currency}$ did your household spend on material costs for $\${activity}$ during the last production cycle? (specify type of material costs per production activity (e.g. phytosanitary products, spraying equipment, protective gears, etc.))

A1.2.5 Non-cocoa income

Enumerator: 'We would like to ask you about your household's different sources of income. Please think about ANY income-generating activity (IGA) done by ANY person in your household, including yourself and income from remittances, pension or cash transfers.'

49. During the last production cycle, in addition to ****COCOA****, did anyone (ANY PERSON) in your household earned/received money from...
(*Enumerator: Read the list*)
(*Select multiple: 1. Selling other cash crops (rubber, coffee, palm, etc.) or food crops (Yam, cassava, plantains, etc.); 2. Selling timber (including shade trees from cocoa farms); 3. Selling livestock or livestock products; 4. Selling fish; 5. Selling forest and bushmeat products; 6. Selling/Renting land; 7. Processing of agri-food; 8. Trade or shop; 9. Other small*

- businesses (e.g. hairdressing, money transfer); 10. Working on other people's cocoa plantations; 11. Working on other people's plantations for other crops (not cocoa); 12. Working for other people in another way (non-agricultural); 13. Employment in a company; 14. Employment in administration or the public sector; 15. Money received from family members living elsewhere (remittances); 16. Pension or other social security program; 17. Cash-transfers (from NGOs, public or private programs); 16. Other; 17. None)
 - a. If you answered Other, please specify the name of this other IGA:
(text)
(save $\${list_income_sources}$)
50. During the last production cycle, which PERCENTAGE of your TOTAL NET HOUSEHOLD income was coming from each of the selected item on $\${list_income_sources}$ (including cocoa)?

(Enumerator: Read the list of selected income sources. Use 20 beads/seeds to visualise the percentages and ask how many to assign to each income source. If the calculated total does not account to 100 percent, please ask the respondent to adjust the allocated percentages).

A1.2.6 Intervention data: training, capacity building and other interventions

51. Did you, or anybody else in your household, participate in any projects, training programmes or other activities organised by for example the cooperative, cocoa companies, NGOs or the government? If yes, please explain what kind of activity or activities you participated in.
 - a. Provision of equipment
 - b. Farm development plans
 - c. Supply of fertiliser and/or pesticides
 - d. Supply of planting material, including tree nurseries
 - e. GAP training, coaching, Farmer Field Schools, demonstration plots
 - f. Organised hired labour, including spraying gangs
 - g. Farm management support/training
 - h. Processing of farm produce, e.g. cocoa juice
 - i. Production and yield interventions on other crops produced by the household

- j. Farm diversification (e.g. food production, livestock, agroforestry)
 - k. Payments for Environmental Services, including carbon credits
 - l. Off-farm employment - own businesses
 - m. Off-farm employment – employment by others
 - n. Community development, including child labour remediation
 - o. Child labour remediation
 - p. Other (specify...)
52. For each of the farm and income diversification activities participated in by the household (51.h – 51.m): what type of support did you receive? (multiple answers possible)
- a. Farm/business management support/training
 - b. Other training
 - c. Provision of inputs/material
 - d. Market access support
 - e. Technical support

A1.2.7 Intervention data: direct activities on prices and incomes

53. On top of the normal cocoa price received, did your household receive from any other financial benefits linked to cocoa production? (name examples from list below). If yes, which?
- a. Farm gate prices
 - b. Premiums and bonuses, including the Living Income Reference Price
 - c. Advance purchase payments
 - d. Cash transfers
54. You indicated that your household participated in $\{price\}$. Do you know by whom this was organised? (multiple options are possible)
- a. *Don't know*
 - b. Actor A
 - c. Actor B
 - d. Actor C, etc. To be added depending on actors working in location of study.
 - e. Other, specify...
55. For the $\{price\}$ organised by $\{actor\}$ could you indicate how much extra money you think your household earned over the past year because of this? (in $\{local\}$)

A1.2.8 Intervention data: support on access to financial services

56. Did you or anyone in your household receive help in accessing financial services? If yes, which?
- a. Credit, including inputs on credit
 - b. VSLAs
 - c. Insurance
 - d. Savings
 - e. Banking

A1.3 Recommended approach

A1.3.1 Household size

Enumerator: 'We would like to ask you a few questions about the people in your household. Household members are the group of adults and children, regardless of their relationship, who contribute to or are dependent on a shared economic pot. Please only include people who live in the household for at least six months per year and dependent children who live elsewhere (e.g. in boarding school) and exclude paid labourers who live with the household.'

1. What is the total number of household members in your household? (save $\{hh_size\}$)

Enumerator: 'Could you indicate the name, gender, age and ... for each of the $\{hh_size\}$ member(s) (including yourself) please?' Start with the head of household then the respondent if he or she is not the head of household:

For each household member ask:

2. What is the name of the household member?
3. What is the sex of $\{name\}$? (M/F)
4. What is the age of $\{name\}$? (number)
5. Has $\{name\}$ lived in your house for at least 6 months in the past year? (Y/N)
6. What is the marital status of $\{name\}$? (*only for adults*)

7. What is the highest level of education attained by $\${name}$? (use locally relevant education levels (e.g. primary school, middle school, high school, higher education) - *(only for adults)*)
8. Has $\${name}$ actively contributed to the household income in the past year? (Y/N) - *(only for adults)*
9. Is $\${name}$ currently enrolled in school? (Y/N) - *(only for children)* (save $\${list_household_members}$)

A1.3.2 Land size

Enumerator: 'We would like to ask you some questions about your land. Total household land size is the total land area that ANY household member either (i) owns (with or without ownership title), (ii) has rights to use (possession, assigned communal land, land reform titles, etc), (iii) has any land-use arrangement with third parties (loans, rentals, lease), (iv) has under a sharecropping arrangement (out or in).'

10. Which area unit do you know best? (1. hectare(s); 2. acre(s); 3. square(s) (0.25ha); 4. Pole; 5. Other) (save as $\${land_unit}$)
 - a. If the respondent indicates Other, please specify the name of the unit (text) and how much of this unit is contained in 1 hectare (integer).
11. During the last production cycle, how many plots (farms) did your households' land consist of? (integer)
12. During the last production cycle, how many $\${land_unit}$ in total did your household have access to? (irrespective of land use) (save $\${land_total}$)
13. During the last production cycle, on how many $\${land_unit}$ did your household cultivated any crop? (save $\${land_cultivated}$)
14. During the last production cycle, on how many $\${land_unit}$ did your household cultivated cocoa? (save $\${land_cocoa}$)
15. During the last production cycle, how much of the $\${land_cocoa}$ did your household own or have user rights over (e.g. possession, assigned communal land, land reform titles, etc)?
16. During the last production cycle, how much of the $\${land_cocoa}$ did you household loan, rent of lease from third parties?
17. During the last production cycle, how much of the $\${land_cocoa}$ did your household sharecrop out? (*Sharecropped out refers to the land that is*

owned by the respondents' household members but cultivated by another household under a sharecropping arrangement)

18. During the last production cycle, how much of the $\${land_cocoa}$ did your household sharecrop in? (*Sharecropped in refers to the land that is owned by another household but cultivated by the respondents' household under a sharecropping arrangement)*
19. Has your cocoa land been GPS measured by someone? (Y/N)
 - a. If yes, how many $\${land_unit}$ is your measured cocoa land?
 - b. If yes, which organization measured your households' cocoa land?
20. During the last production cycle, on how many $\${land_unit}$ did your household cultivated other cash crops?
21. During the last production cycle, on how many $\${land_unit}$ did your household cultivated food crops?
22. During the last production cycle, on how many $\${land_unit}$ did your household use for other farm activities (e.g. livestock, timber trees)?
23. During the last production cycle, how much of the $\${land_cocoa}$ contained cocoa trees younger than 5 years old? (young trees not yet productive)
24. During the last production cycle, how much of the $\${land_cocoa}$ contained cocoa trees 5-25 years old? (mature and fully productive trees)
25. During the last production cycle, how much of the $\${land_cocoa}$ contained cocoa trees older than 25 years old? (older trees with declining productivity)

A1.3.3 Cocoa production volume

Enumerator: 'We would like to ask you some questions about your households' cocoa production during the last production cycle. The production cycle covers both the main and the light season.' Specify the period (months) of the main and light season.

26. Do you keep records (such as Farmer Field Book or passbook) of for instance the volume produced and sold, and costs of your cocoa production? (Y/N)
27. If yes, how often did you record the information ?

Enumerator: If there are such records, ask the farmer if they can get it and capture:

28. Records used to answer questions in this section (Y/N)
29. Which unit do you use to measure cocoa production? (1. bag(s); 2. kg(s); 3. tonne(s); 4. other) (save $\{\text{prod_unit}\}$)
 - a. If the respondent indicates Other, please specify the name of the unit (text). For both Bag(s) and Other Units, indicate how many kg(s) are contained in each unit (integer).
30. During the last main season, how much $\{\text{prod_unit}\}$ of cocoa did your household produced? (save $\{\text{prod_main}\}$)
31. How much of the $\{\text{prod_main}\}$ was lost post-harvest (e.g. due to poor quality, rot, theft, buyer's rejection)?
32. During the last light season, how much $\{\text{prod_unit}\}$ of cocoa did your household produced? ($\{\text{prod_light}\}$)
33. How much of the $\{\text{prod_light}\}$ was lost post-harvest (e.g. due to poor quality, rot, theft, buyer's rejection)?
34. During the last main season, what was the most frequently obtained price for MOST of your cocoa excluding premiums? (local currency)
35. Per which unit did you receive that price? (1. bag; 2. kg; 3. tonne; 4. total; 5.other)
36. During the last main season, for how much of the $\{\text{prod_main}\}$ did you receive a premium?
37. During the last main season, what was the most frequently obtained premium for MOST of your cocoa? (local currency)
38. Per which unit did you receive that premium? (1. bag; 2. kg; 3. tonne; 4. total; 5.other)
39. During the last light season, what was the most frequently obtained price for MOST of your cocoa excluding premiums? (local currency)
40. Per which unit did you receive that price? (1. bag; 2. kg; 3. tonne; 4. total; 5.other)
41. During the last light season, for how much of the $\{\text{prod_light}\}$ did you receive a premium?
42. During the last light season, what was the most frequently obtained premium for MOST of your cocoa? (local currency)
43. Per which unit did you receive that premium? (1. bag; 2. kg; 3. tonne; 4. total; 5.other)

For land that your household sharecrops out to another household:

44. During the last main season how much $\{\text{prod_unit}\}$ of cocoa was produced on this land? (save $\{\text{prod_sc_out_main}\}$)
45. How much of the $\{\text{prod_sc_out_main}\}$ was lost post-harvest (e.g. due to poor quality, rot, theft, buyer's rejection)?
46. During the last light season, how much $\{\text{prod_unit}\}$ of cocoa was produced on this land? (save $\{\text{prod_sc_out_light}\}$)
47. How much of the $\{\text{prod_sc_out_light}\}$ was lost post-harvest (e.g. due to poor quality, rot, theft, buyer's rejection)?
48. How was/were the sharecropper(s) compensated for their work? (1/2 of cocoa volume; 1/3 of cocoa volume; 1/2 of cocoa revenue; 1/3 of cocoa revenues; 1/2 of cocoa profit; 1/3 of cocoa profit; 1/2 of trees (or land); 1/3 of trees (or land); other)

For land that your household sharecrops in from another household:

49. During the last main season, how much $\{\text{prod_unit}\}$ of cocoa was produced on this land? (save $\{\text{prod_sc_in_main}\}$)
50. How much of the $\{\text{prod_sc_in_main}\}$ was lost post-harvest (e.g. due to poor quality, rot, theft, buyer's rejection)?
51. During the last light season, how much $\{\text{prod_unit}\}$ of cocoa was produced on this land? (save $\{\text{prod_sc_in_light}\}$)
52. How much of the $\{\text{prod_sc_in_light}\}$ was lost post-harvest (e.g. due to poor quality, rot, theft, buyer's rejection)?
53. How was your household compensated for its work? (1/2 of cocoa volume; 1/3 of cocoa volume; 1/2 of cocoa revenue; 1/3 of cocoa revenues; 1/2 of cocoa profit; 1/3 of cocoa profit; 1/2 of trees (or land); 1/3 of trees (or land); other)

A1.3.4 Cocoa production costs

Enumerator: We would now like to ask you about your household's cocoa production costs for labour as well as materials.

- a. General questions
 54. During the last production cycle, how much $\{\text{local currency}\}$ did your household spend on loans/rent/lease of land? – (*Only for if the household loans/rents/leases land*)

55. During the last production cycle, who worked on your cocoa farm (for any cocoa production activities)? (1. household members; 2. sharecroppers; 3. permanent labourers; 4. temporary labourers; 5. communal labourers; 6. labour groups) – save $\{\text{labourer_type}\}$
56. During the last production cycle, how many permanent labourers have worked on your cocoa land? (*only permanent labourers*)
57. How much $\{\text{local currency}\}$ have you paid these permanent labourers in TOTAL? (decimal) - (*only permanent labourers*)
58. Per which time period? (1. Week ; 2. Month ; 3. Year)

During the last cycle, how many months did these permanent labourers work on your cocoa land? (integer) - (*Only for permanent labourers*)

b. Additional questions per production activity

Ask for each individual cocoa-producing activity: 1. Land clearing; 2. Planting; 3. Fertiliser application (granular, liquid and organic); 4. Herbicide application; 5. Pesticide/insecticide application; 6. Fungicide application; 7. Weeding (multiple rounds); 8. Pruning; 9. Harvesting; 10. Pod breaking; 11. Fermentation; 12. Drying; 13. Transportation (to drying location, to selling location)

59. During the last production cycle, was activity carried out by your household during the cocoa production? (Y/N) – (save $\{\text{activity}\}$)
60. If yes, which $\{\text{labourer_type}\}$ carried out the selected $\{\text{activity}\}$? (save $\{\text{laborer}\}$)

For each $\{\text{labourer_type}\}$ selected and $\{\text{activity}\}$ carried out, ask:

61. How many $\{\text{labourer_type}\}$ worked on $\{\text{activity}\}$ during the last production cycle?
62. How long (or how many times) did $\{\text{labourer_type}\}$ work on $\{\text{activity}\}$ during the last production cycle?
63. Time period (1. Time(s); 2. Hour(s); 3. Days(s); 4. Week(s); 5. Month(s))
64. How much in $\{\text{local currency}\}$ were temporary labourers paid for $\{\text{activity}\}$ during the last production cycle? (decimal) - (*Only for temporary labourers*)
65. Per which time period? (1. Time(s); 2. Hour(s); 3. Days(s); 4. Week(s); 5. Month(s)) - (*Only for temporary labourers*)
66. How much in $\{\text{local currency}\}$ were labour groups paid for $\{\text{activity}\}$ during the last production cycle? (decimal) - (*Only for labour groups*)

67. Per which time period? (1. Time(s); 2. Hour(s); 3. Days(s); 4. Week(s); 5. Month(s)) - (*Only for labour groups*)
68. How much in $\{\text{local currency}\}$ did your household spend on food and other in-kind costs for communal labourers or labour groups working on $\{\text{activity}\}$ during the last production cycle? (*Only for communal labourers and/or labour groups*)
69. How much $\{\text{local currency}\}$ did your household spend on material costs for $\{\text{activity}\}$ during the last production cycle? (specify type of material costs per production activity (e.g. phytosanitary products, spraying equipment, protective gears, etc.))

A1.3.5 Non-cocoa income

Enumerator: We would like to ask you about your household's different sources of income. Please think about ANY income-generating activity (IGA) done by ANY person in your household, including yourself and income from remittances, pension or cash transfers.

70. During the last production cycle, in addition to **COCOA**, did anyone (ANY PERSON) in your household earned/received money from... (*Enumerator: Read the list*)

(Select multiple: 1. Selling other cash crops (rubber, coffee, palm, etc.) or food crops (Yam, cassava, plantains, etc.); 2. Selling timber (including shade trees from cocoa farms); 3. Selling livestock or livestock products; 4. Selling fish; 5. Selling forest and bushmeat products; 6. Selling/Renting land; 7. Processing of agri-food; 8. Trade or shop; 9. Other small businesses (e.g. hairdressing, money transfer); 10. Working on other people's cocoa plantations; 11. Working on other people's plantations for other crops (not cocoa); 12. Working for other people in another way (non-agricultural); 13. Employment in a company; 14. Employment in administration or the public sector; 15. Money received from family members living elsewhere (remittances); 16. Pension or other social security program; 17. Cash-transfers (from NGOs, public or private programs); 16. Other; 17. None)

- a. If you answered Other, please specify the name of this other IGA:
(text)
(save $\${list_income_sources}$)

If '1. Selling other cash crops or food crops' is selected, ask:

71. In addition to **COCOA**, which crops did your household PRODUCE during the last cycle?
 72. In addition to **COCOA**, which crops did your household SELL during the last cycle?
 73. During the last cycle, apart from **COCOA**, from which crop did you earn the most income in TOTAL? (save as $\${non_cocoa_crop1}$)
 74. During the last cycle, apart from **COCOA**, from which second crop did you earn the most income in TOTAL? (save as $\${non_cocoa_crop2}$)
 (replace '1. Selling other cash crops or food crops' in $\${list_income_sources}$ with $\${non_cocoa_crop1}$, $\${non_cocoa_crop2}$ and, $\${non_cocoa_crop_other}$)

For each of the item selected on $\${list_income_sources}$ (excluding cocoa), ask:

75. During the last cycle, who in the $\${list_household_members}$ was in charge of the selected income source? (save $\${name}$)
 76. How many months per year did $\${name}$ work on this income source during the last cycle?
 77. How many weeks per month did $\${name}$ work on this income source during the last cycle?
 78. How many days per week did $\${name}$ work on this income source during the last cycle?
 79. During the last production cycle, what was your household's total income from this income source in $\${local_currency}$?
 80. For which time period? (Per year, month, week, day)
 81. During the last production cycle, what were your household's total costs for this income source in $\${local_currency}$? (*Enumerator define total costs: costs of rent/lease of land/buildings (e.g. shop), equipment, raw materials and inputs (incl fuel), labour and transport*)
 82. For which time period? (Per year, month, week, day)

For selected remittances, pensions and other cash-transfers, ask the following to capture predictability and stability:

83. What is the TOTAL amount (in $\${local_currency}$) of remittances/pensions/cash transfers received during the last 12 months by household members?
 84. How often did you or another household members receive remittances/pensions/cash transfers in the past 12 months?

A1.3.6 Intervention data: training, capacity building and other interventions

85. Did you, or anybody else in your household, participate in any projects, training programmes or other activities organised by for example the cooperative, cocoa companies, NGOs or the government? If yes, please explain what kind of activity or activities you participated in.
- Provision of equipment
 - Farm development plans
 - Supply of fertiliser and/or pesticides
 - Supply of planting material, including tree nurseries
 - GAP training, coaching, Farmer Field Schools, demonstration plots
 - Organised hired labour, including. spraying gangs
 - Farm management support/training
 - Processing of farm produce, e.g. cocoa juice
 - Production and yield interventions on other crops produced by the household
 - Farm diversification (e.g. food production, livestock, agroforestry)
 - Payments for Environmental Services, including carbon credits
 - Off-farm employment - own businesses
 - Off-farm employment – employment by others
 - Community development, including child labour remediation
 - Child labour remediation
 - Other (specify...)
86. For each of the farm and income diversification activities participated in by the household (71.h – 71.m): what type of support did you receive? (multiple answers possible)
- Farm/business management support/training
 - Other training
 - Provision of inputs/material
 - Market access support

- e. Technical support
87. You indicated that your household participated in $\${intervention}$. Do you know by whom this was organised? (multiple options are possible)
- a. *Don't know*
 - b. Actor A
 - c. Actor B
 - d. Actor C, etc. To be added depending on actors working in location of study.
 - e. Other, specify...

Enumerator: For the $\${intervention}$ organised by $\${actor}$ could you indicate the following?

88. Over the last year, how many times did you participate in this activity?
- a. 1-2 times over the last year
 - b. 2-5 times over the last year
 - c. 5-12 times (e.g. once per month) over the last year
 - d. more than 12 times but less than 50 times over the last year (e.g. biweekly)
 - e. more than 50 times over the last year (e.g. weekly)
89. How many people were participating along with you at the same time?
- a. with a very large group of people > 40
 - b. with a large group of people (max 40)
 - c. with a medium group of people (max 15)
 - d. with a small group of people (max 5)
 - e. one-one-one
90. Do you think that participating in the activity will increase the total amount of money your household earns, after subtracting any costs you have related to agricultural production or to run your business(es)?
- a. I expect that it will decrease the earnings of my household a lot
 - b. I expect that it will decrease the earnings of my household a little
 - c. I expect that the earnings of my household will remain the same
 - d. I expect that it will increase the earnings of my household a little
 - e. I expect that it will increase the earnings of my household a lot

A1.3.7 Intervention data: direct activities on prices and incomes

91. On top of the normal cocoa price received, did your household receive from any other financial benefits linked to cocoa production? (name examples from list below). If yes, which?
- a. Farm gate prices
 - b. Premiums and bonuses, including the Living Income Reference Price
 - c. Advance purchase payments
 - d. Cash transfers
92. You indicated that your household participated in $\${price\ intervention}$. Do you know by whom this was organised? (multiple options are possible)
- a. *Don't know*
 - b. Actor A
 - c. Actor B
 - d. Actor C, etc. To be added depending on actors working in location of study.
 - e. Other, specify...
93. For the $\${price\ intervention}$ organised by $\${actor}$ could you indicate how much extra money you think your household earned over the past year because of this? (in $\${local\ currency}$)

A1.3.8 Intervention data: support on access to financial services

94. Did you or anyone in your household receive help in accessing financial services? If yes, which?
- a. Credit, including inputs on credit
 - b. VSLAs
 - c. Insurance
 - d. Savings
 - e. Banking
95. You indicated that your household received help in accessing $\${financial\ intervention}$. Do you know by whom this was organised? (multiple options are possible)
- a. *Don't know*
 - b. Actor A
 - c. Actor B
 - d. Actor C, etc. To be added depending on actors working in location of study.

- e. Other, specify...
96. For the help with \${financial intervention} organised by \${actor} could you indicate whether you think that participating in the activity will increase the total amount of money your household has available in 5 years?
- I expect that it will decrease the money available to my household a lot
 - I expect that it will decrease the money available to my household a little
 - I expect that the money available to my household will remain the same
 - I expect that it will increase the money available to my household a little
 - I expect that it will increase the money available to my household a lot

A1.3.9 Farmer perceptions on drivers of income change

97. Did your household income in the last cocoa season (specify season xx/xx e.g. 23/24) starting in <month year> change compared to the cocoa season (xx/xx) that started in <month year>? Decrease
- Equal/no change
 - Increase
 - Not applicable (no comparison between seasons possible)
 - I do not know.
98. If you noticed a decrease, what do you see as the main reason or reasons? Multiple responses possible. If there is more than one reason, please list the reasons in order of importance.
- Less inputs available
 - Less cash available for investment in income generating activities
 - Less credit available
 - Less time available by the household to work
 - Less hired labourers available
 - Lower price
 - Lower cocoa volumes produced

- Higher cost of production (inputs, hired labour)
 - More pest and disease pressure
 - Adverse climatic conditions
 - Fewer income opportunities available
 - Less money spent on income generating activities because I spent it on other things
 - Less land available
 - Other (please specify)
 - I do not know
99. If you noticed an increase, what do you see as the main reason or reasons? Multiple responses possible. If there is more than one reason, please list the reasons in order of importance
- More inputs available
 - More cash available for investment in income generating activities
 - More credit available
 - More time available by the household for work
 - More hired labourers available
 - Higher price
 - Higher cocoa volumes produced
 - Lower cost of production (inputs, hired labour)
 - Less pest and disease pressure
 - Better climatic conditions
 - More income opportunities available
 - More money spent on income generating activities because I did not spend it on other things
 - More land available
 - Other (please specify)
 - I do not know
100. You mentioned as reason <xxxx> for the increase in income. Can you explain why you mentioned that reason and how that increase came about? (Open question)

Appendix 2 Conducting regression analyses

A2.1 Comparing between groups while controlling for other factors

Regression analyses can be used for multiple reasons, ranging from analysing the relationship between different indicators to measuring impact. One of them is to compare between groups, while controlling for other factors. When comparing between groups, it involves estimating a formula with the following structure:

$$Y = \beta_0 + \beta_1 FEMALE + \beta_2 X_2 + \epsilon$$

Where:

- Y_0 is the outcome variable of interest;
- *FEMALE* is a variable that is equal to 1 for female-headed households and 0 otherwise.
- X_2 is another variable, such as cocoa farm size. Multiple other variables can be added at the same time.
- β_0 , β_1 and β_2 are coefficients to be estimated by the model.

The β_1 coefficient is the most important coefficient: it gives the difference between male and female-headed households, holding all other variables (land holdings in this case) constant. If the p-value for the *FEMALE* variable would be lower than the significance level (e.g. 5%) in the example above, and the coefficient would be positive, it would indicate that after correcting for everything else, female-headed households indeed have a higher living income gap than male-headed households.

A2.2 Difference-in-Difference

A second reason to use a regression analysis is to conduct the Difference-in-Difference approach as outlined in chapter 3. In this case a formula with the following structure would be estimated:

$$Y = \beta_0 + \beta_1 * Treatment + \beta_2 * Post + \beta_3 * Treatment * Post + \beta_4 * X_4 + \epsilon$$

Where:

- Y_0 is the outcome variable of interest;
- *Treatment* is a dummy variable indicating the treated group (=1) and the counterfactual group (=0);
- *Post* is a dummy variable indicating pre (=0) and post (=1) treatment
- *Treatment * Post* is a dummy variable representing the interaction between (a multiplication of) the *Treatment* and *Post* variables;
- X_4 is another variable to control for, such as cocoa farm size. Multiple other variables can be added at the same time. See section 4.3.1 for examples of commonly used control variables;
- β_0 till β_4 are coefficients to be estimated by the model.

Assuming that the outcome variable is the living income gap, if the p-value for the *Treatment * Post* variable would be lower than the significance level (e.g. 5%) in the example above, and the coefficient would be positive, it would indicate that after correcting for everything else, participation in the intervention is associated with an increase in the living income gap.

To explore
the potential
of nature to
improve the
quality of life



Wageningen Economic Research
P.O. Box 29703
2502 LS The Hague
The Netherlands
T +31 (0)70 335 83 30
E communications.ssg@wur.nl
wur.eu/economic-research

Wageningen Economic Research
REPORT 2024-038

The mission of Wageningen University & Research is "To explore the potential of nature to improve the quality of life". Under the banner Wageningen University & Research, Wageningen University and the specialised research institutes of the Wageningen Research Foundation have joined forces in contributing to finding solutions to important questions in the domain of healthy food and living environment. With its roughly 30 branches, 7,600 employees (6,700 fte) and 13,100 students and over 150,000 participants to WUR's Life Long Learning, Wageningen University & Research is one of the leading organisations in its domain. The unique Wageningen approach lies in its integrated approach to issues and the collaboration between different disciplines.