



**HAL**  
open science

# Biais d'information stratégiques dans la chaîne de valeur : le cas du cacao certifié en Côte d'Ivoire

Marine Jouvin

► **To cite this version:**

Marine Jouvin. Biais d'information stratégiques dans la chaîne de valeur : le cas du cacao certifié en Côte d'Ivoire. Economics and Finance. Université de Bordeaux, 2024. English. NNT : 2024BORD0201 . tel-04833260

**HAL Id: tel-04833260**

**<https://theses.hal.science/tel-04833260v1>**

Submitted on 12 Dec 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

**DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX**

ÉCOLE DOCTORALE ENTREPRISE, ÉCONOMIE ET SOCIÉTÉ (N°42)

SPÉCIALITÉ : SCIENCES ÉCONOMIQUES

Par **Marine JOUVIN**

**Biais d'information stratégiques dans la chaîne de valeur : le cas du cacao  
certifié en Côte d'Ivoire**

Sous la direction de **M. Tanguy BERNARD** et **Mme Delphine BOUTIN**

Soutenue publiquement le 11 octobre 2024

Membres du jury :

**M. Tanguy BERNARD**,  
Professeur, Université de Bordeaux, **Directeur de thèse**

**Mme Delphine BOUTIN**,  
Maître de conférences, Université de Bordeaux, **Co-directrice de thèse**

**M. Kenneth HOUNGBEDJI**,  
Chargé de recherche IRD, Université Paris-Dauphine, **Rapporteur**

**Mme Laure LATRUFFE**,  
Directrice de recherche INRAE, Université de Bordeaux, **Examinatrice, Présidente du jury**

**M. Travis LYBBERT**  
Professeur, University of California, Davis, **Examineur**

**Mme Eva-Marie MEEMKEN**,  
Maître de conférences, ETH Zurich, **Rapporteure**

**M. Joseph LARROSE**,  
Directeur général adjoint, Touton S.A., **Membre invité**



## **Title: Strategic misreporting along the value chain: the case of certified cocoa in Côte d'Ivoire**

**Abstract:** This thesis investigates the challenges raised by information asymmetries in certified cocoa value chains in Côte d'Ivoire. It aims to detect and assess the extent of non-compliance with certification standards among cocoa farmers, middlemen, and cooperatives. The first chapter measures the reliance on child labor among certified cocoa farmers using an indirect questioning method, the list experiment, to mitigate social desirability bias. It reveals significant underreporting of child labor when classic, direct questioning methods are used. The second chapter investigates dishonest behaviors among cocoa middlemen through a lab-in-the-field experiment, and highlights the high prevalence of cheating and the effectiveness of monitoring and sanction mechanisms in reducing such dishonest behaviors. The third chapter explores side-sourcing practices, i.e. whether conventional beans are labeled as certified ones, within certified cocoa supply chains by cross-checking transaction data from farmers, middlemen, and cooperatives, uncovering widespread discrepancies and potential side-sourcing from cooperatives. Overall, these findings highlight the shortcomings of current certification standards and emphasize the need for enhanced monitoring and enforcement mechanisms, to ensure the credibility and effectiveness of certification in the cocoa industry.

**Keywords:** Agricultural value chains, certification, information asymmetries, non-compliance, child labor, misreporting.

---

## **Titre : Biais d'information stratégiques dans la chaîne de valeur : le cas du cacao certifié en Côte d'Ivoire**

**Résumé :** Cette thèse s'intéresse aux défis liés à la présence d'asymétries d'information dans la chaîne de valeur certifiée du cacao en Côte d'Ivoire. Son objectif est de détecter et d'évaluer l'ampleur du non-respect des normes de certification parmi les producteurs de cacao, les intermédiaires, et les coopératives. Le premier chapitre mesure la proportion de producteurs de cacao certifiés ayant recours au travail des enfants en utilisant une méthode de questionnement indirect, le "list experiment", afin de réduire le biais de désirabilité sociale. Il révèle une sous-déclaration significative du recours au travail des enfants par rapport aux méthodes classiques de questionnement direct. Le deuxième chapitre examine les comportements malhonnêtes des intermédiaires du cacao lorsqu'ils font face à des asymétries d'information. Utilisant un "lab-in-the-field", il met en lumière une prévalence élevée de la triche parmi les intermédiaires et souligne l'efficacité des mécanismes de surveillance et de sanction pour réduire de tels comportements. Le troisième chapitre explore les pratiques de "side-sourcing" au sein des chaînes de valeur certifiées, c'est-à-dire lorsque du cacao conventionnel est labellisé comme certifié, en recoupant des données de transactions provenant de producteurs, d'intermédiaires, et de coopératives. Il montre d'importants écarts en termes de volume de cacao certifié reporté par les producteurs et les coopératives, soulignant un important recours au "side-sourcing" par ces dernières. Dans l'ensemble, ces résultats mettent en évidence les limites des normes de certification actuelles et soulignent la nécessité de renforcer les mécanismes de surveillance et de sanction, pour garantir la crédibilité et l'efficacité des standards de certification dans l'industrie du cacao.

**Mots-clés :** Chaînes de valeur agricoles, certification, asymétries d'information, non-conformité, travail des enfants.



*In a world of asymmetry, honesty is a precious commodity.*

- George A. Akerlof



# Remerciements

C'est avec plaisir que j'exprime ici ma profonde gratitude envers toutes les personnes qui ont contribué de près ou de loin à la réalisation de cette thèse.

Je tiens d'abord à remercier l'ensemble des membres de mon jury d'avoir accepté d'y prendre part et d'évaluer mon travail. Merci à Eva-Marie Meemken et Kenneth Hounbedji d'avoir accepté d'être les rapporteurs de cette thèse, et à Laure Latruffe et Travis Lybbert, pour avoir pris le temps de lire mon travail et de participer au jury de soutenance.

Je souhaite également remercier chaleureusement mes directeurs de thèse, Tanguy Bernard et Delphine Boutin, pour m'avoir encadrée et guidée durant ces années de doctorat. Tanguy, merci de m'avoir accordé en premier lieu ta confiance pour mener à bien cette thèse CIFRE. Merci à vous deux pour votre grande disponibilité, votre bienveillance, et votre optimisme contrastant avec mon réalisme ! Votre accompagnement et vos conseils durant ces années de thèse ont joué un rôle fondamental dans la réalisation de ce travail, ainsi que dans ma formation de chercheur.

Mes remerciements vont également à Touton S.A., et plus particulièrement à Joseph Larrose, pour m'avoir fait confiance tout au long de ce partenariat, et m'avoir donné la possibilité de mettre un pied dans ce monde si complexe, mais passionnant, qu'est le secteur du cacao. Merci à Touton Négoce Côte d'Ivoire (TNCI), de m'avoir accueillie pour mes nombreuses missions de terrains, et plus particulièrement à Achille Kouamé, dont la rigueur, le soutien, et les connaissances intarissables sur le sujet m'ont été précieux tout au long de ce travail. Merci à Sénèque Kouadio, sans qui aucune mission de terrain n'aurait été possible. J'ai sincèrement apprécié toutes les missions que nous avons partagées, malgré les difficultés, les imprévus, et les défis. Je tiens également à remercier l'ensemble des membres du Département Approvisionnement Durable (DAD) à TNCI, qui, d'une manière ou d'une autre, m'ont accompagnés dans chacun de mes projets de recherche. Je souhaite adresser de chaleureux remerciements à tous les enquêteurs avec qui j'ai eu la chance de travailler au cours de ces années de thèse. Je tiens à remercier plus personnellement Danon, Gérald, Inès, Jacques, et Hyppolite. Travailler et évoluer à vos côtés fut un plaisir ; vos connaissances précieuses sur le secteur du cacao et votre expertise en tant qu'agents enquêteurs n'ont cessé de m'impressionner. Votre dévouement, malgré les difficultés (récurrentes...) rencontrées sur le terrain force l'admiration, et mon travail ici présenté repose en très grande partie sur le vôtre.

Qu'aurait été cette aventure sans l'équipe des jeunes Toutonnais ? Ingrid, Patrick, Laura, Thibaud, Hugo, Jean-Philippe : vous avez grandement participé à maintenir mon moral au beau fixe tout au long de ces années de thèse, et vous m'avez fait rire à bien des égards. Florian, merci pour ta disponibilité et ton engagement : deux des trois articles de recherche présentés dans cette thèse n'existeraient pas sans ton travail et ton appui. Enfin, je remercie cette aventure de t'avoir mis sur ma route, Oumaïma. Si nos missions de terrains en Côte d'Ivoire (pensée particulière pour Alépé !) et nos métiers aux nombreux points communs nous ont amenés à nous rapprocher, c'est sur les valeurs que nous partageons que s'est construite notre amitié. Merci pour ton soutien inconditionnel et tes conseils avisés. Tu m'as permis de surmonter de difficiles moments au cours de ces dernières années. Je souhaite à tout le monde de croiser un jour ton chemin.



Je tiens à remercier l'ensemble de mes collègues doctorants (que nous nous soyons rencontrés au GREThA ou à BSE), qui m'ont si bien accueillie au début de ma thèse, et avec qui j'ai partagé de bien beaux moments tout au long de mon parcours : Louis, Sun, Lucile, Valentina, Jérémy, et Élodie. Des remerciements particuliers à Lucie, Sébastien, Jeanne, Diane, et Arnaud, pour avoir toujours été à mes côtés au cours de cette entreprise commune. Merci pour la constante écoute bienveillante, les conseils avisés, les séances de re-motivation, et votre amitié. Vous avez été un pilier certain tout au long de mon doctorat.

Anaïs, merci pour ton amitié depuis plus d'une décennie ! Ton écoute attentive et ton soutien inconditionnel ont été précieux pendant ces cinq années de thèse. L'arrivée d'Eva et d'Eliott restera un moment de joie certain dans cette longue aventure. Sherylin, merci pour ces moments de retrouvailles toujours remplis de joie ! Notre amitié est désormais ponctuée d'une aventure commune, le doctorat... Jamais nous n'aurions anticipé ce destin lié lors de notre rencontre, il y a quinze ans maintenant !

Je souhaite également remercier mes parents, pour m'avoir toujours soutenue dans mes études, même lorsque mes choix sortaient des sentiers battus. Je ne serai pas là aujourd'hui sans votre soutien et vos sacrifices. Merci à mon frère, Jonathann, et sa femme, Julie, pour leurs encouragements, leur écoute, et nos discussions à refaire le monde ! Merci à Adeline, ma sœur, et Florent, mon beau-frère, pour tous ces moments de joie et de légèreté partagés. Une pensée inévitable pour Éléa, Théo, Lina, Tymaën, et Élaïne, qui ont grandement contribué à me faire relativiser sur de nombreux sujets et m'ont apporté beaucoup de joie pendant ces années de travail.

Damien, je n'aurai jamais pu achever cette expérience sans ton soutien quotidien indéfectible. Mille mercis pour ta présence, ton écoute, et ta patience, qui m'ont permis de surmonter les nombreux moments de doute, de stress, et de remise en question. Merci d'avoir eu confiance en moi et en mes capacités lorsque je n'y croyais plus moi-même. Merci pour tout ces moments de joie et d'aventures partagés. Je vous en souhaite encore d'innombrables !





# Contents

<b>General Introduction</b>	1
<b>Chapter 1.</b> _____	
Addressing Social Desirability Bias When Measuring Child Labor Use: An Application to Cocoa Farms in Côte d'Ivoire	13
<b>Chapter 2.</b> _____	
Information Asymmetries, Risk and Dishonesty: A Behavioral Experiment Among Middlemen in Certified Cocoa Value Chain	53
<b>Chapter 3.</b> _____	
Traceability Along the Value Chain: the Case of Certified Cocoa in Côte d'Ivoire	105
<b>General Conclusion</b>	171
Detailed Table of Contents	177
List of Tables and Figures	179



# General Introduction

On April 19th, 2023, the European Parliament enacted a law prohibiting the import of a range of commodities if their production is associated with deforestation or forest degradation.<sup>1</sup> More recently, both the European Parliament and the European Council have also supported a new regulation banning imports of products associated with forced labor.<sup>2</sup> This recent tightening of European legislation aligns with growing concerns among civil society and regulators in high-income markets regarding the ethical and sustainable sourcing of commodities, often originating from low-income countries. Such changes in legislation increase pressure on exporters of targeted commodities to provide evidence that the products they trade are not associated with deforestation or forced labor. To this end, exporters mostly rely on certification schemes through which one can trace commodities to producers certified for compliance with environmental and/or social requirements. In line with the Sustainable Development Goals (SDGs) set by the United Nations<sup>3</sup>, certification initiatives have received significant support from international development agencies, non-governmental organizations (NGOs), and governments through capacity building, technical assistance, and financial incentives to support farmers' adoption of certification standards, mainly in developing countries.

Agricultural certification schemes are designed to ensure that goods meet specific social, environmental, and economic sustainability standards. These programs include precise criteria that farmers must comply with, including the adoption of environmentally sustainable farming practices and non-reliance on forced and child labor. To help meet these criteria, farmers

---

<sup>1</sup>See The Deforestation Regulation (EU) 2023/1115, which took effect on June 29th, 2023. The commodities covered by this new legislation are cocoa, coffee, oil palm, soya, and rubber. Wood and cattle are also targeted.

<sup>2</sup>See the Proposal for a regulation of the European Parliament and the European Council on prohibiting products made with forced labor on the Union market (COM(2022)0453 – C9-0307/2022 – 2022/0269(COD)).

<sup>3</sup>Such as poverty alleviation, environmental conservation, and sustainable agriculture.

are provided with extensive training on topics ranging from good agricultural practices to the harmful effects of child labor on their health. Certified farmers receive extra income for adhering to certification guidelines, called a certification premium. Independent organizations conduct regular audits to verify compliance with these standards, providing consumers with assurance that the products they purchase meet the stated criteria. While the global area of agricultural land covered by such standards remains small and is estimated to be less than 2% (Meemken et al., 2021), it is significant for specific commodities. This is especially the case for coffee and cocoa, for which an estimated 45.3% and 34.5% of cultivated areas worldwide are covered by certification schemes, respectively (Kemper et al., 2023).

Whether and to what extent certification standards are associated with benefits for farmers in low-income countries is unclear. If evidence suggests that certification has a positive effect on prices received by certified farmers (Meemken, 2020), effects on other economic indicators, such as profits, yields, production costs, or household living standards, are mixed, depending on the country, commodity, and type of certification (Meemken, 2020; Oya et al., 2018; Akoyi and Maertens, 2018). In line with the main objective of certification, the literature has focused on assessing its potential benefits for farmers. Comparatively, little attention has been given to farmers' compliance with certification requirements or potential issues with the implementation of certification. Yet, farmers involved in certification schemes usually operate within extended supply chains involving multiple intermediaries whose relationships are prone to information asymmetry, reinforced by the impossibility of observing from the product itself whether it was produced in line with the requested criteria - a credence attribute, that raises concerns about mislabeling.

In Côte d'Ivoire, the world's leading cocoa-producing country which accounts for 40% of global cocoa production, recent evidence suggests that 60% of farmers are now linked to certification standards (The World Bank, 2019). Yet, it is also estimated that the industry still employs 790,000 children and is the largest contributor to deforestation in the country (Sadhu et al., 2020; Kalischek et al., 2023). As a result, the sector faces increasing pressure to achieve traceability at the farmer level, particularly as cocoa is one of the commodities targeted by the new European import restrictions. However, most industry players only reach traceability

at the cooperative level as achieving first-mile traceability is costly and challenging (Huetz-Adams and Fountain, 2022). The Ivorian cocoa supply chain is made of several geographically dispersed players, whose relationships are fraught with information asymmetries. Exporters looking to supply certified cocoa cannot legally purchase cocoa directly from farmers in Côte d'Ivoire.<sup>4</sup> They are thus required to buy from farmer groups (cooperatives) or buying centers, who in turn mandate middlemen to buy cocoa from thousands of geographically dispersed smallholder farmers and deliver it to cooperatives or buying centers' warehouses. As a result, multiple levels of principal-agent relationships characterize the cocoa supply chain in Côte d'Ivoire: exporters are the principals, and cooperatives have a double agency role, as agents for the exporters and principals for the middlemen, who in turn serve as agents for the cooperatives. In this complex setting, exporters face challenges ensuring that certification standards are respected throughout the supply chain without a reliable monitoring system.

Compliance with certification standards is currently monitored at several levels in the Ivorian cocoa sector. Exporters and cooperatives maintain their own internal inspection and monitoring systems to ensure compliance with certification requirements related to issues such as child labor, deforestation, and sustainable farming practices. Exporters often provide dedicated staff within cooperatives to further ensure compliance and collect data and documentation to support certification requirements. However, as all value chain actors potentially benefit from non-reliable traceability, the key part of the monitoring framework lies in third-party audits conducted by certification bodies. These theoretically independent audits are intended to provide an unbiased assessment of whether cooperatives and their associated farmers comply with certification standards. Typically conducted on an annual basis, third-party audits involve comprehensive evaluations of cooperative practices. However, their effectiveness is considered limited due to their pre-scheduled nature and the limited number of farmers effectively visited. Existing monitoring mechanisms are therefore likely to be inefficient and only cover a small number of players in the cocoa supply chain, leaving room for them to engage in opportunistic behavior, as the probability of being detected and sanctioned is considered low.

---

<sup>4</sup>See ordinance n°2011-481 of December 28th, 2011, and decree n°2012-1008 of October 17th, 2012.



Agency relationships combined with the lack of a large-scale, effective monitoring system, open room for opportunistic behaviors at every stage of the cocoa supply chain. Thus, it is possible that some farmers do not comply fully with the requirements set by certification schemes, which might be costly to implement, such as applying good agricultural practices or avoiding the use of child labor. Since certified cocoa is sold at a higher price than conventional one through certification premiums, it is also possible that certified cocoa farmers gather conventional cocoa from other non-certified farmers and sell it all as certified cocoa, thereby passing part of the certification premium to uncertified cocoa farmers. Similarly, middlemen are likely to misreport the true nature of the cocoa they buy from farmers, whether certified or conventional, to capture the rent associated with certified cocoa. The same logic holds true for cooperatives when providing cocoa to exporters (Stoop et al., 2021).

This thesis investigates opportunistic behaviors in the certified cocoa supply chain in Côte d'Ivoire and relies on original, first-hand data collected in the supply chain of a top-five global cocoa exporter. It consists of three chapters, all of which aim to shed light on compliance with certification standards of cocoa farmers, middlemen, and cooperatives, by detecting and measuring opportunistic behavior among these players.

## **Outline and contributions**

**Chapter 1.** The first chapter aims to measure reliance on child labor by cocoa farmers belonging to certified cocoa supply chains. Measuring the use of child labor through direct questioning of farmers is prone to social desirability bias, as the use of child labor is prohibited both by the Ivorian government and certification schemes, and farmers who fail to comply with the ban on child labor are exposed to sanctions. We therefore use a list experiment (LE), an indirect questioning method designed to reduce social desirability bias by embedding the sensitive question in a series of other, less sensitive ones, in addition to the direct questioning traditionally used in surveys designed to measure the use of child labor. We surveyed 4,458 UTZ and Rainforest Alliance farmers from December 2019 to March 2020, located in different cocoa-producing areas: Agnéby-Tiassa, Cavally, Grands Ponts, Guémon, Gôh, Lôh-Djiboua, Nawa, and Tonkpi.

We find that cocoa farmers belonging to certified value chains are indeed prone to social desirability bias when directly questioned about their use of child labor on cocoa farms. The proportion of farmers using child labor to perform dangerous tasks on cocoa farms is twice as high when they are surveyed using the LE as when they are asked directly. Our findings provide evidence that the use of child labor is common among certified cocoa farmers. 24% of them were helped by at least one child under 16 for harvesting and breaking the cocoa pods over the 12 months preceding the survey, 21% for preparing their farm, and 25% employed and paid at least one child to perform any task on their cocoa farm. Drawing on an additional survey of 13,040 additional certified farmers conducted at the end of 2022, we further show that the LE approach is robust to changes in design, such as the length of the recall period, the type of non-sensitive items, and the addition of a comprehension test before its implementation.

Chapter 1 makes three contributions to the literature. First, it highlights an additional source of error associated with measuring child labor use: social desirability bias. This constitutes a substantial contribution to the literature on the measurement of child labor and child labor use, which relies mainly on direct questioning methods, both in surveys conducted by researchers and those carried out by the private cocoa sector. The results suggest that using direct questioning alone is likely to bias policies implemented to reduce child labor due to potential poor targeting. Second, this chapter shows that certified cocoa farmers rely to a significant extent on child labor, despite the ban from the Ivorian government and the certification standards. This finding raises concerns about the effectiveness of monitoring by certified cooperatives and certification bodies. Finally, Chapter 1 contributes to the extensive literature on the LE further showing the relevance of its use to measure reliance on child labor and its robustness to design changes in certain contexts. Findings from Chapter 1 further suggest that one should be cautious about using the LE to measure the same phenomenon over time. Researchers using this approach must be able to identify any potential confounding factors, to ensure that only the variation in sensitive behavior over the years is captured by the LE estimator.

**Chapter 2.** The second chapter investigates dishonest behaviors among cocoa middlemen in the presence of information asymmetries, and their responses to monitoring and sanction mechanisms. We draw on data from a lab-in-the-field experiment that we designed and conducted among 151 cocoa middlemen operating in certified value chains. Specifically, we implemented a modified version of the "die-under-the-cup" task where middlemen were invited to roll a die in private and then report the outcome to the enumerator, who used the reported outcome to determine the player's payout. Players have an incentive to lie in the event of a losing die, to maximize their winnings. Middlemen played 15 rounds across which the risk of being monitored and sanctioned for cheating was sequentially included.

We find that cheating is prevalent among cocoa middlemen in such a context, with 78% of players cheating at least once during the game. 22% of the middlemen, however, never cheated during the 15 rounds of the game, even when their true die roll outcome could not be observed by enumerators and they risked no penalty. This result is in line with the literature on cheating games, which states that some individuals have a high cheating cost such that, despite the associated benefits, they will never cheat. We find heterogeneous cheating patterns among middlemen who cheated: 59% of them consistently cheated when faced with a losing outcome, even when the risk of detection and sanction was high, while 19% only cheated occasionally. We further identify key individual characteristics, such as age, religion, and attitudes toward risk, that influence the likelihood of cheating. Last, findings from Chapter 2 suggest that the introduction of monitoring and sanctions leads to a significant reduction in cheating, highlighting the effectiveness of such mechanisms in deterring dishonest behaviors in situations of information asymmetries.

The findings in Chapter 2 contribute to the experimental literature on cheating behavior in two ways. First, our computerized version of the "die-under-the-cup" task allows us to observe individual cheating behaviors rather than aggregate ones, making it possible to investigate individual determinants of cheating. Second, in addition to looking at the effect of monitoring players' cheating behavior, we also investigate the effect of financial sanctions applied in the event of cheating, which has been neglected in the literature so far. Chapter 2 also contributes to the literature on middlemen's role and behavior in agricultural value chains. It provides

interesting insights into the activity of cocoa middlemen and suggests that while most of them exploit information asymmetries when possible, a fair proportion of middlemen exhibit a strong preference for honesty, regardless of the incentives.

**Chapter 3.** Chapter 3 investigates potential side-sourcing, i.e. whether players in the value chain buy conventional cocoa and sell it as certified, and its extent within the certified cocoa value chain. We further conduct a randomized controlled trial to assess the effects of the implementation of a new digital traceability system (DTS) upstream in the supply chain, aimed at reducing the opacity of transactions between middlemen (and therefore cooperatives) and cocoa farmers. We use a large set of original and first-hand data: farmer and middleman surveys, DTS data collected by middlemen, and traceability registers provided by cooperatives. We cross-check transaction data reported by cocoa farmers, middlemen, and cooperatives, covering the same farmers and period, to assess whether any reporting differences arise. We then use farmers and middlemen survey data, as well as the mid- and end-season versions of the cooperatives' traceability registers for the 2022-2023 cocoa season, to investigate the effects of DTS on cooperatives sourcing and misreporting strategies.

We find significant reporting discrepancies in terms of volumes declared as sold/purchased by farmers/middlemen on the one hand, and volumes declared as purchased by cooperatives on the other. Reporting discrepancies are systematic within cooperatives, ranging between 123% to 870% of the farmers' reports, indicating that cooperatives engage extensively in side-sourcing. The implementation of the DTS experienced significant reluctance on the part of cooperatives and middlemen, leading to a rather low uptake of the DTS within treated cooperatives and sections. Yet, we find that the implementation of a DTS has led cooperatives to adjust their cocoa sourcing strategy, as middlemen bought more cocoa from farmers located in control sections of treated cooperatives than in treated ones. Focusing on cooperatives' traceability registers, we find that treated cooperatives report far more transactions and overall volume than control ones, particularly at the beginning of the season (October/November). This is in line with field observations that this period is most prone to side-sourcing activities. Comparing transactions reported at the beginning of the cocoa season

in the first traceability register with those reported for the same period in the final one, we provide evidence that cooperatives manipulate transaction data extensively throughout the season, especially treated cooperatives. We further show that treated cooperatives engage in different misreporting strategies depending on treatment intensity.

Chapter 3 makes three contributions to the literature. First, by shedding light on the existence and magnitude of side-sourcing in certified cocoa value chains, it contributes significantly to the literature on certification schemes in agricultural value chains. Second, it provides evidence that intermediaries in the cocoa value chain (cooperatives and, to a lesser extent, middlemen) engage in dishonest behavior on a large scale, exploiting information asymmetries on credence characteristics, thus building additional evidence into the literature on the roles of intermediaries in agricultural value chains. Finally, Chapter 3 also contributes to the literature on traceability systems in agricultural value chains, by providing empirical evidence of supply chain actors' adherence to such tools and their impact on the behavior of upstream supply chain players.

## **Thesis context**

This thesis was funded through a CIFRE grant (Industrial Contract for Training through Research) by the French National Association of Research and Technology (ANRT), in collaboration with Touton S.A, a French commodities trading company that ranks among the global top 5 cocoa trading companies. Operating in 14 countries across 4 continents, Touton S.A. annually processes 350,000 tonnes of cocoa and 105,000 tonnes of coffee. Over 130,000 farmers from whom they source cocoa, coffee, or ingredients are part of a certification program in key producing countries such as Côte d'Ivoire, Ghana, Kenya, Madagascar, Nigeria, and Uganda. This thesis is part of Touton S.A.'s aim to increase its due diligence efforts and to identify and measure potential internal issues within its supply chain, such as child labor and deforestation. In this context, one of the company's goals is to improve the traceability of products, an objective recently further strengthened by the commercial risks posed by stricter European import regulations, given that the European market accounts for 55% of their sales.

As Corporate Social Responsibility policy researcher at Touton S.A., my work was divided into two main roles: applied research and advisory work in data management, data quality, and survey implementation. This work extended beyond the cocoa sector in Côte d'Ivoire to include projects in Ghana and Madagascar, but also projects for coffee in Uganda. My research focused on key issues related to the company's sustainability strategy. I first sought to improve the measurement of child labor use among certified farmers, Touton S.A. being well aware that the direct questioning used in classic survey questionnaires leads to respondents misreporting (Chapter 1). Such biased reporting hinders the identification of at-risk farmers and priority areas for Child Labor Monitoring and Remediation Systems (CLMRS) implementation. I therefore used the LE, an indirect questioning method, to obtain a less biased measure of farmers' use of child labor. I was responsible for designing the household survey administered to the entire pool of Touton-certified farmers in Côte d'Ivoire, as well as designing and integrating a LE into the survey. I trained over 40 enumerators in survey practices and the LE approach and monitored their work throughout the household surveys, which were conducted approximately six months each year over three years.

The second area of my research focused on cocoa traceability, a fundamental issue for the industry and Touton S.A., even before new European regulations restricted the import of certain commodities into Europe. Through extensive fieldwork and meetings with all supply chain actors (farmers, middlemen, and cooperatives), as well as engaging in extensive internal discussions within Touton S.A. about traceability, I gained valuable insights into the most opaque aspects of traceability in the cocoa supply chain. To improve downstream cocoa traceability, the company developed an internal digital traceability application, designed for middlemen use. I played an active role in helping to design and configure this tool with the IT team, testing it in the field, and adjusting it during the pilot phase. I chose to investigate the effect of introducing this new traceability tool using a randomized controlled trial (RCT) (Chapter 3), which involved extensive discussions and convincing each player to take part in the project. I was responsible for introducing and training local Ivorian teams on how to use the application and then monitoring their usage over nearly two years. We faced significant resistance from cooperatives and middlemen in adopting the application, leading

to operational delays. Initially scheduled for the 2021-2022 main cocoa season, the RCT was postponed to the 2022-2023 season.

Anticipating resistance from the cocoa middlemen to use the traceability application, I conducted an experimental game in August 2021 to investigate their behavior in situations of asymmetric information (Chapter 2). Specifically, this experimental cheating game aimed to shed light on their behavior when mechanisms similar to those of the traceability application (i.e., monitoring), were introduced. Running this experiment involved training 5 enumerators on the protocol of the experimental game, and on daily data monitoring over one month. The results of this experiment, designed to inform the RCT outcomes originally planned for October-March 2021-2022, could not be linked to the RCT that finally took place in 2022-2023, as the cocoa middlemen who took part in the game were significantly different from those who took part in the RCT.

Overall, working for a global top 5 cocoa trading company for four years provided me with a solid insight into the industry, the stakeholders, and the issues at stake in the cocoa sector. I had the opportunity to engage with every actor in the supply chain: farmers, middlemen, cooperatives, the trading company, and its partners. I conducted applied research on real issues faced by players in the cocoa sector, and benefited from complete independence in both the direction and conclusions of my research.

## References

- Akoyi, K. T. and Maertens, M. (2018). Walk the talk: private sustainability standards in the Ugandan coffee sector. *The Journal of Development Studies*, 54(10):1792–1818. [2]
- Huetz-Adams, F. and Fountain, A. C. (2022). 2022 Cocoa Barometer. Technical report, the Cocoa Barometer Consortium. [3]
- Kalischek, N., Lang, N., Renier, C., Daudt, R. C., Addoah, T., Thompson, W., Blaser-Hart, W. J., Garrett, R., Schindler, K., and Wegner, J. D. (2023). Cocoa plantations are associated with deforestation in Côte d'Ivoire and Ghana. *Nature Food*, 4(5):384–393. [2]
- Kemper, L., Sampson, G., Larrea, C., Schlatter, B., Luna, E., Dang, T. D., and Willer, H. (2023). The State of Sustainable Markets 2023: Statistics and Emerging Trends. Technical report, ITC, Geneva. [2]
- Meemken, E. M. (2020). Do smallholder farmers benefit from sustainability standards? A systematic review and meta-analysis. *Global Food Security*, 26(April):100373. [2]
- Meemken, E. M., Barrett, C. B., Michelson, H. C., Qaim, M., Reardon, T., and Sellare, J. (2021). Sustainability standards in global agrifood supply chains. *Nature Food*, 2(10):758–765. [2]
- Oya, C., Schaefer, F., and Skalidou, D. (2018). The effectiveness of agricultural certification in developing countries: A systematic review. *World Development*, 112:282–312. [2]
- Sadhu, S., Kysia, K., Onyango, L., Zinnes, C., Lord, S., Monnard, A., and Arellano, I. R. (2020). Sadhu et al., 2020. Technical report, National Opinion Research Center (NORC), University of Chicago. [2]
- Stoop, P., Ramanan, N., Geens, H., Lambrecht, A., and Dekeister, S. (2021). Technical brief on cocoa traceability in West and Central Africa. [4]
- The World Bank (2019). Cote d'Ivoire Economic Update. Technical report, The World Bank. [2]





---

# **ADDRESSING SOCIAL DESIRABILITY BIAS WHEN MEASURING CHILD LABOR USE: AN APPLICATION TO COCOA FARMS IN CÔTE D'IVOIRE**

*Published in the World Bank Economic Review.*

## **Abstract.**

This paper proposes new estimates of child labor use in Côte d'Ivoire's cocoa farms that are certified free of child labor. The study relies on list experiments (LE) to overcome social desirability bias associated with measuring sensitive issues, implemented on a sample of 4 458 Ivorian cocoa farmers. Findings show that 24% of them were helped by at least one child under 16 for harvesting and breaking the cocoa pods during the past 12 months, 21% for preparing their farm, and 25% employed and paid at least one child to perform any task on their cocoa farm. These results are twice as high as those declared by farmers when directly questioning them about their child labor use. This study provides evidence that the LE method, while more cognitively demanding than the direct questioning method, can be successfully understood by populations with low levels of education. Findings further show that, in this setting, the LE estimates are robust to specific LE design changes.

## 1. Introduction

Recent evidence based on household surveys suggests that the Ivorian cocoa sector is still struggling with child labor, with an estimated 790,000 children between 5 and 17 years laboring on cocoa farms in 2019 (Sadhu et al., 2020). The cocoa industry thus remains under pressure from national and international regulators, as children working on cocoa farms are widely exposed to hazardous work, one of the worst forms of child labor according to the International Labour Organization (ILO). In Côte d'Ivoire, 97% of children working on cocoa farms are thus estimated to be involved in such hazardous child labor (Sadhu et al., 2020). The ILO defines the latter as “work which, by its nature or the circumstances in which it is carried out, is likely to harm the health, safety or morals of children”.<sup>1</sup> Pressure from various civil society stakeholders to reduce child labor in cocoa has in part been fueled by a growing body of research related to the adverse effects of child labor on children’s physical and mental health (Guarcello et al., 2004; Fekadu et al., 2006; Roggero et al., 2007; Trinh, 2020) or education (Udry, 2006; Beegle et al., 2009; Putnick and Bornstein, 2015) since the early 2000s.<sup>2</sup>

In the late 2000s, traders and manufacturers began to use independent certification labels to meet growing consumer demand for ethical and sustainable cocoa.<sup>3</sup> These labels reward cocoa farmers with the farmgate price and a certification premium, the latter rewarding a more sustainable cocoa production, without deforestation and free of child labor.<sup>4</sup> Farmers that are members of certified cooperatives benefit from various training and are sensitized to the negative consequences of child labor, especially hazardous tasks, on children’s well-being. Third-party auditors annually audit certified cooperatives to ensure that their affiliated farmers comply with certification programs’ social and environmental requirements. A farmer who does not respect these requirements risks being excluded from the certified cocoa value chains, which implies losing all the benefits of certification, including the certification premium. Certification schemes and the services they provide to farmers have gradually attracted the attention of researchers in the past few years, mainly focusing on assessing the impact of

---

<sup>1</sup>Article 3 of ILO’s Convention No. 182.

<sup>2</sup>See Boutin and Jouvin (2022) for a comprehensive review.

<sup>3</sup>UTZ and Rainforest Alliance for sustainable development labels, Fairtrade International (Max Havelaar) for fair trade label, or Organic for cocoa that follows organic production methods. All sustainable development programs set up by chocolate producers themselves (e.g., Cocoa Life for Mondelez) are excluded here.

<sup>4</sup>In Côte d'Ivoire, the cocoa farmgate price is set at the beginning of each season by the Conseil Café Cacao (CCC). During the survey period (2019–2020 cocoa season), the farmgate price per kilo of cocoa was set at 825 CFA (1.48 dollars).

certification on farmers' incomes. As far as we know, no studies have investigated whether child labor use remains prevalent in the specific context of certified cocoa farms.

Investigating farmers' use of child labor on certified cocoa farms implies addressing methodological challenges to correctly measure this complex and sensitive phenomenon. Guarcello et al. (2012) highlight that child labor prevalence fluctuates significantly from one year to another in various countries, calling into question the way in which it is measured. The subsequent literature on the measurement of adult and child labor has identified key factors that may cause labor measurement errors, although evidence remains scarce for Côte d'Ivoire.

The choice of the respondent to be asked about child labor is an important consideration. Proxy respondents, such as household heads, generally under-report when asked about child labor issues compared to self-report from children (Dillon, 2010; Bardasi et al., 2011; Galdo et al., 2018; Janzen, 2018). Findings from Galdo et al. (2020) also suggest that discrepancies between child and proxy respondent responses might depend on the child's gender. However, whether children can provide accurate answers remains unclear, as the literature on children's cognitive development provides mixed evidence (Borgers et al., 2000; Fuchs, 2008). The recall period over which respondents are asked about labor issues is also likely to generate erroneous answers, as it can be difficult for them to recall accurately and over long periods the amount of labor used, or even their own working time. Several studies have confirmed the existence of such recall bias in labor surveys in the African context, suggesting that the choice of recall period sometimes leads to underestimate or overestimate labor use or working time, depending on the context (Beegle et al., 2012; Arthi et al., 2018; Gaddis et al., 2021). Bardasi et al. (2011) further highlight that labor-related estimates are sensitive to the length of the module covering labor issues, thus suggesting that survey design is also a factor that may influence labor-related estimates.

It is also likely that, depending on the type of child labor concerned, national legislation, and social standards, the measurement of child labor related behaviors may be prone to social desirability bias. This means that respondents may not honestly answer child labor questions to provide socially desirable answers, and avoid negative legal or social consequences (Tourangeau and Yan, 2007). This behavior is common when respondents are questioned about sensitive topics, both from a personal or societal perspective (Tourangeau and Yan, 2007), such as politics (Çarkoğlu and Aytaç, 2015), health (De Cao and Lutz, 2018; Moseson et al., 2015) or intimate partner violence (Cullen, 2020).

Attempting to detect cocoa households relying on child labor on their farms, the cocoa industry is increasingly implementing child labor monitoring systems among farmers. These systems, which aim at identifying and addressing cases of child labor, usually rely on surveying cocoa households specifically on their use of child labor, in a dedicated child labor questionnaire that focuses on both light and hazardous work, and uses direct questioning methods.<sup>5</sup> We argue that a social desirability bias is likely at play when cocoa farmers from certified cooperatives are asked about their use of child labor, given that it is strictly prohibited and that they could suffer negative economic repercussions, by being excluded from certified cocoa value chains. This bias may be important since the questionnaires used so far: 1) are usually conducted as part of a CLMRS, *i.e.*, a system specifically designed to identify and address child labor, 2) ask questions exclusively or almost exclusively about child labor, 3) use direct questioning methods which explicitly signals to the respondent that the focus is on child labor issues.

This social desirability bias has not received much attention in the measurement of child labor issues, and has rarely been considered in reports assessing the prevalence of child labor use in mainstream or certified cocoa industry. Yet, estimates of child labor use or prevalence obtained through a direct questioning method could significantly underestimate the phenomenon in the case of a proven social desirability bias.

This paper relies on the list experiment (LE) approach, an indirect questioning method, to measure the proportion of certified cocoa farmers in Côte d'Ivoire who use child labor to perform hazardous work on their farms. By hiding the respondent's answer to the sensitive question among other answers to innocuous questions, this method is expected to reduce social desirability bias (Imai, 2011). This approach is used to measure farmers' use of children under the age of 16 on their cocoa farms. Specifically, respondents are asked whether a child helped them with harvesting and pod breaking, preparing the farm, or if they had employed and paid a child for any work on their cocoa farm. Results from the LEs are compared to those obtained from directly questioning farmers about their use of child labor for hazardous tasks, as done in existing surveys aiming at measuring child labor use. This study covers 4 458 cocoa farmers certified by the UTZ and RA labels, surveyed between December 2019 and March 2020. Two additional survey waves conducted in 2022 are used to test the robustness of the LE estimates to design changes, and farmers' level of understanding of the LE method.

Findings show that the prevalence of child labor use estimated using the indirect measure is twice as large as that obtained from direct questioning, indicating that cocoa farmers likely exhibit social

---

<sup>5</sup>Usually called Child Labour Monitoring and Remediation Systems (CLMRS) survey.

desirability bias when asked about their use of child labor. Using LEs, this paper provides evidence that 24% of certified cocoa farmers were helped by a child under 16 for harvesting and breaking cocoa pods during the past 12 months, 21% for preparing their farm, and 25% employed and paid a child to perform any task on their cocoa farm. Findings further show that the LE method is well understood by cocoa farmers, despite their low level of education. There is no evidence of a better understanding of LE among educated farmers compared to those with no education. Last, this paper suggests that LE estimates are robust to design changes in this setting. Specifically, light modifications to the control items, as well as reduction of the recall period, have a negative average effect on the number of items to which farmers respond “yes”, but in a uniform way between control and treated farmers, thus not affecting the difference-in-means between the two groups.

This paper contributes to the literature on child labor in two ways. First, unlike most studies that attempt to measure child labor related behavior, an indirect method of questioning is used to reduce the social desirability bias among respondents and obtain a more accurate measure of child labor use. Findings support the existence of a social desirability bias and question the relevance of using only the direct questioning method in this context, which can contribute to the design of poorly targeted policies. Accordingly, surveys used by the cocoa industry, such as Child Labour Monitoring and Remediation Systems (CLMRS) surveys, likely underestimate the magnitude of child labor and farmers’ reliance on children to work on cocoa farms. More broadly, the use of an indirect questioning method, such as the list experiment, should be considered along with direct questioning when individuals are asked about child labor, and have incentives to hide the truth. Second, this study focuses on measuring hazardous child labor use on certified cocoa farms. These findings contribute to the literature on the use of certification schemes in agricultural supply chains, by providing evidence that certification bodies are failing to keep Ivorian cocoa farms free of child labor, despite their commitment. It raises concerns about the effectiveness of certification labels in reducing child labor in the cocoa sector, as well as the effectiveness of currently implemented monitoring systems.

This study also contributes to the list experiment literature, suggesting that this indirect questioning method can be used to measure the use of child labor in particularly sensitive contexts. Although considered more cognitively demanding than the direct questioning method, this paper shows that cocoa farmers with low levels of education can successfully understand the method. Results from the robustness test also contribute to the discussion regarding the LE estimates’ sensitivity to design changes. Although the recent literature highlights a significant sensitivity of LE estimates to the choice

of control items (Chuang et al., 2021), findings from this setting show that, along with a reduction in the recall period, slight modifications of the control items have no significant impact on the difference-in-means estimator.

While the list experiment has become increasingly popular for measuring sensitive behaviors, this indirect questioning method also presents some limitations. Future research should focus on assessing the robustness of the list experiment to survey design and implementation, as evidence on this matter remains limited. It may help researchers and practitioners appreciate the contexts where the use of LE would be most appropriate, and adopt better implementation practices. In addition, any attempt to validate the list experiment method against objective data would also increase confidence in using this method to better measure child labor related behaviors, and overcome the associated social desirability bias.

The remainder of this article is organized as follows. Section II provides background on child labor in the cocoa sector in Côte d'Ivoire. Section III presents the data, the design and implementation of the list experiment, and the empirical strategy. Results are discussed in section IV, along with robustness checks. Section V concludes.

## **2. Child labor in the Ivorian cocoa sector**

### **2.1 Cocoa sustainability standards and child labor**

Child labor has become a major issue in the cocoa sector since the early 2000s, when the cocoa industry committed to reducing the worst forms of child labor by 70% by 2020.<sup>6</sup> Despite ongoing efforts to achieve this goal by both governments and cocoa companies, the sector has failed to meet its commitment. Indeed, many children continue to work on cocoa farms in West Africa, many of whom perform dangerous tasks that may threaten their health, such as land clearing, pruning, pod-breaking with a sharp tool, harvesting with a machete, or manipulating agrochemicals (Sadhu et al., 2020). In Côte d'Ivoire, all these activities are prohibited for children under the age of 18 and recognized as hazardous

---

<sup>6</sup>In September 2001, several companies from the cocoa industry further ratified the Harkin-Engel Protocol, an international public-private partnership aimed at eradicating the worst forms of child labor in cocoa value chains in West Africa.

under Ivorian legislation.<sup>7</sup> In response to growing international demand for sustainable cocoa, as well as increasing legal requirements in both importing and producing countries, cocoa traders are constantly trying to improve the detection and remediation of child labor in their supply chains.

Since the late 2010s, the various independent cocoa certification labels (Fairtrade, RA) represent one of the potential tools for reducing child labor in the cocoa sector, now covering nearly 60% of the country's cocoa farmers (The World Bank, 2019). These labels allow certified farmers to obtain an additional premium and to benefit from various trainings<sup>8</sup>, related to good agricultural and social practices. In return, farmers must comply with social and environmental requirements, including non-reliance on child labor on their farms. This ban is associated with awareness campaigns to inform certified cocoa farmers about the harmful effects of child labor, the benefits of schooling for children, and the legal minimum age at which children can start working. Certified cocoa cooperatives must also implement systems to prevent, monitor, and address child labor at the community level. Similarly, cocoa-buying companies are required to implement a child labor monitoring system<sup>9</sup> among certified cocoa farmers in their supply chain. For this, they implement household surveys in which cocoa farmers are asked about their labor practices on their cocoa farms, including child labor. These requirements are assumed to contribute to reducing child labor on cocoa farms.

The impact of certification schemes on farmers has received increased attention from researchers in recent years. Mainly studied in the coffee sector, the economic effects are mixed and differ according to countries, types of certification, and context (de Janvry et al., 2015; Oya et al., 2018; Meemken, 2020). In the cocoa sector, studies assessing the impact of certification on cocoa farmers show a positive effect on cocoa productivity, income, and consumption (Astrid Fenger et al., 2017; Ingram et al., 2018; Knößlsdorfer et al., 2021). Nevertheless, these positive results seem conditioned by the quality of the services and the supervision that farmers receive from the cooperative to which they are affiliated. Recently, Sellare et al. (2020) confirmed the importance of accounting for the heterogeneity of cocoa cooperatives to improve the evaluation of certification schemes in Côte d'Ivoire. Mainly focused on the impacts of certification on farmers' income or productivity, the literature has paid only little attention to other outcomes likely impacted by farmers' participation in a certification program (Chen et al., 2017).

---

<sup>7</sup>Dangerous work can be allowed from the age of 16 provided that children have received specific instruction or vocational training.

<sup>8</sup>As an example, Fairtrade provides training guides and materials for farmers on sustainability, biodiversity, gender equality, and child labor, available on their dedicated website: [producerlibrary.org](http://producerlibrary.org)

<sup>9</sup>Usually based on the Child Labour Monitoring and Remediation Systems (CLMRS) developed by the International Cocoa Initiative (ICI).



Child labor, which is prohibited by all cocoa certification schemes in Côte d'Ivoire, is one of those topics on which research remains limited.

## 2.2 Social desirability bias among certified cocoa farmers

While it is prohibited by both national legislation and certification bodies in the context of certified cocoa value chains, certified cocoa farmers may still rely on child labor for a number of tasks on their farms.

In some cases, the use of child labor may be the only available option for some agricultural households. According to Cigno (2009), the poorest may not be able to afford adult labor, and others may need the additional wage earned by their children to reach a decent level of income. Empirical evidence among cocoa households in Côte d'Ivoire supports the idea that child labor is rooted in household poverty (Nkamleu and Kielland, 2006). However, previous literature also indicates that land-rich agricultural households are also likely to use child labor (Bhalotra and Heady, 2003), as each additional hectare of land increases the household's need for labor. This wealth-paradox has further been requalified as an inverted U-shaped relationship between land-wealth and child labor (Basu et al., 2010). A combination of market failures, particularly in rural areas of developing countries (Dumas, 2013), may explain this paradox. While several studies highlight the positive impact of certification on cocoa farmers' income (Astrid Fenger et al., 2017; Ingram et al., 2018), the literature on child labor thus suggests that the relationship between income and child labor is still unclear and depends on many other factors.

In Côte d'Ivoire, cocoa-producing households face a significant need for labor as cocoa-related activities are not mechanized and thus labor-intensive. While some studies have found a positive impact of certification on cocoa productivity (Ingram et al., 2018), it may nonetheless put upward pressure on farmers' need for any type of labor, including children. Moreover, while cocoa is a year-round labor-intensive activity, the labor demand peaks during the main cocoa season, when cocoa production is most intensive (Davis Pluess, 2018). Cocoa farmers are, however, usually clustered in cocoa-growing communities, which means that during peak season, farmers from the same community simultaneously seek additional labor for their farms. Considering the remoteness of cocoa communities, accessing an external labor force might be an issue for cocoa farmers. All these factors increase the likelihood that a cocoa-producing household will face difficulties finding available adult labor in a tight labor market.<sup>10</sup>

---

<sup>10</sup>The difficulty in finding adult labor was also discussed in the 2020 Cocoa Barometer.

Constraints on the labor market could, in principle, be mitigated by efficient land or credit markets. The sale or lease of land may reduce farmers' need for labor while increasing their income. However, farmers in Côte d'Ivoire continue to hold customary land titles despite successive policy attempts to facilitate the official registration of property rights.<sup>11</sup> Land exchanges are therefore limited because they are fraught with uncertainty. Access to credit can facilitate income smoothing, limiting the need to take children out of school for financial reasons, or to make them work to increase household income (Guarcello et al., 2009). However, few Ivorian cocoa farmers are affiliated with a bank or microcredit institute, and most of their loans come from informal sources (Lonie et al., 2018) and are of small amounts. The lack of school infrastructure at the community level or within a reasonable distance may also increase education costs and reduce parents' incentive to send their children to school (Fafchamps and Wahba, 2006).

All these market failures are present in both non-certified and certified cocoa value chains, suggesting that despite some directly observable income and productivity benefits of certification for cocoa farmers, their structural difficulties in accessing labor remain. Beyond economic factors explaining why certified cocoa farmers might be using child labor despite its interdiction by certification schemes, the existing monitoring difficulties might also play against the ban. Cocoa farms are often isolated and difficult to reach, making the monitoring process more costly and less effective.<sup>12</sup> Enforcement of the ban, which relies on the awareness, involvement, and cooperation of farmers and various stakeholders at the community level, is also likely to be limited. Indeed, the very definition of a child, and therefore of child labor, established by the international community is inconsistent with that of Ivorian cocoa communities. The latter does not assign tasks to children according to their age, but rather to their physical capacities, which they are able to appreciate (Buono and Babo, 2013). At the community level, it is therefore unlikely that the ILO standard would prevail, as it sometimes conflicts with social traditions of Ivorian cocoa households. The efficiency of community-based systems that detect and monitor child labor is thus probably undermined by the discrepancies between standards. Yet, previous literature on child labor has shown that the ability to enforce a policy is a key factor for its success. Consequently, where enforcement is low, policies to reduce child labor are shown to have little or no effect (Edmonds and Shrestha, 2012; Bargain and Boutin, 2021).

The literature on the determinants of child labor suggests that cocoa farmers may be using child labor on their farms, even if they belong to certified schemes. Yet, by using child labor on their certified cocoa

---

<sup>11</sup>Ministry of Agriculture and Rural Development, Declaration Of Rural Land Policy Of Cote d'Ivoire, 2017.

<sup>12</sup>Audits performed by certification bodies have recently been declared ineffective in addressing child labor (Fountain and Huetz-Adams, 2020)

farms, farmers face the risk of being excluded from certified cocoa value chains. This would imply losing all the associated benefits, including the extra income they receive from selling certified cocoa. Thus, certified cocoa farmers are likely to exhibit a social desirability bias when directly questioned about their use of child labor, especially when it involves hazardous tasks. While stakeholders in the cocoa sector conduct household or CLMRS<sup>13</sup> surveys to collect data on child labor from cocoa farmers, they only rely on direct questioning methods.

### 2.3 The list experiment

List experiments<sup>14</sup> are one of the main methods used as an alternative to direct questioning about sensitive topics, seeking to reduce respondents' social desirability bias. It consists of showing participants a list of statements (usually 4 or 5) and asking how many apply to their case. Since the answer provided by the respondent is a global response to a set of questions, the enumerator has no way of knowing precisely which statements apply to the respondent's case, which gives the latter confidentiality and should encourage him to respond more honestly. In the simplest version of the list experiment, respondents are randomly divided into two groups. The control group receives a list of  $n$  non-sensitive items, and the treatment group receives a list with the exact same  $n$  sensitive items as the control group, plus a sensitive one. The proportion of the sample engaging in sensitive behavior is then estimated by a simple difference in mean responses between the two groups.

Initially proposed by Raghavarao and Federer (1979) in a slightly different version, the list experiment has since been used in many areas to study various topics such as politics (Çarkoğlu and Aytaç, 2015; Comşa and Postelnicu, 2013; Corstange, 2010; Holbrook and Krosnick, 2010), health (Lépine et al., 2020; Moseson et al., 2015), homophobia (Lax et al., 2016; Co Man et al., 2017), intimate partner violence (Cullen, 2020) or food security (Tadesse et al., 2020) among others. The use of list experiments has revealed the existence of a social desirability bias for various sensitive topics, raising concerns about previous estimates obtained from responses to direct questioning. As an example, Gonzalez-Ocantos et al. (2012) found that 24% of Nicaraguan voters exchanged their vote for a gift or service using the list experiment approach, compared to only 2% when asked directly. Similarly, Co Man et al. (2017) found a 67% increase in the proportion of homophobic sentiment in the workplace through the list experiment in an online survey compared to the direct question. The better performance of the list experiment

---

<sup>13</sup>Child Labour Monitoring and Remediation Systems survey.

<sup>14</sup>Also known as the Unmatched Count Technique or the Item Count Technique.

compared to other methods of questioning (direct and indirect) has been highlighted by several other empirical studies (Çarkoğlu and Aytaç, 2015; Holbrook and Krosnick, 2010; Moseson et al., 2015; Wolter and Laier, 2014).

Yet, the list experiment sometimes leads to insignificant results or of the same magnitude as those obtained via the direct question (Arentoft et al., 2016; Krebs et al., 2011; Lax et al., 2016), which may occur when experiments are underpowered. Indeed, one limitation of the list experiment is that it requires a sufficiently large sample size to detect social desirability bias, as it produces estimates with relatively high variance due to the non-sensitive items. A thoughtful design of the LE is thus essential for its performance. In particular, it is recommended to prevent the use of exclusively high or low prevalence items (Blair and Imai, 2012; Glynn, 2013; Tsuchiya et al., 2007), to avoid a ceiling or floor effect issue (further explained in section 3.3.1.), and to propose negatively correlated items to reduce the variance of LE estimators (Glynn, 2013). Recently, Chuang et al. (2021) also showed that the choice of non-sensitive items could also influence respondent compliance and the results obtained with the list experiment. Focusing on sexual and reproductive behavior in Côte d'Ivoire, they show that choosing non-sensitive items related to the sensitive one implied a better LE performance than non-sensitive items fully innocuous and unrelated to the studied phenomenon. The lack of difference between the list experiment approach and direct questioning may also reveal the absence of social desirability bias, contrary to the researchers' initial assumptions (Blair et al., 2020).

Although used in many research settings, the list experiment has not yet been used to measure child labor-related behavior.

### **3. Research design**

#### **3.1 The data**

This paper uses data from a socioeconomic survey, designed by the author and carried out by an international cocoa trading company, among Ivorian certified cocoa households that are part of its supply chain. Two questionnaire versions are used to implement a list experiment.<sup>15</sup> The survey is divided into two complementary questionnaires: (i) A household-level questionnaire, which collects

---

<sup>15</sup>Randomization of the survey version among farmers was also carried out and monitored by the author.

household characteristics such as size, composition, and basic characteristics of its members, along with information on banking, savings and credit practices, and wealth. (ii) A farm-level questionnaire which covers issues such as farm size, farm age, pesticide use, the incidence of pests and diseases, use of family and hired labor, and cocoa-related income. The survey is only addressed to the certified cocoa farmer, who is the head of the household in most cases. No other respondents, such as other household members, were surveyed. Data collection started during the main cropping season in December 2019 and stopped in mid-March 2020. The enumerators collected information in several cocoa-growing areas in the regions of Agnéby-Tiassa, Cavally, Grands Ponts, Guémon, Côh, Lôh-Djiboua, Nawa, and Tonkpi. All enumerators were fluent in French but were also proficient in local languages<sup>16</sup>, making it possible to survey non-French-speaking farmers. The translation of the list experiment questions and proposed items into different local languages was discussed and defined during the enumerators' training.

The final sample consisted of 4 458 UTZ<sup>17</sup> (88%) and RA (12%) certified cocoa farmers<sup>18</sup>, randomly selected among sampled cooperatives.<sup>19</sup> The great majority of them are Ivorian (74%) or Burkinabe (24.5%), and almost all of the farmers are men (95%) (see Table 5. in Appendix). Cocoa farmers are on average 42 years old, generally married (89%), and live in a household with at least one child (90%). The level of farmers' education remains relatively low: 57% of them never attended school, and only 17% of them have reached secondary or higher education. The median size of the cocoa farm is 2.09 ha, but cocoa farmers also grow other crops on their farm in 51% of cases.

### 3.2 Design and implementation of the list experiment

The implemented list experiments focus on three different categories of work performed by children under 16 on certified cocoa farms. First, this study aims to estimate the proportion of cocoa farmers being helped by a child for tasks classified as hazardous by the ILO. The focus is on hazardous work related to farm preparation (applying pesticides, clearing, tree felling, burning plots), and harvesting and pod breaking. Second, we seek to estimate the proportion of cocoa farmers who hire and pay a child to

---

<sup>16</sup>Such as Baule, Dioula, or Senoufo.

<sup>17</sup>Although the survey ended before reaching the target sample size due to the COVID-19 outbreak, post-hoc power calculations show that our sample allows us to detect a minimum detectable effect of 8%.

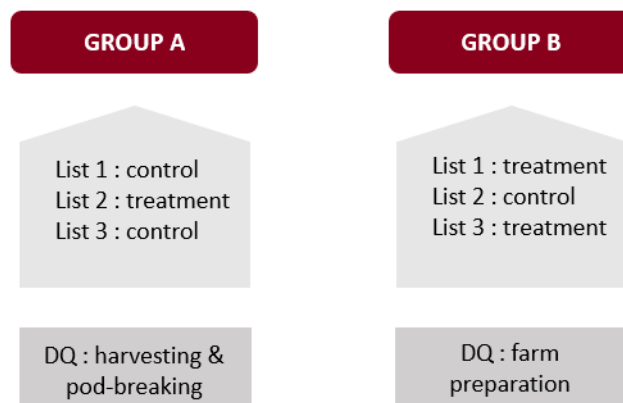
<sup>18</sup>Farmers in our sample benefited from several awareness sessions on the negative effects of child labor. Each cocoa farmer and their household are individually sensitized to child labor when they start working for a certified cooperative. A second awareness session, this time at the community level, is conducted every year.

<sup>19</sup>The cooperatives sampled were all certified cooperatives that were part of the supply chain of the cocoa trading company with whom the author worked with.

perform any task(s) on their farm. The light work children perform on cocoa farms is not considered here, as this is not harmful to the child and is commonly encouraged by their family. Those tasks, tolerated by the ILO, are generally perceived by cocoa farming households as an integral part of the child's education and personal development (Buono and Babo, 2013). The scope of this study is also restricted to children under 16, in line with Ivorian legislation which, under certain conditions, allows children between 16 and 18 to perform hazardous work.<sup>20</sup> Thus, this study focuses on the worst forms of child labor.

Three list experiments were proposed to cocoa farmers, one for each of the aforementioned categories of work (Table 1). Cocoa farmers were randomly assigned to two groups (Figure 1). Group A serves as a control group for list experiments 1 and 3, *i.e.*, farmers in group A are not exposed to the sensitive item in lists 1 and 3; and as a treatment group for list 2. The inverse is true for group B, which serves as a control group for list experiment 2 and treatment group for list experiments 1 and 3. The list experiments were always administered to respondents in the same order, regardless of the respondent group. Section 4.2.1 discusses whether and how this choice may have affected the reliability of the LE estimates.

**Figure 1:** Sequencing of the list experiments and direct questions



Source: Author.

**Note:** DQ stands for direct question. Farmers in group A are asked about the three list experiments. Since they belong to the control group for lists 1 and 3, these two lists do not contain the sensitive item. They are asked directly about child labor for harvesting and breaking the pods.

The non-sensitive items used in the list experiments are related to the sensitive item's topic, *i.e.*, the farmer's use of labor on his certified cocoa farm. We argue that including non-sensitive items unrelated to the sensitive one would only increase its salience to respondents, which is in line with findings from

<sup>20</sup>See decree No. 2017-017 MEPS/CAB of June 2, 2017, determining the list of hazardous work prohibited for children.

**Table 1:** List experiments and direct questions

Wording of the questions	
Harvesting & pod-breaking	<p><b>List Experiment #1</b></p> <p>During the last campaign, for harvesting and breaking the pods:</p> <ol style="list-style-type: none"> <li>1) No woman over 16 helped me.</li> <li>2) At least one member of my family helped me.</li> <li><b>3) One or more people under 16 helped me.</b></li> <li>4) A group of communal laborers helped me.</li> </ol> <p>In total, how many statements apply to your case ?</p>
	<p><b>Direct Question</b></p> <p>During the last campaign, for harvesting and breaking the pods, did one or more people under 16 help you ?</p>
	<hr/>
Farm Preparation	<p><b>List Experiment #2</b></p> <p>During the last campaign, to prepare my plot (applying pesticides, clearing, felling of trees, burning of plots):</p> <ol style="list-style-type: none"> <li>1) No woman over 16 helped me.</li> <li>2) At least one member of my family helped me.</li> <li><b>3) One or more people under 16 helped me.</b></li> <li>4) A group of communal laborers helped me.</li> </ol> <p>In total, how many statements apply to your case ?</p>
	<p><b>Direct Question</b></p> <p>During the last campaign, to prepare your plot (applying pesticides, clearing, felling of trees, burning of plots), did one or more people under 16 help you ?</p>
	<hr/>
Employed child labor	<p><b>List Experiment #3</b></p> <p>During the last campaign, to do some work on my farm (for all types of activities):</p> <ol style="list-style-type: none"> <li>1) I have not employed any woman over 16.</li> <li>2) At least one member of my family helped me.</li> <li><b>3) I have employed and paid one or more people under 16 who are not members of my family.</b></li> <li>4) I employed a group of communal laborers.</li> </ol> <p>In total, how many statements apply to your case ?</p>
	<p><b>No Direct Question</b></p>
	<hr/>

Source: Author.

**Note:** For each list experiment, the bolded item is the sensitive one, related to child labor. It is only read to cocoa farmers who belong to the treatment group for that list experiment.

Chuang et al. (2021). In their study about LEs and sexual practices in Côte d'Ivoire, they found that the LE performed worse when the control items were non-sensitive or unrelated to the sensitive item. The LEs used in this study are hence focusing on different types of labor that cocoa farmers may use on their farms. Thus, enumerators ask cocoa farmers about their overall labor practices, rather than specifically about their use of child labor. The non-sensitive items also relate to types of labor that cocoa farmers are typically asked about in questionnaires administered by different stakeholders in the sector.<sup>21</sup> Although the LEs also mention child labor and may still lead to farmers' mistrust, it is expected that LE's format and wording, as well as the proposed control items, help reduce the salience of the sensitive item (*i.e.*, the use of child labor), as it is framed as one of several types of labor possibly used by cocoa farmers.

The non-sensitive items were furthermore designed following the recommendations provided by the list experiment literature. Two negatively correlated items (items 1 and 2) were chosen to minimize the variance of the LE estimator, as recommended by Glynn (2013).<sup>22</sup> This also prevents a treated farmer from answering "yes" or "no" to all items, thus revealing one's answer to the sensitive item (Tsuchiya et al., 2007). The non-sensitive items used in the LEs were also slightly modified from those designed initially, based on feedback from enumerators who had the opportunity to test the LEs during a one-week test with certified farmers. The modifications resulted in more relevant and understandable non-sensitive items for cocoa farmers.

Before proceeding with the list experiment, each enumerator explained to the respondent how this particular question worked, as follows:

*"Please put your hand behind your back. I am going to read 3 or 4 statements about the labor used on your cocoa farm in the last 12 months. If the statement read is true, and applies to your case, lift a finger behind your back. If the statement is not true, do not lift a finger. After I read you all the statements, show me your hand and the number of fingers you raised. I do not want to know which statements apply to you, only the number of statements that are true for you. Please do not answer "yes" or "no" when I read statements to you. Your answers must remain confidential."*

Enumerators were further encouraged to provide an example to the farmer in case he seemed confused about the question's wording. The proposed example was a simple one, using familiar terms and

<sup>21</sup>Certified cocoa farmers are subjected to different types of surveys including modules on their labor use, such as CLMRS surveys conducted by cocoa traders or surveys conducted by researchers (NORC survey).

<sup>22</sup>Items 1 and 2 are, indeed, negatively correlated. In most cases, the farmer is at least helped by his wife on his cocoa farm. Based on the information collected about household members in the survey, we can also state that the wives of the sampled farmers are also older than 16 years old.



representing a common situation for farmers. Acknowledging that the list experiment may involve a higher cognitive load than a direct question, farmers were not asked to count the number of statements that applied to their case, but only to raise a finger behind their backs for each positive response to an item. Adding up the number of positive answers given to each item was left to the enumerators.

In addition to the list experiments, farmers from group A were asked directly about the use of child labor for harvesting and pod breaking, and farmers from group B were asked directly about their reliance on children for plot preparation. To be able to compare the answers given to the list experiment and the direct question, the wording of the direct questions was identical to that of the corresponding list experiment (Table 1). No direct question was asked about employed child labor. During the author's fieldwork, this issue was found to be far too sensitive and could have caused farmers to refuse to continue the survey, as the employment of children is sometimes associated with child trafficking and forced labor.

The design and sequencing of the direct questions were used to minimize potential response bias from farmers. First, farmers in each group were asked directly about only one type of child labor, as asking several direct sensitive questions might raise their suspicions or be perceived as offensive, which could distort the answers given to the second direct question and lead to downward bias. Second, farmers assigned to the treatment group for one type of child labor with the LE were not asked directly about that same sensitive topic, thus minimizing the potential salience effect and downward bias affecting the response to the direct question.

### 3.3 Estimation strategy

A difference-in-means estimator is used to obtain the estimated prevalence of use of the three types of child labor mentioned above. This estimator is used to obtain the average number of items to which respondents answered "yes" in both the control and treatment groups. Any difference in means between the treatment and control groups is attributed to the sensitive item. Following Blair and Imai (2012), the estimator is defined as follows:

$$\hat{\tau} = \frac{1}{N_1} \sum_{i=1}^N T_i Y_i - \frac{1}{N_0} \sum_{i=1}^N (1 - T_i) Y_i \quad (1.1)$$

With  $T_i$  taking the value of 0 if the respondent is in the control group, 1 otherwise.  $Y_i$  is the answer given by the respondent, which can take the value from 0 to  $J$ ,  $J$  being the maximum possible value, *i.e.* the total number of proposed items.  $N_1$  and  $N_0$  correspond to the size of the treatment and control group, respectively.

The validity of the difference-in-means estimator relies on three key assumptions (Blair and Imai, 2012): a random allocation of respondents between the control group and treatment group (1), the assumption of no design effect (2), and the assumption of no liars (3). If these three assumptions are not simultaneously satisfied, the estimator does not yield valid estimates.

Assumption 1 cannot be rejected, as farmers who received questionnaires A and B display similar characteristics, supporting the successful random allocation of treatment between the two farmer groups (Table 2).

**Table 2:** Basic characteristics of respondents in groups A and B

	Control	Treatment	(C-T) p-value	Obs.
age	41.99	41.88	0.749	4 458
married	0.89	0.89	0.951	4 458
no education	0.57	0.56	0.289	4 458
ivorian	0.73	0.74	0.947	4 458
number of children	2.62	2.61	0.726	4 458
number of adults	3.29	3.38	0.117	4 458
cocoa income (XOF)	995 001	992 207	0.913	4 458
yields (kg/ha)	656.57	639.16	0.392	4 458
other crop	0.51	0.51	0.911	4 458

Source: Author's calculation.

Assumption 2 implies that the inclusion of the sensitive item does not affect the respondents' answers to control items. That is, the treatment group's answers to non-sensitive items should be approximately equivalent to those of the control group. Additionally, the average total number of items in the treatment group must be equal to or greater than that of the control group, but at most by one item. This condition corresponds to the null hypothesis of the test proposed by Blair and Imai (2012) to detect the presence of a design effect. Keeping their initial notations, we have:

$$H_0 : \begin{cases} \Pr(Y_i \leq y | T_i = 0) \geq \Pr(Y_i \leq y | T_i = 1) & \text{for all } y = 0, \dots, J - 1 \text{ and} \\ \Pr(Y_i \leq y | T_i = 1) \geq \Pr(Y_i \leq y - 1 | T_i = 0) & \text{for all } y = 1, \dots, J \end{cases} \quad (1.2)$$

This test yields estimates of the different proportions of each type of respondent ( $\hat{\pi}_{yz}$ ), characterized by the total number of affirmative answers to the non-sensitive items ( $y$ , which can take a value from 0 to 3) and their answer for the sensitive item ( $x$ , which takes the value of 0 for a negative answer, and 1 for a positive one). If all these proportions are positive, the null hypothesis of no design effect cannot be rejected (Blair and Imai, 2012). This is the case for the three list experiments, as shown in Table 3.

**Table 3:** No-design effect test results

y value	Harvesting & pod-breaking				Farm preparation				Employed child labor			
	$\hat{\pi}_{y0}$	SE	$\hat{\pi}_{y1}$	SE	$\hat{\pi}_{y0}$	SE	$\hat{\pi}_{y1}$	SE	$\hat{\pi}_{y0}$	SE	$\hat{\pi}_{y1}$	SE
0	2.41	0.003	1.01	0.005	2.91	0.003	0.11	0.005	2.74	0.003	0.69	0.005
1	24.75	0.010	4.81	0.014	27.93	0.010	3.73	0.014	25.73	0.010	0.92	0.014
2	44.50	0.013	8.37	0.012	39.28	0.014	5.89	0.013	40.35	0.014	8.61	0.013
3	4.21	0.010	9.94	0.006	8.13	0.011	12.02	0.007	5.55	0.011	15.41	0.008

Source: Author's calculation.

**Note:** The table shows the estimated proportion of respondents characterized by the total number of affirmative answers to the non-sensitive items ( $y$ ) and their answer for the sensitive item (1 stands for a positive response to the sensitive item, 0 for a negative one). These proportions are estimated using the R package proposed by Blair and Imai (2012). *For example:* The value in the fourth row and third column suggests that an estimated 9.94% of farmers answered “no” to all non-sensitive items but “yes” to the sensitive item. Standard errors are also provided for each estimated proportion. We use Blair and Imai's (2012) notations.

Finally, assumption 3 implies that respondents answer honestly to the sensitive item. While there is no statistical test to verify this hypothesis, the LE approach should, by design, enhance honest responses from respondents to the sensitive item. Nevertheless, the validity of this assumption may be questioned when the respondent would like to answer “yes” (ceiling effect) or “no” (floor effect) to all the items. In this case, the enumerator can infer the respondent's answer to the sensitive item, which removes the private nature of the answer and may encourage respondents to misreport their true answer. Such concealment strategy, known as strategic measurement errors, calls for specific care to avoid floor or ceiling effects when designing a LE. In particular, researchers should avoid proposing lists of high-prevalence or low-prevalence items (Blair and Imai, 2012; Glynn, 2013; Tsuchiya et al., 2007), as it increases the likelihood that a respondent will answer “yes” or “no” to all the items. Glynn (2013) advises proposing negatively correlated non-sensitive items, which has the added advantage of reducing the variance of the difference-in-means estimator, which is known to be less powered than the direct method. For each list experiment, these recommendations were followed in the choice of non-sensitive items (Table 1). Yet, it is likely that a ceiling effect occurred, with 15% to 20% of respondents answering “yes” to all items in the control group (Table 4). Thus, it is possible that some farmers from the treatment group who used child labor, but who had already answered “yes” to all the control items, did

not answer truthfully to the sensitive one. This suggests that the difference-in-means estimates are to be interpreted as lower bounds of the true prevalence of child labor use.

**Table 4:** Distribution of farmers according to their answer to the list experiments

y value	Harvesting & pod-breaking				Farm preparation				Employed child labor			
	Control		Treatment		Control		Treatment		Control		Treatment	
	Freq.	(%)	Freq.	(%)	Freq.	(%)	Freq.	(%)	Freq.	(%)	Freq.	(%)
0	79	3.43	52	2.41	65	3.02	67	2.91	79	3.43	59	2.74
1	681	29.56	555	25.77	682	31.66	646	28.04	614	26.65	569	26.42
2	1 218	52.86	1 062	49.30	973	45.17	991	43.01	1 128	48.96	889	41.27
3	326	14.15	271	12.58	434	20.15	323	14.02	483	20.96	305	14.16
4			214	9.94			277	12.02			332	15.41
Total	2 304		2 154		2 154		2 304		2 304		2 154	

Source: Author's calculation.

**Note:** The table shows the distribution of farmers based on the total number of items to which they answered yes. For example: 52.86% of farmers in the control group answered affirmatively to 2 items for harvesting and pod-breaking, compared to 49.30% in the treatment group.

## 4. Results

### 4.1 Differences between direct questioning and LE responses

The results of the list experiment (column 3), the direct question (column 4), and the difference between the two measures (column 5) for each type of child labor are reported in Table 5. The difference-in-means between the treatment and control groups is positive, statistically significant, and relatively large for all three types of child labor. The findings suggest that the use of child labor is not an isolated practice among certified cocoa farmers: 24.1% of the farmers in the study sample used child labor for harvesting and breaking the cocoa pods during the past 12 months, 21.7% for farm preparation, and 25.6% of them employed and paid at least one child under 16 to perform any task on their cocoa farm.

Column 4 of Table 5 reports the prevalence of farmers who responded positively to the direct question regarding child labor for harvesting and pod-breaking, and farm preparation. When cocoa farmers are asked directly, 12.76% of them report having been helped by a child under 16 for harvesting and breaking the pods in the last 12 months, and 18% of them declared that they used child labor for farm preparation. These results are half as high as those obtained using the list experiment, supporting the

**Table 5:** Proportion of farmers using child labor estimated by list experiment and direct questioning method

	(1) Control mean	(2) Treatment mean	(3) Difference- in- means (%)	(4) Direct question (%)	(5) Difference (pp)
List 1: harvesting & pod-breaking	1.777	2.018	24.12	12.76	11.36***
N	2 304	2 154	4 458	2 304	
List 2: preparing the plot	1.824	2.042	21.75	11.28	10.47***
N	2 154	2 304	4 458	2 154	
List 3: employed child labor	1.874	2.130	25.63		
N	2 304	2 154	4 458		

Source: Author's calculation.

Note: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

existence of a social desirability bias among certified cocoa farmers when directly asked about their use of child labor to perform hazardous tasks. Although farmers were not asked directly about hiring and paying children under 16 to perform any task on their farm, it is likely that this work category is prone to similar social desirability bias as hazardous child labor. Indeed, employing and remunerating a child under 16, regardless of the work performed, is legally prohibited under Ivorian law.<sup>23</sup> This information is also common knowledge in cocoa farming communities that benefit from child labor awareness-raising campaigns.

The difference of over 10 percentage points between the direct method and the list experiment indicates that at least half of the farmers who do use child labor are not willing to declare it (column 5). These results are in line with the initial hypothesis about the existence of a social desirability bias. From the author's field observations, these findings are interpreted as a rational response to farmers' double pressure regarding hazardous child labor, which is both prohibited in Côte d'Ivoire and by the certification schemes.

<sup>23</sup>Article 23.2 of the Labour Code.

## 4.2 Robustness checks

Estimating the proportion of certified cocoa farmers relying on child labor is challenging as it is prone to measurement errors, as discussed in Section 4.2.1. Sections 4.2.2 and 4.2.3 test for the robustness of our LE estimates to design changes and the extent to which the LE method is understood by farmers.

### 4.2.1 Potential source of measurement errors

While respondent type may impact answers given to child labor-related questions (Bardasi et al., 2011), this possibility is ruled out in this setting. Since the survey was administered to certified cocoa farmers, the risk of getting inaccurate reporting associated with questioning proxy respondents is removed.

Memory bias might be present due to the length of our recall period, which refers to the last cocoa season, *i.e.*, the past 12 months. Although such a long recall period may increase the likelihood of memory bias, this risk is lowered for two reasons. First, the annual cocoa season is divided into two 6-month periods: the main crop, when cocoa production peaks, and the light crop, when cocoa production is lower. These two seasons were discussed at length by the enumerators throughout the survey, facilitating the recovery of farmers' memories, including those related to their use of labor. Second, while the literature on recall bias associated with the measurement of farm labor shows that recall period length matters, whether a longer recall period overestimates or underestimates the phenomenon is still unclear, depending on the context (Gaddis et al., 2021; Arthi et al., 2018). Further, it is likely that measuring the occurrence of a binary behavior (whether or not the farmer used child labor on his farm) decreases the likelihood of recall bias compared to other more complex measures of farm labor, such as hours worked or the number of workers. Last, the fact that the same recall period was used for both the direct question and the list experiment suggests that if any memory bias occurs, it should affect both measures in the same way.

The use of the list experiment is expected to reduce social desirability bias, which remains an important issue associated with estimating child labor-related measures. Yet, this method itself may be prone to measurement errors, mainly because the LE is considered more cognitively demanding for respondents than a direct question. People with cognitive difficulties or low levels of education may struggle to understand this indirect questioning method, which may lead them to give a random answer to the enumerator. As a result, measurement errors would correlate with respondents' education levels. To

mitigate this issue, farmers were only asked to lift a finger behind their backs when a LE statement applied to their case, thus removing the need for them to count and aggregate their answers. Section 4.2.3 reports the results of a comprehension test of the LE from another survey also carried out with Ivorian certified cocoa farmers, which supports a good level of understanding of the LE among this population, even for the less educated farmers.

One might also be concerned about any measurement issue arising from the fact that not all of the items composing the LEs are mutually exclusive. Thus, control items 2 and 4, which are used in the three list experiments and asked to all farmers, whether in the control or treatment group, can overlap. Similarly, farmers belonging to the treatment group for LEs 1 (harvesting and pod-breaking) and 2 (farm preparation) receive items 2 and 3, which can also overlap if one was helped by a child under 16 who also belongs to one's family. It is therefore possible that a farmer might answer "yes" to both control item 2 and sensitive item 3 when considering the same person (a child of his family under 16). This could be an issue under the following two conditions: 1) farmers who were helped by a child under 16 belonging to their family have different response patterns to item 2, *i.e.*, in some cases, they answer "yes" to item 2, while in other cases they do not, 2) these different response patterns are associated with treatment or control group assignment. As the farmers in the control and treatment groups have socio-economic characteristics that are not statistically different from each other on average (Table 2), their average response to item 2 should also be similar, thus not impacting the difference-in-means between the two groups. The same reasoning holds for the overlap of control items 2 and 4.

Respondents' answers to LEs may also be influenced by the order in which they are presented, as each farmer is administered the three list experiments. To avoid any order effect, cocoa farmers were first presented with the list experiment related to tasks that are *a priori* the most often performed by children on cocoa farms (Sadhu et al., 2020). The last list experiment asks about the most sensitive practice (*i.e.*, hiring children outside one's family). This sequence was designed to avoid offending the respondent by asking first a question on the most sensitive topic. The statements of the list experiments were also neutral and did not contain any judgments or terms that could offend the farmers. We did not use the term "children" or describe the tasks we were interested in as "hazardous". Despite these precautions, an order effect may persist, but it is likely to be limited. It is unlikely that cocoa farmers in Group A experience an order effect, as they are only exposed to one list experiment as a treatment group, implying that they are only presented with the sensitive item once. If any order effect occurred, it would likely be among farmers belonging to group B, who act as treatment groups for two LEs, thus being

exposed to the sensitive item twice, possibly increasing its salience. This may lead them to give less truthful answers to the sensitive item of the second list experiment for which they serve as a treatment group. Acknowledging that an order effect may exist in that case, we consider that, if anything, LE estimates on children's employment (list 3) are probably a lower bound, as it may have been affected by such a downward effect.

#### 4.2.2 Sensitivity of LE estimates to design changes

Two new survey rounds were conducted between December 2021 and May 2022 among certified cocoa farmers, similar to 2019. The farmers surveyed in these new survey rounds are either (1) farmers who were not surveyed in the 2019 survey due to the unexpected end of data collection due to the COVID-19 outbreak, or (2) new farmers who entered the cocoa trading company's supply chain after the 2019 survey. These new survey waves involved different respondents than in 2019, so they cannot be considered panel data. They can nevertheless offer insight into the sensitivity of the LE to design modifications. The different LE designs used in each survey wave are summarized in Table 6.

In the first survey wave of 2022 (referred to as "2022-V1" hereafter), changes were made to the LEs used in 2019, both to the design of the non-sensitive items, making them all mutually exclusive, and to the recall period, which was shortened from 12 to 6 months. Two LE comprehension tests preceding the three real LEs on child labor were also included. These modifications were made across the entire sample, making it impossible to determine a specific effect related to each design change, but rather an overall effect. The second wave of 2022 (2022-V2) restored the same LE characteristics as in the 2019 survey.

The two additional datasets successfully pass the balance tests for farmer characteristics between the control and treatment groups (Tables A2 & A3 in Appendix), thus implying successful randomization of the LEs across farmers.

Specifically, this analysis aims to test whether the changes in LE design are associated with different response behaviors for farmers belonging to the treatment group for LEs. If so, the difference-in-means between the treatment and control groups would reflect not only a social desirability bias, but also measurement errors related to LE design features.

We rely on the following pooled regression for each type of child labor:

$$Y_i = \alpha + \beta_0 T_i + \beta_1 Year + \beta_2 Year * T_i + \beta_3 Design + \beta_4 Design * T_i + \delta_i + \gamma_i + \epsilon_i \quad (1.3)$$



**Table 6:** Survey characteristics across waves

	2019	2022 - V1	2022 - V2
<i>Survey characteristics</i>			
<b>Sampled individuals</b>	Certified cocoa farmers	Certified cocoa farmers	Certified cocoa farmers
<b>Sample size</b>	4 458	7 447	5 593
<b>Sampled regions</b>	Agnéby-Tiassa Cavally Gôh Grands Ponts Guémon Lôh-Djiboua Nawa	Agnéby-Tiassa Cavally Gôh Grands Ponts Guémon Haut-Sassandra La Mé Lôh-Djiboua Sud-Comoé	Agnéby-Tiassa Cavally Gôh Guémon Haut-Sassandra La mé Lôh-Djiboua Sud-Comoé San-Pédro
<b>Data collection period</b>	December 2019 - March 2020	December 2021 - February 2022	February 2022 - May 2022
<i>List Experiment (LE) design</i>			
<b>Non-sensitive items</b>	Initial	Modified	Initial
<b>LE test</b>	No test	2 tests before LE	1 test after LE
<b>Recall period</b>	12 months	6 months	12 months

Source: Author.

where  $Y_i$  is the farmer's response to the LE question, *i.e.*, the number of items to which the individual  $i$  answered "yes" (ranging from 0 to 3 for farmers in the control group, and 0 to 4 for farmers in the treatment group).  $T_i$  is a dummy indicating whether the cocoa farmer was assigned to the control or treatment group,  $Year$  is a dummy equal to 1 if the survey took place in 2022, 0 if it was conducted in 2019,  $Design$  is a dummy equal to 1 if the LE design is that of 2022-V1, 0 for the 2019 and 2022-V2 version.  $\delta_i$  stands for a vector of household control variables (including farmer's gender, age, education, and cocoa income), and  $\gamma_i$  for enumerators' fixed effects. Standard errors are clustered at the regional level.

The results from these regressions are reported in Table 7 for the three types of child labor, using the overall sample (2019, 2022-V1, and 2022-V2).

**Table 7:** Regression analysis of list experiments (LEs) across survey waves

	Harvesting & pod-breaking	Preparing the farm	Employed child labor
Treatment group	0.210*** (0.027)	0.209*** (0.046)	0.238*** (0.041)
2022	-0.099 (0.075)	-0.040 (0.122)	-0.023 (0.093)
2022 x Treatment group	-0.186** (0.045)	-0.232** (0.055)	-0.197*** (0.024)
LE design 2022-V1	-0.680*** (0.121)	-0.721*** (0.122)	-0.769*** (0.110)
LE design 2022-V1 x Treatment group	0.014 (0.026)	0.057* (0.023)	-0.014 (0.022)
Constant	1.809*** (0.099)	1.693*** (0.087)	1.689*** (0.117)
N	13 211	13 165	13 068
R-squared	0.447	0.453	0.473
Controls	Yes	Yes	Yes

Source: Author's calculation.

**Note:** Clustered robust standard errors at the regional level are shown in parentheses. Level of significance: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . Included controls: enumerators fixed effects and household characteristics (farmer's gender, age, education, and cocoa income).

Findings show that being surveyed with LEs during the main cocoa season in 2022 versus 2019 has no significant average effect on the number of items farmers answer "yes" to, but it does for the farmers in the treatment group. Since the only difference between the control and treatment groups is the sensitive item related to child labor use, two assumptions may explain this result. Either there have been

structural factors leading to a decline in child labor use between 2019 and 2022, or differences in training and implementation of the LE may have impacted how farmers answered to the LE. The descriptive statistics reported in Table A4 suggest a reduction in the use of child labor between 2019 and 2022 when looking at answers to direct questions. Although these results should be interpreted with caution, as they do not come from panel data and rely on direct questioning, they provide some indication of the overall trend. However, this decline in the use of child labor is not associated with a decrease in the use of family or hired labor.

Changes in the design of the list experiment (reduction of the recall period to 6 months, addition of LE comprehension tests, and modification of non-sensitive items to make them mutually exclusive) have a large negative average effect on the number of items to which farmers respond “yes”. As expected, this result is driven by the reduction of the recall period from 12 to 6 months, which mechanically led to a downward effect. This finding is also likely driven by the fact that modifying non-sensitive items to make them mutually exclusive implied adding more restrictive items (Tables A5 & A6 in Appendix). Descriptive statistics show that the distribution of the number of items to which farmers answered “yes” is thus much more concentrated than with the original design (2019, and 2022-V2), in both the control and treatment groups (Tables A7, A8, & A9 in Appendix). However, the interaction coefficient between the treatment variable and the LE design version indicates that changes in LE design have the same effect on control and treatment groups for the LEs related to harvesting and pod-breaking, and employed and paid child labor. The reason why changes in LE design are associated with a higher number of items to which the farmers answer “yes” in the treatment group for child labor related to farm preparation remains, however, unclear.

Overall, findings provide evidence that changes in LE design (length of recall period, non-sensitive items, and addition of a comprehension test before LEs) are strongly associated with a lower average number of items to which farmers respond “yes”. This is likely driven by the reduction of the recall period and the addition of more restrictive non-sensitive items. However, this downward trend in LE responses is similarly observed across farmers in both the control and treatment groups. That is, while changes in the LE design are associated with a reduction in the number of average items to which farmers respond “yes”, this reduction is uniform between the control and treatment groups, thus not affecting the difference-in-means between the two groups, *i.e.*, the LE estimates.

Another insightful finding is that using the LE across years to measure the same sensitive behavior can lead to different results. In this study setting, it is unclear why being surveyed in 2022 versus 2019 is associated with a reduction in positive answers provided to the LEs only among farmers in the treatment group. This may be driven by a structural downward trend in the use of child labor, or differences in LE survey training and implementation across the years.

#### 4.2.3 List experiment and cognitive load

The inclusion of two new survey waves also provided an opportunity to test the understanding of the LE method on a similar sample as in 2019, *i.e.*, certified cocoa farmers. For this purpose, the new questionnaire included either two (2022-V1) or one (2022-V2) comprehension test for the LE. Tests are conducted as follows: a list experiment is used to ask cocoa farmers about a non-sensitive topic, followed by direct questions for each item used in the LE test. The number of items the farmer answered “yes” to when asked the LE, was compared to the number of direct questions to which the farmer answered “yes”. If both numbers are equal, then it is considered that the cocoa farmer has understood the LE method correctly. If answers differ, it is assumed that the farmer misunderstood the LE or made a mistake while answering one of the two questioning methods. The wording and proposed items for each of the LE comprehension tests are presented in Appendix (Table A10).

The test results indicate a high rate of understanding of the LE method by the cocoa farmers, around 97% (Table 8). Although these tests were not conducted during the 2019 survey, they provide evidence of farmers’ ability to understand the LE method, even though the cognitive load is greater than that of direct questions. Descriptive statistics also show no particular difference in LE comprehension between cocoa farmers who have never been to school and those with primary education or higher.

## 5. Conclusions

Eradicating child labor remains a major challenge for both public and private sectors. To provide effective and well-targeted policies, as well as to monitor progress and evaluate the effectiveness of interventions aimed at reducing child labor use, it is essential to measure this phenomenon accurately. Yet, measuring child labor use is challenging given its sensitive nature: misreporting the use of child labor is likely common, given the social, economic, and legal sanctions that one who relies on child labor can be

**Table 8:** Proportion of farmers who passed the list experiment comprehension test

	1st test		2nd test	
	(1) %	(2) N	(3) %	(4) N
Overall sample	96.96	13 040	97.31	7 447
<b>By education level</b>				
No education	96.83	7 727	97.38	4 737
Primary education or higher	97.16	5 308	97.19	2 706

*Source:* Author's calculation.

**Note:** Column 1 reports the success rate of the 1st list experiment comprehension test, administered in the 2022-V1 and 2022-V2 surveys. Column 3 reports the success rate of the 2nd list experiment comprehension test, which was only administered in the 2022-V1 survey.

exposed to. This paper relies on the list experiment approach, implemented among 4 458 Ivorian cocoa households to estimate the proportion of farmers using child labor on cocoa farms certified as child labor free. This indirect survey method is used to overcome the social desirability bias associated with direct questions related to hazardous child labor, which is prohibited both by national legislation and cocoa certification schemes. Two additional survey rounds are used to test the robustness of the estimates to LE design changes, and assess farmers' understanding of this indirect questioning method.

LE estimates show that 24% of certified cocoa farmers were helped by at least one child under 16 for harvesting and breaking the cocoa pods during the 2019 cocoa season, 21% for preparing their farm, and 25% employed and paid a child to perform any task on their cocoa farm. These results are twice as high as those declared by farmers when directly questioning them about their child labor use, indicating a social desirability bias among cocoa farmers when asked about such a sensitive topic. Through the implementation of a comprehension test of the LE method, it is shown that despite its higher cognitive load than direct questioning, the LE is well understood by farmers. Findings further provide evidence that, in this setting, if changes in LE design are associated with a lower average number of LE items to which farmers respond "yes", this reduction is uniform between the control and treatment groups, thus not affecting the LE estimates.

These paper's results are of interest for three reasons. First, this study highlights an additional source of error associated with measuring child labor use that has so far received little attention in the existing literature. The findings support the existence of a social desirability bias associated with directly questioning certified cocoa farmers about their use of child labor on their farms, in a context where hazardous child labor is prohibited by both the government and the certifying schemes. The results

suggest that direct questioning leads to underestimating the prevalence of child labor use by at least 50%.

Second, findings of this study highlight that in Côte d'Ivoire, certified cocoa farmers are using child labor despite the ban imposed by the government and certification schemes, even to perform dangerous tasks on their farms. This raises questions about the effectiveness of certification schemes to reduce child labor in the cocoa sector, and the monitoring processes implemented by certification bodies.

Last, results from the robustness tests suggest that the LE method, which is cognitively more demanding than direct questioning, can be understood even by populations with low levels of education, if implemented in a way that lowers issues of misunderstanding. In this study setting, respondents were not required to remember or add up their positive responses to the LE items. Instead, they only had to raise a finger behind their back when an item applied to their case, and show their hand to the enumerator at the end of the LE question. Adding up the respondent's positive answers was the sole responsibility of the enumerator.

Several caveats are in order. First, while direct questioning provides a lower bound on the use of child labor at an individual level, indirect questioning methods can provide a broader identification of farmers using child labor at an aggregate level. Nevertheless, this method is probably more appropriate for measuring a simple binary phenomenon, *i.e.*, whether a farmer relies on child labor, than for more complex measures of child labor use, such as the number of hours worked by children, due to a higher cognitive load than the direct questioning method. Second, we were not able to further validate the measures obtained through indirect questioning methods with objective measures of child labor. The inability to do so represents the main limitation of this study. Such an exercise would make it possible to assess the remaining gap between the true value of child labor use and the estimates obtained using the LE approach.

## References

- Arentoft, A., Van Dyk, K., Thames, A. D., Sayegh, P., Thaler, N., Schonfeld, D., Labrie, J., and Hinkin, C. H. (2016). Comparing the unmatched count technique and direct self-report for sensitive health-risk behaviors in HIV+ adults. *AIDS Care - Psychological and Socio-Medical Aspects of AIDS/HIV*, 28(3):370–375. [23]
- Arthi, V., Beegle, K., De Weerd, J., and Palacios-López, A. (2018). *Not your average job: Measuring farm labor in Tanzania*, volume 130. The World Bank. [15], [33]
- Astrid Fenger, N., Skovmand Bosselmann, A., Asare, R., and de Neergaard, A. (2017). The impact of certification on the natural and financial capitals of Ghanaian cocoa farmers. *Agroecology and Sustainable Food Systems*, 41(2):143–166. [19], [20]
- Bardasi, E., Beegle, K., Dillon, A., and Serneels, P. (2011). Do labor statistics depend on how and to whom the questions are asked? Results from a survey experiment in Tanzania. *World Bank Economic Review*, 25(3):418–447. [15], [33]
- Bargain, O. and Boutin, D. (2021). Minimum Age Regulation and Child Labor: New Evidence from Brazil. *World Bank Economic Review*, 35(1):234–260. [21]
- Basu, K., Das, S., and Dutta, B. (2010). Child labor and household wealth: Theory and empirical evidence of an inverted-U. *Journal of Development Economics*, 91(1):8–14. [20]
- Beegle, K., Carletto, C., and Himelein, K. (2012). *Reliability of recall in agricultural data*, volume 98. The World Bank. [15]
- Beegle, K., Dehejia, R., and Gatti, R. (2009). Why should we care about child labor? The education, labor market, and health consequences of child labor. *Journal of Human Resources*, 44(4):871–889. [14]
- Bhalotra, S. and Heady, C. (2003). Child farm labor: The wealth paradox. *World Bank Economic Review*, 17(2):197–227. [20]
- Blair, G., Coppock, A., and Moor, M. (2020). When to Worry about Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments. *American Political Science Review*, 114(4):1297–1315. [23]
- Blair, G. and Imai, K. (2012). Statistical analysis of list experiments. *Political Analysis*, 20(1):47–77. [23], [28], [29], [30]
- Boutin, D. and Jouvin, M. (2022). Child Labour Consequences on Education and Health: A Review of Evidence and Knowledge Gaps. *Bordeaux Economics Working Papers, BxWP2022-14*, (2022-14). [14]
- Buono, C. and Babo, A. (2013). Travail des enfants dans les exploitations de cacao en Côte d'Ivoire. Pour une réconciliation entre normes locales et normes internationales autour du « bic », du balai et de la machette. *Mondes en Développement*, 41(3):69–84. [21], [25]
- Çarkoğlu, A. and Aytaç, S. E. (2015). Who gets targeted for vote-buying? Evidence from an augmented list experiment in Turkey. *European Political Science Review*, 7(4):547–566. [15], [22], [23]
- Chen, L., Zhao, X., Tang, O., Price, L., Zhang, S., and Zhu, W. (2017). Supply chain collaboration for sustainability: A literature review and future research agenda. *International Journal of Production Economics*, 194:73–87. [19]
- Chuang, E., Dupas, P., Huillery, E., and Seban, J. (2021). Sex, lies, and measurement: Consistency tests for indirect response survey methods. *Journal of Development Economics*, 148:102582. [18], [23], [27]

- Cigno, A. (2009). The economics of child labor. *Labor and Employment Law and Economics*, pages 399–417. [20]
- Co Man, K. B., Co Man, L. C., and Keith, M. M. E. (2017). The size of the LGBT population and the magnitude of antigay sentiment are substantially underestimated. *Management Science*, 63(10):3168–3186. [22]
- Comşa, M. and Postelnicu, C. (2013). Measuring social desirability effects on self-reported turnout using the item count technique. *International Journal of Public Opinion Research*, 25(2):153–172. [22]
- Corstange, D. (2010). Vote Buying under Competition and Monopsony: Evidence from a List Experiment in Lebanon. In *Annual Conference of the American political Science Association*. Citeseer. [22]
- Cullen, C. (2020). Method matters: Underreporting of intimate partner violence in Nigeria and Rwanda. *World Bank Policy Research Working Paper*, (9274). [15], [22]
- Davis Pluess, J. (2018). Children’s Rights in the Cocoa-Growing Communities of Côte d’Ivoire. [20]
- De Cao, E. and Lutz, C. (2018). Sensitive Survey Questions: Measuring Attitudes Regarding Female Genital Cutting Through a List Experiment. *Oxford Bulletin of Economics and Statistics*, 80(5):871–892. [15]
- de Janvry, A., McIntosh, C., and Sadoulet, E. (2015). Fair trade and free entry: Can a disequilibrium market serve as a development tool? *Review of Economics and Statistics*, 97(3):567–573. [19]
- Dillon, A. (2010). Measuring child labor: Comparisons between hours data and subjective measures. *Research in Labor Economics*, 31:135–159. [15]
- Dumas, C. (2013). Market Imperfections and Child Labor. *World Development*, 42(1):127–142. [20]
- Edmonds, E. V. and Shrestha, M. (2012). The impact of minimum age of employment regulation on child labor and schooling\*. *IZA Journal of Labor Policy*, 1(1):1–28. [21]
- Fafchamps, M. and Wahba, J. (2006). Child labor, urban proximity, and household composition. *Journal of development Economics*, 79(2):374–397. [21]
- Fekadu, D., Alem, A., and Hägglöf, B. (2006). The prevalence of mental health problems in Ethiopian child laborers. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 47(9):954–959. [14]
- Fountain, A. and Huetz-Adams, F. (2020). Cocoa Barometer. Technical report. [21]
- Gaddis, I., Oseni, G., Palacios-Lopez, A., and Pieters, J. (2021). Measuring Farm Labor: Survey Experimental Evidence from Ghana. *World Bank Economic Review*, 35(3):604–634. [15], [33]
- Galdo, J., Dammert, A. C., and Abebaw, D. (2018). Child Labor Measurement in Agricultural Households: Seasonality, Proxy Respondent and Gender Information Gaps in Ethiopia. (43). [15]
- Galdo, J., Dammert, A. C., and Abebaw, D. (2020). *Gender Bias in Agricultural Child Labor : Evidence from Survey Design Experiments*. The World Bank. [15]
- Glynn, A. N. (2013). What can we learn with statistical truth serum? *Public Opinion Quarterly*, 77(S1):159–172. [23], [27], [30]
- Gonzalez-Ocantos, E., de Jonge, C. K., Meléndez, C., Osorio, J., and Nickerson, D. W. (2012). Vote buying and social desirability bias: Experimental evidence from Nicaragua. *American Journal of Political Science*, 56(1):202–217. [22]



- Guarcello, L., Kovrova, I., Lyon, S., Manacorda, M., and Rosati, F. C. (2012). Towards Consistency in Child Labour Measurement: Assessing the Comparability of Estimates Generated by Different Survey Instruments. *SSRN Electronic Journal*. [15]
- Guarcello, L., Lyon, S., Rosati, F., and Others (2004). Impact of working time on children's health. *Understanding Children's Work (UCW Programme)*, (September). [14]
- Guarcello, L., Mealli, F., and Rosati, F. C. (2009). Household vulnerability and child labor: The effect of shocks, credit rationing, and insurance. *Journal of Population Economics*, 23(1):169–198. [21]
- Holbrook, A. L. and Krosnick, J. A. (2010). Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly*, 74(1):37–67. [22], [23]
- Imai, K. (2011). Multivariate regression analysis for the item count technique. *Journal of the American Statistical Association*, 106(494):407–416. [16]
- Ingram, V., van Rijn, F., Waarts, Y., and Gilhuis, H. (2018). The impacts of cocoa sustainability initiatives in West Africa. *Sustainability (Switzerland)*, 10(11):4249. [19], [20]
- Janzen, S. A. (2018). Child labour measurement: Whom should we ask? *International Labour Review*, 157(2):169–191. [15]
- Knößlsdorfer, I., Sellare, J., and Qaim, M. (2021). Effects of Fairtrade on farm household food security and living standards: Insights from Côte d'Ivoire. *Global Food Security*, 29:100535. [19]
- Krebs, C. P., Lindquist, C. H., Warner, T. D., Fisher, B. S., Martin, S. L., and Childers, J. M. (2011). Comparing sexual assault prevalence estimates obtained with direct and indirect questioning techniques. *Violence Against Women*, 17(2):219–235. [23]
- Lax, J. R., Phillips, J. H., and Stollwerk, A. F. (2016). Are survey respondents lying about their support for same-sex marriage? Lessons from a list experiment. *Public Opinion Quarterly*, 80(2):510–533. [22], [23]
- Lépine, A., Treibich, C., Ndour, C. T., Gueye, K., and Vickerman, P. (2020). HIV infection risk and condom use among sex workers in Senegal: Evidence from the list experiment method. *Health Policy and Planning*, 35(4):408–415. [22]
- Lonie, S., Martinez, M., Oulai, R., and Tullis, C. (2018). Opportunities for Digital Financial Services in the Cocoa Value Chain in Côte d'Ivoire. *Opportunities for Digital Financial Services in the Cocoa Value Chain in Côte d'Ivoire*. [21]
- Meemken, E. M. (2020). Do smallholder farmers benefit from sustainability standards? A systematic review and meta-analysis. *Global Food Security*, 26:100373. [19]
- Moseson, H., Massaquoi, M., Dehlendorf, C., Bawo, L., Dahn, B., Zolia, Y., Vittinghoff, E., Hiatt, R. A., and Gerdt, C. (2015). Reducing under-reporting of stigmatized health events using the List Experiment: Results from a randomized, population-based study of abortion in Liberia. *International Journal of Epidemiology*, 44(6):1951–1958. [15], [22], [23]
- Nkamleu, G. B. and Kielland, A. (2006). Modeling farmers' decisions on child labor and schooling in the cocoa sector: A multinomial logit analysis in Côte d'Ivoire. *Agricultural Economics*, 35(3):319–333. [20]
- Oya, C., Schaefer, F., and Skalidou, D. (2018). The effectiveness of agricultural certification in developing countries: A systematic review. *World Development*, 112:282–312. [19]
- Putnick, D. L. and Bornstein, M. H. (2015). Is child labor a barrier to school enrollment in low- and middle-income countries? *International Journal of Educational Development*, 41:112–120. [14]

- Raghavarao, D. and Federer, W. T. (1979). Block Total Response as an Alternative to the Randomized Response Method in Surveys. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(1):40–45. [22]
- Roggero, P., Mangiaterra, V., Bustreo, F., and Rosati, F. (2007). Erratum: The health impact of child labor in developing countries: Evidence from cross-country data (American Journal of Public Health (2007) 97, (271-275) DOI: 10.2105/AJPH.2005.066829. *American Journal of Public Health*, 97(3):393. [14]
- Sadhu, S., Kysia, K., Onyango, L., Zinnes, C., Lord, S., Monnard, A., and Arellano, I. R. (2020). Sadhu et al., 2020. Technical report, National Opinion Research Center (NORC), University of Chicago. [14], [18], [34]
- Sellare, J., Meemken, E. M., Kouamé, C., and Qaim, M. (2020). Do Sustainability Standards Benefit Smallholder Farmers Also When Accounting For Cooperative Effects? Evidence from Côte d'Ivoire. *American Journal of Agricultural Economics*, 102(2):681–695. [19]
- Tadesse, G., Abate, G. T., and Zewdie, T. (2020). Biases in self-reported food insecurity measurement: A list experiment approach. *Food Policy*, 92:101862. [22]
- The World Bank (2019). Cote d'Ivoire Economic Update. Technical report, The World Bank. [19]
- Tourangeau, R. and Yan, T. (2007). Sensitive Questions in Surveys. *Psychological Bulletin*, 133(5):859–883. [15]
- Trinh, T. A. (2020). Mental Health Impacts of Child Labour: Evidence from Vietnam and India. *Journal of Development Studies*, 56(12):2251–2265. [14]
- Tsuchiya, T., Hirai, Y., and Ono, S. (2007). A study of the properties of the item count technique. *Public Opinion Quarterly*, 71(2):253–272. [23], [27], [30]
- Udry, C. (2006). Child labor. *Understanding Poverty*, 4:3607–3709. [14]
- Wolter, F. and Laier, B. (2014). The effectiveness of the item count technique in eliciting valid answers to sensitive questions. An evaluation in the context of self-reported delinquency. In *Survey Research Methods*, volume 8, pages 153–168. [23]

## Appendix

**Table A1:** Descriptive statistics of basic farmers characteristics

	N	Proportion (%)
<b>Age</b>		
Less than 35 years	1 237	27.75
Between 35 & less than 50 years	2 195	49.24
50 years and above	1 026	23.01
<b>Marital status</b>		
Single	377	8.46
Divorced	24	0.54
Married/in a relationship	3 975	89.17
Widowed	82	1.84
<b>Education</b>		
No education	2 524	56.62
Primary education	1 163	26.09
Middle school	574	12.88
High school & higher	197	4.42
<b>Nationality</b>		
Ivoirian	3 297	73.96
Burkinabe	1 093	24.52
Other	68	1.52
<b>Children in the household</b>		
Yes	4 002	89.77
No	456	10.23
<b>Polygamous household</b>		
Yes	727	16.31
No	3 731	83.69
<b>Cocoa income</b>		
Less than 500 000 XOF	1 276	28.62
Between 500 000 & 1 000 000 XOF	1 637	36.72
More than 1 000 000 XOF	1 545	34.66
<b>Cocoa farm size</b>		
Less than 1.5 Ha	1 321	29.63
Between 1.5 Ha & 3 Ha	1 919	43.05
More than 3 Ha	1 218	27.32
<b>Other crop</b>		
Yes	2 282	51.25
No	2 171	48.75
<b>Region</b>		
Agnéby-Tiassa	195	4.37
Cavally	1 157	25.95
Gôh	322	7.22
Grands Ponts	459	10.30
Guémon	624	14.00
Lôh-Djiboua	776	17.41
Nawa	490	10.99
Tonkpi	435	9.76

Source: Author's calculation.

**Table A2:** Basic characteristics of respondents across groups - 2022-V1

	Control	Treatment	Placebo	P-value	Obs.
age	44.53	44.38	44.14	0.419	7 447
married	0.93	0.93	0.92	0.232	7 447
no education	0.64	0.64	0.62	0.430	7 447
ivorian	0.70	0.71	0.72	0.430	7 447
number of children	2.35	2.35	2.32	0.767	7 447
number of adults	3.11	3.14	3.10	0.481	7 447
cocoa income (XOF)	1 758 013	1 737 334	1 738 668	0.786	7 447
other crop	0.68	0.67	0.68	0.519	7 447

Source: Author's calculation.

Note: P-value comes from the F-test of equality of the means across the three groups.

**Table A3:** Basic characteristics of respondents across groups - 2022-V2

	Control	Treatment	Placebo	P-value	Obs.
age	44.46	44.71	44.48	0.613	5 593
married	0.94	0.94	0.95	0.633	5 593
no education	0.54	0.54	0.52	0.531	5 593
ivorian	0.85	0.85	0.85	0.889	5 593
number of children	2.52	2.49	2.49	0.722	5 593
number of adults	3.16	3.21	3.17	0.323	5 593
cocoa income (XOF)	1 613 451	1 617 350	1 622 078	0.964	5 593
other crop	0.68	0.69	0.69	0.668	5 593

Source: Author's calculation.

Note: P-value comes from the F-test of equality of the means across the three groups.

**Table A4:** Farmers' use of labor (using direct questions)

	Child labor			Any type of labor	
	Harvesting & pod-breaking	Preparing the farm	Employed child labor	Household labor (any tasks)	Employed labor (any tasks)
<b>2019</b>					
% of "yes"	12.76	11.28		76.96	66.69
N	2 304	2 154		4 458	4 458
<b>2022 - V1</b>					
% of "yes"	5.85	2.71	1.11	92.46	72.28
N	7 447	7 447	7 447	7 447	7 447
<b>2022 - V2</b>					
% of "yes"	3.97	0.99	0.59	92.01	76.70
N	3 750	3 750	3 750	5 593	5 593

Source: Author's calculation.

Note: In 2019 and 2022-V2, the recall period used for direct questions about child labor is 12 months, while in 2022-V1, the recall period is set at 6 months in order to match that of the list experiment. For columns 4 & 5, recall periods are 12 months.

**Table A5:** List experiments in 2022-V1

	Since the beginning of this main crop, for harvesting and breaking the pods :
Harvesting & pod-breaking	1) I worked alone, no one helped me. 2) At least one adult in my family helped me. <b>3) One or more people under 16 helped me.</b> 4) The community leader helped me.
	In total, how many statements apply to your case ?
	For this main crop, to prepare my plot (applying pesticides, clearing, felling of trees, burning of plots) :
Farm Preparation	1) I worked alone, no one helped me. 2) At least one adult in my family helped me. <b>3) One or more people under 16 helped me.</b> 4) The community leader helped me.
	In total, how many statements apply to your case ?
	Since the beginning of this main crop, to do some work on my farm (for all types of activities):
Employed child labor	1) I worked alone, no one helped me. 2) At least one adult in my family helped me. <b>3) I have employed and paid one or more people under 16 who are not members of my family.</b> 4) I employed the community leader.
	In total, how many statements apply to your case ?

Source: Author.

**Note:** For each list experiment, the bolded item is the sensitive one, related to child labor. It is only read to cocoa farmers who belong to the treatment group for that list experiment.

**Table A6:** List experiments in 2022-V2

	During the last campaign, for harvesting and breaking the pods :
Harvesting & pod-breaking	1) No adult women helped me. 2) At least one adult in my family helped me. <b>3) One or more people under 16 helped me.</b> 4) A group of communal laborers helped me.
	In total, how many statements apply to your case ?
	During the last campaign, to prepare my plot (applying pesticides, clearing, felling of trees, burning of plots) :
Farm Preparation	1) No adult women helped me. 2) At least one adult in my family helped me. <b>3) One or more people under 16 helped me.</b> 4) A group of communal laborers helped me.
	In total, how many statements apply to your case ?
	During the last campaign, to do some work on my farm (for all types of activities):
Employed child labor	1) I have not employed any adult women. 2) At least one adult in my family helped me. <b>3) I have employed and paid one or more people under 16 who are not members of my family.</b> 4) I employed a group of communal laborers.
	In total, how many statements apply to your case ?

Source: Author.

**Note:** For each list experiment, the bolded item is the sensitive one, related to child labor. It is only read to cocoa farmers who belong to the treatment group for that list experiment.

**Table A7:** Distribution of farmers according to their answer to the list experiment - "harvesting and pod-breaking"

	2019		2022 - V1		2022 - V2	
	Control	Treatment	Control	Treatment	Control	Treatment
0	3.43	2.41	2.13	2.42	0.22	0.27
1	29.56	25.77	87.82	82.05	28.45	27.17
2	52.86	49.30	9.86	15.02	65.96	66.43
3	14.15	12.58	0.19	0.52	5.37	6.13
4		9.94		0		0
N	2 304	2 154	2 586	2 484	1 842	1 844

Source: Author's calculation.

**Table A8:** Distribution of farmers according to their answer to the list experiment - "farm preparation"

	2019		2022 - V1		2022 - V2	
	Control	Treatment	Control	Treatment	Control	Treatment
0	3.02	2.91	2.52	2.94	0.31	0.27
1	31.66	28.04	87.21	84.57	26.64	27.58
2	45.17	43.01	10.22	12.41	65.44	64.60
3	20.15	14.02	0.04	0.08	7.60	7.55
4		12.02		0		0
N	2 154	2 304	2 377	2 586	1 907	1 842

Source: Author's calculation.

**Table A9:** Distribution of farmers according to their answer to the list experiment - "employed child labor"

	2019		2022 - V1		2022 - V2	
	Control	Treatment	Control	Treatment	Control	Treatment
0	3.43	2.74	2.78	1.64	0.38	0.16
1	26.65	26.42	85.55	85.70	26.63	23.75
2	48.96	41.27	11.63	12.54	61.44	63.56
3	20.96	14.16	0.04	0.13	11.55	12.48
4		15.41		0		0.05
N	2 304	2 154	2 484	2 586	1 844	1 907

Source: Author's calculation.

**Table A10:** Wording of the list experiments tests and proposed items

1st test	2nd test
During the past week :	During the past week :
1) You harvested pods on your cocoa farm.	1) You bought pesticides.
2) You helped a neighbor on his cocoa farm.	2) You received shade trees.
3) You cleared your cocoa farm.	3) You bought some fertilizer.
In total, how many statements apply to your case ?	In total, how many statements apply to your case ?

*Source:* Author.

**Note:** The equivalent direct questions followed the same wording as the proposed items for each list experiment. For example, regarding the 1st item of the 1st list experiment test, the direct question was : "During the past week, have you harvested pods on your cocoa farm ?".





---

# INFORMATION ASYMMETRIES, RISK AND DISHONESTY: A BEHAVIORAL EXPERIMENT AMONG MIDDLEMEN IN CERTIFIED COCOA VALUE CHAIN

*Co-authored with Delphine Boutin and Louis Olié.*

## **Abstract.**

This chapter investigates dishonest behaviors among cocoa middlemen in Côte d'Ivoire, focusing on the impact of observability and financial penalties in deterring such behavior. Drawing on experimental data collected in a lab-in-the-field experiment where we implement a modified version of the "die under the cup" task, we examine the cheating behavior of 151 cocoa middlemen over several rounds of interaction. Our findings reveal that cheating is prevalent among cocoa middlemen, with 78% of players cheating at least once during the game. However, we found heterogeneous cheating patterns: 59% of cocoa middlemen consistently cheated when faced with a losing outcome, even when the risk of detection and sanction is high, while 22% of them never cheated, and 19% did so occasionally. We identify key individual characteristics, such as age, religion, and attitudes toward risk, that influence the likelihood of cheating. Further, we show that introducing monitoring and sanctions leads to a significant reduction in cheating, highlighting the effectiveness of such mechanisms in deterring dishonest behaviors.

## 1. Introduction

Intermediaries play an essential role in agricultural global value chains, connecting producers to international markets. However, the prevalence of information asymmetries within such value chains, which are often made up of many geographically dispersed players, and the weak enforcement mechanisms, are likely to encourage dishonest behavior, such as the misreporting of quantities and/or qualities (Rustagi and Kroell, 2022; Mitra et al., 2018). Despite the potential for widespread misconduct, the level of dishonest behavior among intermediaries is likely to be heterogeneous. While conventional economic theory posits that *homo economicus* engage in dishonest behavior based on a cost-benefit analysis and should therefore lie whenever it is beneficial (Becker, 1968), recent experimental research challenges this paradigm. Individuals do not always lie or do not lie to the maximum extent, even when it would benefit them, and they run no risk of being caught, indicating the existence of an intrinsic cost of lying (Fischbacher and Föllmi-Heusi, 2013). The decision to engage in dishonest behavior is thus shaped by several individual characteristics, such as the preference for truth-telling, guilt aversion, self and social image concerns, and social norms (Khalmetski and Sliwka, 2019; Abeler et al., 2019; Bašić and Quercia, 2022; Huber et al., 2023). Understanding the drivers of dishonest behavior among middlemen is crucial for designing effective interventions, to enhance the integrity and transparency of agricultural value chains prone to information asymmetries.

This paper considers the case of cocoa middlemen in Côte d'Ivoire, operating in certified value chains. Their role in the cocoa supply chain is crucial, as they are required to buy and transport cocoa from many small, remote farmers on behalf of the local cooperatives or buying centers. They rank first at the bottom of the cocoa value chain after farmers, prior to the cocoa being distributed to cooperatives followed by local processors or exporters. Because they operate in certified value chains, i.e. the farmers they source from comply with certification standards, the information they have on the cocoa's origin (certified or conventional) is crucial, as they are the first to introduce it into the supply chain. However, since whether the cocoa is certified is a credence characteristic, i.e., indiscernible to the naked eye, and that certified cocoa is sold at a higher price on the market than conventional one, middlemen have an incentive to misreport the nature of the cocoa they bring to the cooperatives. Such behavior is likely, as recent evidence confirms large differences between the quantities of certified cocoa introduced into the supply chain by farmers and those reported at the cooperative level (Bernard et al., 2024).

We implemented an adapted version of the “die-under-cup task”, an experimental game commonly used to study dishonesty (Fischbacher and Föllmi-Heusi, 2013), in which participants are asked to roll a die in private and report the outcome to the experimenter. Players are subsequently rewarded based on their reported outcome, providing them an incentive to lie to increase their financial rewards. In our settings, players do not roll a physical but a digital die in private, using a tablet. This allows the actual result of each player’s die roll to be recorded, enabling individual cheating to be determined afterward.<sup>1</sup> The payouts associated with each dice outcome were predefined based on a risk-elicitation experiment similar to Holt and Laury (2002). Implementing an experimental game gives us the opportunity to investigate middlemen’s cheating behavior in a controlled environment, thus overcoming the current difficulties in identifying their real-life behavior. Indeed, monitoring middlemen in real-life settings would require an effective traceability system to capture information at each supply chain level, and prevent them from misreporting the nature of the cocoa in the system, but such a tool does not yet exist. Similarly, asking them directly about their opportunistic behavior would have been prone to social desirability bias.

Beyond individual characteristics, the experimental literature has recently focused on contextual factors likely to increase or discourage dishonest behavior. Along with the payoff frame (wins vs. losses) and size, the potential observability of dishonest behavior is one of the contextual factors likely to influence an individual decision to engage in such behavior (Garbarino et al., 2019; Charness et al., 2019; Kajackaite and Gneezy, 2017; Fries et al., 2021). Being observed usually contributes to increased social concern, as individuals do not want to be seen as misbehaving by others. Observability has thus been found to have a deterrent effect on cheating behavior (Gneezy et al., 2018; Fries et al., 2021; Bašić and Quercia, 2022). However, existing studies looking at the impact of observability on cheating are usually set in a one-shot game, thus offering no means of examining the impact of observability over time on individual cheating behavior in the setting of repeated interaction games. To date, little attention has also been paid in the economic literature to financial sanctions as a tool for reducing dishonest behavior, whereas this could prove effective for people who are not prone to social image concern but rather to loss aversion (Thielmann and Hilbig, 2018; Dai et al., 2018). Our paper aims to address these gaps by introducing a new game feature to assess whether implementing a monitoring system, namely the possibility of being observed and subsequently penalized if caught lying, can help reduce cheating behavior. The experiment was divided into three treatments, each featuring a different level

---

<sup>1</sup>Using a tablet for dice rolling rather than a cup may lead participants to question the confidentiality of their behaviors. In such a scenario, participants are likely to engage in less dishonesty. Thus, we argue that our findings represent a conservative estimate of dishonest behavior among participants.

of monitoring. In the first treatment, players were never observed by the experimenter. In the second treatment, the experimenter could observe the die-roll outcome after players reported the number displayed on their tablets. Observation was not systematic but determined by the outcome of a physical die-roll, with an even die leading to monitoring by the experimenter. In the third treatment, in addition to the 50% risk of being observed, we introduce the possibility of incurring a financial penalty if the experimenter detected that the number displayed on the tablet did not match the one declared by the player. Each treatment consisted of five rounds, with the probability of losing, i.e. the probability of getting a number that yields no reward, progressively increasing over rounds. Final payouts were determined by randomly selecting one round for payment in each treatment.

Overall, 151 cocoa middlemen operating for 19 cooperatives sourcing certified cocoa participated in the game, and took part in a 20-minute survey about their activities as cocoa middlemen before the experiment. Our empirical analyses are drawn from 2,265 observations, corresponding to 15 rounds played by each of the 151 middlemen. We find that cheating is a widespread behavior among cocoa middlemen, occurring in 39% of the rounds played and rising to 73% when considering only losing rounds. Given the possibility of retrospectively observing each player's behavior across 15 rounds, players are classified into three categories based on their cheating behavior. Most players systematically cheat when faced with losing rounds (59%), 22% never cheat, while 19% cheat only occasionally. These results are in line with previous literature, showing that some individuals do not cheat when they have an opportunity to do so and that it would benefit them, even when there is no risk of being caught. Using a multinomial logit, we find that being a "sometimes" or "always" cheater is influenced by a few individual characteristics, such as age, religion, and attitudes toward time and risk. We investigate whether introducing observation and sanction affects the cheating behavior of "occasional cheaters" using a conditional logit with player fixed effects. We find that the mere risk of observation, and observation combined with a penalty, is associated with a reduction in cheating among players. Players are 42% less likely to cheat when they risk being observed and 52% less likely when they risk being observed and penalized for cheating. Findings also show that having actually been observed while cheating previously is negatively correlated with cheating in the current round, reducing the likelihood of cheating by 89%. Besides, the more financial penalties players get, the less they cheat. We find no significant effect of previous losses and cheating history on cheating behavior in the current round.

Our contributions to the literature are twofold. First, we contribute to the experimental literature on cheating behavior, particularly the emerging literature using lab-in-the-field experiments to study fraud

and corruption behaviors in developing countries (Armand et al., 2023; Rustagi and Kroell, 2022; Harris et al., 2024; Hanna and Wang, 2017). In line with Abeler et al. (2019) and Rustagi and Kroell (2022), our computerized version of the "die-under-cup-task" allows us to observe individual cheating behaviors, rather than aggregated ones. We provide additional evidence that, even when players face no risk of being caught cheating and that doing so benefits them, some of them never engage in cheating. Age, and particularly religion, seem to play an important role in the decision to engage in cheating behavior, likely by influencing both self and social image concerns. Additionally, players classified as risk-takers are significantly more likely to engage in cheating than others. Beyond better distinguishing intrinsic dishonesty from the randomness of die rolls (Rustagi and Kroell, 2022), observing individual behaviors rather than aggregate ones provides an opportunity to analyze the effect of monitoring and sanction mechanisms on individual lying. Thus, we also contribute to the smaller body of literature that seeks to understand which mechanisms can inhibit dishonest behaviors (Gneezy et al., 2018; Kajackaite and Gneezy, 2017; Fries et al., 2021; Bolton et al., 2021; Jiang and Villeval, 2022). In this paper, we show that for players who have a finite cost of lying, and are thus likely to react to incentives aimed at reducing cheating, both simple monitoring of behavior and monitoring combined with a penalty for cheating are effective mechanisms.

Second, we contribute to the literature investigating middlemen's role and behavior in agricultural value chains with information asymmetries (Rubinstein and Wolinsky, 1987). Recent empirical evidence in developing countries shows that intermediaries are prone to exploit market failures, such as low market competition or price information imbalances, to increase their profit margins (Renier et al., 2023; Casaburi and Reed, 2022; Bergquist and Dinerstein, 2020; Bergquist and McIntosh, 2021; Chatterjee, 2023). We provide evidence that such behavior can be observed in an experimental setting and concerns a large proportion of middlemen. Focusing on the cocoa value chain, we offer quantitative insights into middlemen's socio-economic characteristics and activities, as well as insights into their cheating behavior when faced with asymmetric information, through our lab-in-the-field experiment. This provides valuable information about cocoa middlemen, who have been neglected in the literature on cocoa and certified value chains despite being considered the most opaque tier in the cocoa supply chain (Stoop et al., 2021).

The remainder of this article is organized as follows. Section II provides contextual elements that frame our experimental game. Section III presents the experimental design and procedure, and describes

cocoa middlemen activities based on first-hand survey data. Results are discussed in sections IV and V. The last section concludes.

## **2. Context**

### **2.1 Traceability in the certified cocoa supply chain**

Our lab-in-the-field experiment takes place in Côte d'Ivoire, the world's largest cocoa producer, where approximately one million small-scale farmers generate an annual output of over 2 million tons of cocoa beans and contribute to 40% of the global cocoa production (Renier et al., 2023). We implemented our cheating game among middlemen, important players of the complex Ivorian cocoa value chain, which involves many geographically dispersed actors. Cocoa farmers sell their cocoa to middlemen, operating on behalf of cooperatives. Exporting companies then directly acquire cocoa from these cooperatives. We focus on middlemen working in certified value chains, i.e., those responsible for sourcing cocoa from certified farmers. The latter are farmers who comply with a set of criteria established by certification bodies, such as Rainforest Alliance or Fairtrade. Although the concept of certified cocoa theoretically implies the physical traceability of cocoa back to the farmer, farm-level traceability remains an important challenge in the cocoa industry. Despite nearly 60% of Ivorian cocoa farmers affiliating with certification standards (World Bank, 2019), most industry players only reach traceability at the cooperative level. In fact, as of 2022, 50% of global cocoa production remained completely untraceable (Huetz-Adams and Fountain, 2022; Renier et al., 2023).

Achieving cocoa traceability down to the farmer level is difficult, as interactions between supply chain players are fraught with information asymmetries. Indeed, two types of cocoa are available on the market, conventional or certified, the latter being sold at a higher price than uncertified cocoa. As certified cocoa is a credence good, i.e. it is impossible to determine its type with the naked eye, this situation gives rise to information asymmetries. Middlemen working in certified cocoa value chains have thus a strong incentive to engage in side-sourcing, which involves sourcing uncertified cocoa but reporting it as certified to the cooperative in order to increase their profits. The supply of truly certified cocoa to cooperatives, therefore, depends on these middlemen's honesty. Yet, they are considered the

most opaque link in the cocoa value chain (Stoop et al., 2021), with little information available about their activities, which are difficult to monitor.

Significant reporting discrepancies were observed in terms of volume of certified cocoa between farmers and cooperatives for a set of 22 cooperatives in Côte d'Ivoire (Bernard et al., 2024). In fact, the cocoa volumes recorded in the cooperatives' traceability registers were at least two to ten times higher than those declared by the farmers for the same period. This suggests a parallel supply of supposedly certified cocoa (defined as side-sourcing) to the volume sourced from the cooperatives' farmers. We argue that the likelihood of side-sourcing occurring at the farmer level is low since cocoa farms located in the same community tend to be certified, making it difficult and costly for farmers to collect non-certified cocoa. Thus, discrepancies between volumes reported by farmers and cooperatives registers found in Bernard et al. (2024) suggest that middlemen or cooperatives engage in side-sourcing. This companion paper also investigates the impact of a mobile traceability application designed to increase transparency in cocoa transactions between farmers and middlemen. Not all middlemen used this tool despite the absence of any technical issues, suggesting reticence among middlemen to use a tool that monitors their cocoa supply behavior. Among those who did use it, only small volumes of certified cocoa were digitized in the application, well below the volumes actually purchased by middlemen. The study, however, does not reveal whether middlemen are prone to dishonest behavior, the determinants of such behavior, or whether monitoring or sanctions would be effective tools to increase their honesty.

## **2.2 Detecting dishonest behavior**

Measuring intermediaries' propensity to engage in side-sourcing using traditional observational data or surveys is challenging because such behaviors are hidden and subject to social desirability bias (Rustagi and Kroell, 2022). Direct questions about opportunistic behaviors often lead to underreporting, as individuals tend to provide socially acceptable rather than truthful answers, especially when their responses might be judged or result in negative consequences. Additionally, the fact that whether the cocoa is certified or conventional is a credence feature, that the middlemen are involved with many farmers, and that purchasing activities are relatively informal and opaque, further makes accurate monitoring difficult. Even when traceability information does exist, the fact that it is partly manipulated by those who generate it adds complexity to the process of measuring fraud or dishonest behavior from supply chain players (Bernard et al., 2024).

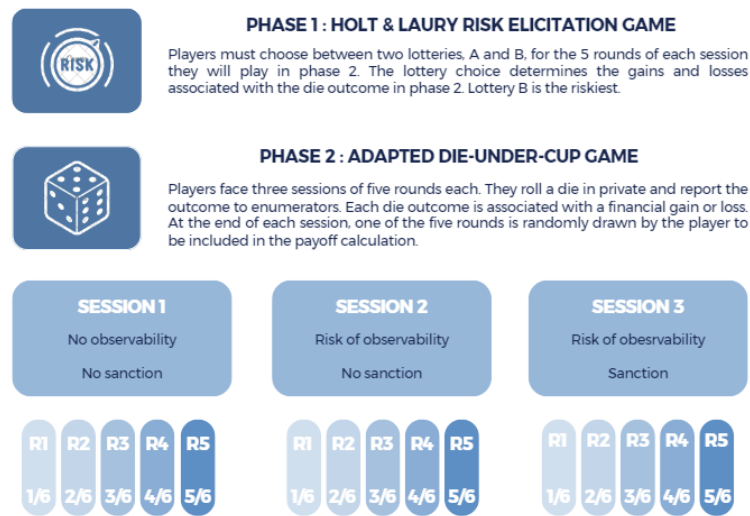


We thus rely on a lab-in-the-field experiment on cheating to precisely measure dishonest behaviors and analyze the underlying factors of engaging in dishonesty, such as personal characteristics and economic incentives. Indeed, experimental approaches offer a controlled alternative to obtain a measure of individuals' dishonesty. Armand et al. (2023); Mitra et al. (2018); Rustagi and Kroell (2022); Harris et al. (2024) and Stoop et al. (2021) demonstrate how experimental games can predict real-world behaviors. Armand et al. (2023) conducted a corruption game in Mozambique where citizens could send bribes to local leaders. They found a significant correlation between the bribes sent in the game and the actual misappropriation of community funds, indicating that the game accurately reflected real-world corrupt behaviors. Stoop et al. (2021) used a "Mind game" where participants guessed the color on the backside of a card to achieve high payoffs, showing a significant correlation between lab cheating and the likelihood of not reporting overpayment in real-world scenarios. Mitra et al. (2018) used a subsidy application game with potato farmers in India, showing that those who underreported yields in the game also did so in real life, thus establishing the game's predictive power for real-world dishonesty. Rustagi and Kroell (2022) investigated milk quality reporting using a lab game where participants could adulterate milk, finding that lab behaviors correlated with actual market misreporting. Harris et al. (2024) used corruption games with police officers, showing that experimental measures of corruption correlated with observed corruption behaviors. Collectively, these studies underscore the predictive power of experimental games in understanding and forecasting dishonest behaviors across various real-world contexts.

### **3. Experimental design**

#### **3.1 Protocol**

The game's design is grounded in the assumption that the decision to engage in dishonest behavior involves a trade-off between anticipated gains and costs, which vary according to the probability of being caught, the severity of sanctions upon detection, and individuals' willingness to take risks (i.e., their risk attitude). Consequently, we merged our cheating game with a conventional risk elicitation task to compare and assess individuals' general risk attitudes in such a game. To this end, we divided the experiment into two phases, summarized in Figure 1 below.

**Figure 1:** Structure of the experimental game

Source: Authors.

**Phase 1: Risk elicitation game.** The initial phase of the game is an adapted version of the classic Holt and Laury (2002) experiment, which was designed to assess risk preferences. During this phase, participants face a series of five paired lottery choices (Table 1), where they had to choose between two options, A or B, with lottery B being the riskier choice. These paired lottery choices were designed to reflect different probabilities of success. The lottery selection phase was of significant importance, as the participants' payoffs in the next part of the experiment depended on the chosen lottery option. Therefore, we can differentiate the preference for dishonesty from risk attitudes. Both players and experimenters recorded the players' lottery choices on paper, which was kept on hand throughout the game to prevent any mistakes related to expected payouts (see player sheet in Appendix A).

**Phase 2: Adapted die-under-cup game.** The second phase of our experimental setup draws inspiration from the cheating game originally introduced by Fischbacher and Föllmi-Heusi (2013), the "die-under-cup task", in which participants are asked to roll a die in private and report the outcome to the experimenter. Participants are then rewarded according to the reported outcome, giving them an incentive to lie to increase their financial rewards. We implemented a modified version of this game.

First, we introduce a feature designed to observe retrospectively the true individual outcomes of the die rolls. Participants roll a six-sided fair die using a specially designed tablet application that records the true die outcomes after the experiment ends. Similar to Abeler et al. (2019) and Rustagi and Kroell (2022), we are thus able to identify individual cheating behavior. The remaining steps of the

**Table 1:** The five paired lottery-choice decisions

	Option A	Expected Payoff Option A	Option B	Expected Payoff Option B	Expected Payoff Difference
Lottery 1	5/6 of FCFA 1,500 1/6 of FCFA -1,000	1,083	5/6 of FCFA 3,000 1/6 of FCFA -2,000	2,167	-1,083
Lottery 2	4/6 of FCFA 1,500 2/6 of FCFA -1,000	667	4/6 of FCFA 3,000 2/6 of FCFA -2,000	1,333	-667
Lottery 3	3/6 of FCFA 1,500 3/6 of FCFA -1,000	250	3/6 of FCFA 3,000 3/6 of FCFA -2,000	500	-250
Lottery 4	2/6 of FCFA 1,500 4/6 of FCFA -1,000	-167	2/6 of FCFA 3,000 4/6 of FCFA -2,000	-333	167
Lottery 5	1/6 of FCFA 1,500 5/6 of FCFA -1,000	-583	1/6 of FCFA 3,000 5/6 of FCFA -2,000	-1,167	583

Source: Authors' calculation.

experiment follow the standard protocol: participants are asked to report the number displayed on the screen. However, because the experimenter does not have visual access to the tablet, participants can lie to maximize their earnings. During the experiment, players and experimenters face each other, maintaining a distance of at least two meters (Figure A1). No other person is in the classroom, ensuring only players can see the content on their tablet.

Second, we introduce the risk that experimenters will observe the result of players' dice rolls, and the risk of players incurring a financial penalty in the event of cheating. Thus, the game is divided into three sessions. In the first one, participants are never observed by the experimenter. The second session introduces the possibility of experimenters observing players' true outcomes directly on their tablets. The observation is not guaranteed but depends on the outcome of a physical die-roll, resulting in a 50% chance of being observed.<sup>2</sup> At this stage, no penalty is applied if players are caught cheating. The third session combines the risks of observability and of being financially penalized. Following the same observation procedure as in session 2, if experimenters detect that the number displayed on the tablet does not match the one declared by participants, a penalty of -2,000 FCFA is applied for that round, representing  $2/3$  of the maximum amount a player can win per round.

<sup>2</sup>In sessions 2 and 3, after rolling the die on the tablet, players then roll a physical die to determine whether the experimenter will observe their actual outcome. An even die on this second dice roll leads to monitoring by the experimenter.

Finally, when most die-under-cup experiments take place in a one-shot game setting, we introduce repetition to simulate real-world scenarios with recurring opportunities for dishonesty. The participants played a series of three sessions, each consisting of five rounds. The probability of losing increases in each round by the same amount over the three sessions. Players have thus a one-in-six chance of losing in the first round compared with a five-in-six chance in the fifth round. This allows us to observe whether the likelihood of cheating is higher when the risk of loss increases. The potential earnings are contingent upon the lottery choices that participants made during the previous risk-elicitation game.

Overall, participants roll the die fifteen times over three sessions, and payouts are randomly determined at the end of each session. Players draw a chip from a bag containing five chips numbered 1 to 5, corresponding to the five rounds (Figure A1). The round drawn is the one selected for payment for this session. Once the three sessions are completed, the experimenter calculates the total winnings or losses based on the three selected rounds.

### **3.2 Field implementation**

One of the main cocoa purchasing companies in Côte d'Ivoire gave us access to its partner cooperatives. We selected 19 cooperatives from a pool of 29 (Figure A2), considering time, budget, and logistical constraints, and excluding those situated in isolated areas. These 19 cooperatives provided us with comprehensive lists of cocoa middlemen working for them, referred to as *délégués*, and/or their assistants who perform the same activities but under their supervision, known as *sous-délégués*. We randomly selected middlemen and assistants from each cooperative's list to participate in our experiment. We informed the selected individuals, as well as the cooperatives they worked for, that they would be participating in a quantitative survey on their activity as cocoa middlemen and of their opportunity to play a game in which they could win money. Except for travel expenses incurred to reach the survey and experiment location, we did not inform them regarding the potential amounts they could earn from their participation. The field experiment involved 17 sessions from August 16 to September 10, 2021, across nine cities in key cocoa-producing areas: Soubré, Gnipi, Cecchi, Alépé, Sikensi, Lakota, Gagnoa, Daloa, and Guiglo. The sessions were conducted in empty classrooms in primary or secondary schools, and a total of 215 middlemen attended informational sessions. As most of them lived in villages away from the experiment's cities, where the schools were located, we compensated them with 5,000 FCFA to cover their travel costs, equivalent to about 8 dollars. Four experimenters and one supervisor were

assigned the responsibility of ensuring the attendance of the chosen participants on the day of the experiment.

The experiment started with a brief quantitative survey aimed at gathering socio-demographic and activity-related information. Middlemen received a reward of 6,000 FCFA for completing this 20-minute short survey. Subsequently, all participants gathered in a classroom for an introductory explanation of the experiment, providing them with a comprehensive understanding of the game, including details about potential associated payouts. Practice sessions with volunteer participants were conducted to ensure all middlemen understood the rules. Participation in the experiment was entirely voluntary, and those who agreed to play staked the 6,000 FCFA received from the survey. At this stage, 41 middlemen chose not to proceed and were released from the experiment. To ensure strict confidentiality and to prevent any potential peer influence that could influence participation in the game, we assigned each enumerator-player pair to a separate classroom, where the potential payoffs were displayed on a blackboard (Figure A1). At this stage, each experimenter briefly re-explained the game rules (without mentioning the possibility of lying), and the entire experiment was first tested with no money at stake to ensure that each player had a clear understanding of the game's steps and rules. Following this test version, the actual game, with money involved, begins. While 174 middlemen agreed to participate in the test, only 151 decided to take part in the actual experiment. Except for age and education, we observe no significant difference between individuals who chose to participate and those who opted out (Appendix B, Table B1). In total, the average duration of the survey and the experiment was about two hours.

#### **4. Pre-lab quantitative survey**

Before the experiment, middlemen were surveyed regarding their socio-economic characteristics and cocoa trading activity throughout the entire 2020-2021 cocoa season. This individual questionnaire gathered data such as their age, household composition, migration history, time preferences, socio-economic ties with farmers, work experience, working conditions, and income from their middlemen activity.

**Player profile.** A total of 151 middlemen took part in the actual experiment. On average, players are approximately 39 years old (Table C1 in the Appendix), and most of them are married (92%). Players

come from large households, with an average of 4.5 adults and 4.4 children. 19% of middlemen have not received any formal education, 27% have completed primary education, 24.5% have attained middle school education, and 29% have achieved high school or higher education. The players are mainly Christians (47%) and Muslims (41.5%), with the remainder being animists.

On average, players have been working as cocoa middlemen for their current cooperative for four years. However, 62% of them had previous experience as middlemen, either working independently (as *pisteurs*) or for another cooperative. Although not all participants had prior middlemen experience, 94% of them had previously worked in the cocoa sector, mainly as cocoa farmers (in 95% of cases). 59.5% of the players are main *délégués*, the remainder being assistants. It is worth noting that 94% of middlemen are involved in additional activities, with 90% of them being involved in cocoa farming and 25% in selling other agricultural products. Only 40% of participants consider that being a cocoa middleman is their main activity.

**Cocoa middleman activity.** Not all middlemen source cocoa from the same number of farmers, as they are allocated a group of farmers within a specific geographical area defined by the cooperative they work for. On average, *délégués* work with approximately 114 farmers, although some source cocoa from up to 500 farmers (Table C2 in Appendix). In our sample, middlemen delivered on average 90.5 tons of cocoa to their cooperative during the cocoa season preceding the survey,<sup>3</sup> 2/3 of it being collected during the main crop (from October to March) and 1/3 during the light crop (from April to September). Half of the middlemen were set delivery targets by the cooperative, and 90% of them failed to meet these targets. On average, *délégués* received a commission of 34 FCFA per kilo of certified cocoa delivered during the main crop. Annual income generated solely from cocoa middleman activities averaged around 1.5 million FCFA (equivalent to 2,476\$) during the 2020-2021 cocoa season, with a relatively high standard deviation, indicating a significant income disparity among *délégués*. Some of them reported yearly earnings of up to 5,000,000 FCFA (8,255\$), suggesting the presence of highly successful middlemen. Consistent with previous figures indicating that *délégués* delivered 2/3 of their annual volume during the main crop, the income generated by their cocoa trading activity during the main crop also accounts for 2/3 of their total income over the cocoa season.

---

<sup>3</sup>Covering October 2020 to September 2021.

**Challenges in cocoa sourcing.** Most middlemen (79%) reported difficulties in purchasing cocoa from their regular group of cocoa farmers over the past 12 months. Transportation issues due to poor road conditions (77%) and the lack of operational means of transport (46%) are the main difficulties experienced by traders, along with changes in the farm-gate cocoa price over the cocoa season (70%). Furthermore, 31% of middlemen also reported financial difficulties when purchasing cocoa. In fact, 30% of *délégués* receive no cash advance from their cooperative, and 80% of those who do receive such an advance report that the money lent by the cooperative is never or rarely sufficient to cover all their cocoa purchases. In such cases, 78% of middlemen encountering cash shortages opt to pay farmers at a later time, and 32% use money from their personal savings.

**Costs incurred by middlemen.** Middlemen's work involves significant costs mainly related to transportation, such as fuel expenses, buying or leasing a vehicle, paying road tolls, or paying drivers. Additionally, 41% of middlemen employ external helpers, such as friends or family members (commonly known as "coxeurs"), incurring salary expenses. While middlemen operating in geographical areas close to the cooperative benefit from the advantage of storing purchased cocoa directly in the cooperative's warehouse (19%) or in storage facilities provided by the cooperative and located in the cocoa communities (39%), others must either own or rent their own storage facilities (42%). On average, participants reported that operating as cocoa middlemen cost them 951,533 FCFA over the past 12 months.

**Role of middlemen as credit providers.** Middlemen also take on the role of credit providers to farmers, with 92% of them reporting that they lend money to the farmers they source cocoa from (Table C3 in Appendix). Among *délégués* who lend money to cocoa farmers, 35% provides loans up to 25% of their farmers' pool, and 47% up to 50%. This suggests that middlemen cover the credit needs of a significant proportion of the farmers they work with. Farmers mainly approach *délégués* for credit to cover expenses such as their children's school fees (89%).<sup>4</sup> Other reasons for seeking credit include purchasing agricultural inputs (56%), addressing personal shocks, such as funeral expenses (47.5%), covering medical costs (85%), and buying food (28%).

---

<sup>4</sup>This high proportion highlights the timing mismatch between farmers' school fee payments (usually in September) and their peak income period during the main crop season (typically from October to December)..

**Payment transactions.** Despite 93% of cocoa middlemen having a mobile money account, only 13% of them report paying cocoa farmers by this means. The majority (90%) of those not using mobile money with farmers express a preference for cash transactions, the latter remaining untraceable. Besides, 48% indicate that farmers themselves do not wish to be paid by mobile money.<sup>5</sup>

## 5. Cheating behavior of cocoa middlemen

### 5.1 Cheating rates

Our final sample consists of 2,265 observations, corresponding to 15 rounds played by each of the 151 middlemen. The digital dice used by the players are fair, as they had an equal number of losing and winning rounds (an average of 7.7 losing rounds), and the distribution of the real die-roll outcomes, recorded on the tablet, closely approximates the theoretical distribution of a random dice roll (represented by the red line in Figure 2).

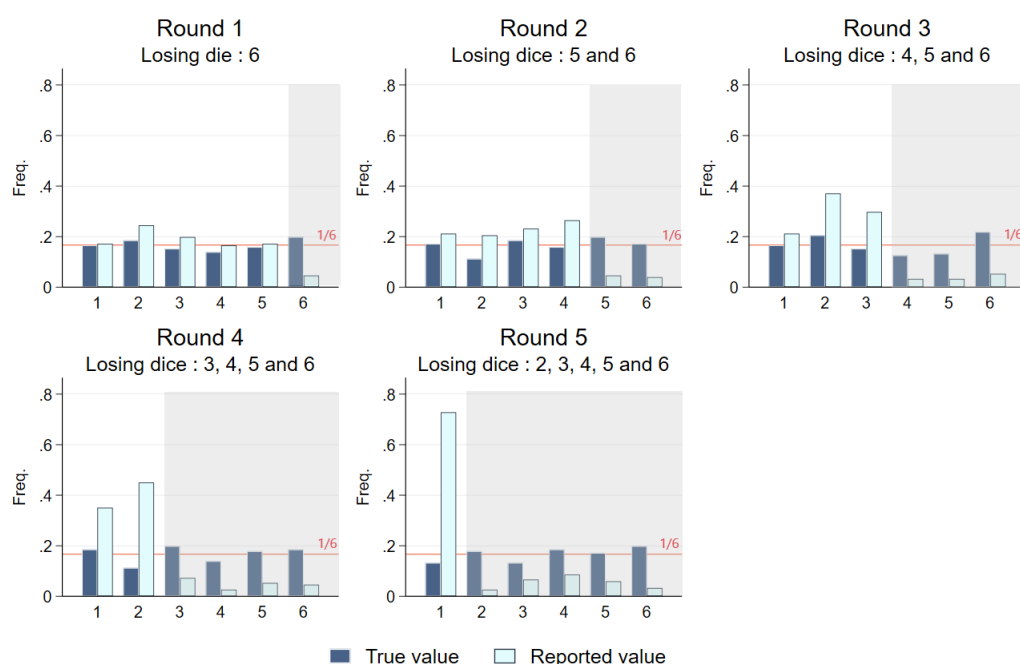
**Deviation from theoretical distribution.** Figures 2, D1, and D2 show for each session the difference between the distribution of actual dice outcomes captured by the tablet and the outcomes reported by players. The grey area indicates the dice outcomes that result in a loss for the corresponding rounds, i.e., those that do not lead to a monetary win. Results show that players less frequently report numbers associated with losses, and that numbers associated with wins are therefore over-represented compared with the distribution of a fair die. The shift to the left in the distribution of outcomes reported to enumerators compared to the theoretical distribution indicates some misreporting from players.<sup>6</sup> When players provided the experimenter with a different number from the actual one they rolled during a losing round, they tended to report the threshold winning number, i.e. the minimum required number for them to secure a win, in 65.6% of cases (Figures 2, D1 and D2). This observation provides further

---

<sup>5</sup>A survey of 4,458 cocoa farmers carried out by the authors in 2020 showed that 68% of farmers had a mobile money account, and among them, 81% were in favor of cocoa being paid for by their middlemen via mobile money.

<sup>6</sup>Note that in a negligible proportion of rounds (0.5%), the reported number is higher than the one obtained, with a significant difference between winning rounds (1%) and losing rounds (0%). Over-reporting is not rational, as it decreases the probability of earning some money. It may be attributed to a lack of attention or reporting errors by the player or enumerator. About 4.6% of players over-report at least once during the experiment. Such errors typically occur only once per player, except for one of them who repeated this error four times over 15 rounds. Consequently, we consider over-reporting as random noise, which should not affect our analyses.



**Figure 2:** Distribution of true die roll outcomes and outcomes declared by players for Session 1

Source: Authors' calculation.

**Note:** The red horizontal line indicates the theoretical distribution of a random draw. This graph shows the distribution of actual and reported roll-dice outcomes for session 1, where the observability and penalty features are not integrated. Similar patterns are observed in sessions 2 and 3 (in Appendix).

descriptive evidence of cheating, strongly suggesting that the numbers reported by players are not the result of a random process.

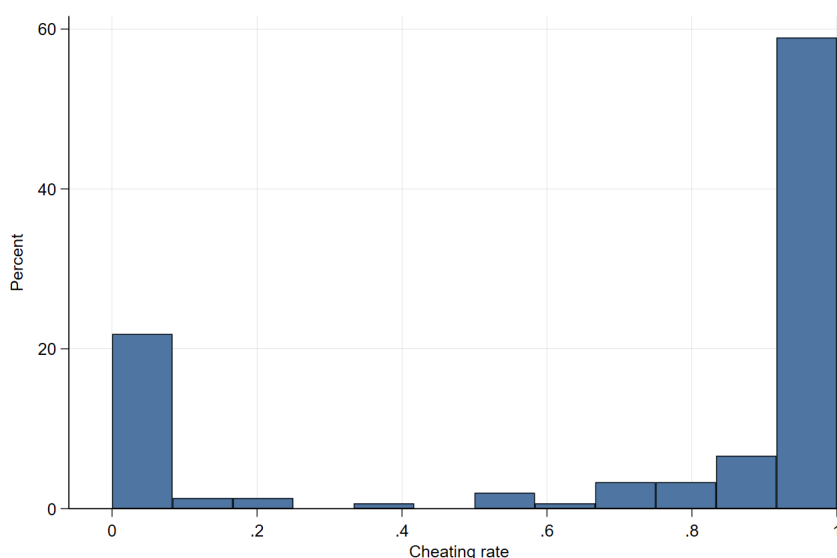
**Overall cheating rates.** The unconditioned cheating rate, which indicates the proportion of rounds, whether winning or losing, in which players behave dishonestly, is 39%.<sup>7</sup> When focusing on losing rounds, players under-report their die-roll outcome in 72.5% of cases. In contrast, under-reporting occurs in only 3% of cases during winning rounds. The substantial occurrence of misreporting in losing rounds indicates that players were well aware of the potential for dishonesty and occasionally exploited this opportunity to increase their payouts. This discrepancy of behaviors between winning and losing rounds further stresses that the misreporting observed during losing rounds was likely a strategic choice. The individual overall cheating rate is quite high, as 78% of middlemen cheated at least once when facing a losing die.

<sup>7</sup>This is relatively low compared with other studies involving a die-rolling task in a lab experiment. For instance, in a meta-analysis of 129 die-rolling tasks by Gerlach et al. (2019), the average proportion of individuals exhibiting dishonest behavior was 52% (CI=[47%; 57%]). A closely related study of ours is Rustagi and Kroell (2022), which investigates the role of innate honesty in explaining the variation in milk quality sold in naturally occurring markets in India, also employing a die-rolling experiment. They report that 50 percent of the milkmen exhibited dishonest behavior.

**Influence of the probability of losing.** The proportion of players engaging in dishonest behavior increases as the probability of losing increases over the rounds, regardless of whether it is a losing or a winning round. For instance, in round 2 (where the probability of losing is  $2/6$ ), the probability of cheating increases by approximately 11 percentage points compared to round 1 (where the losing probability is  $1/6$ ) (Table D1). In the last round, where the probability of losing is  $5/6$ , the likelihood of cheating rises by 42 percentage points compared with the first round. A more straightforward interpretation of these findings is directly considering the theoretical losing probabilities. Transitioning from a null losing probability to a certain losing probability results in an increase of 63 percentage points in the likelihood of cheating. On average, this corresponds to an increase of approximately ten percentage points for every  $1/6$  increment in the probability of losing (i.e., for each successive round).

## 5.2 Cheater profiles

**Individual cheater profiles.** An analysis of middlemen's behavior over the 15 rounds uncovers three distinct player profiles: those who never cheat (22%), those who cheat occasionally (19%), and those who consistently cheat when facing a losing roll-die outcome (59%) (see Figure 3). Players who consistently refrain from cheating, even when it could yield benefits, are characterized as the "ethical type" in the existing literature. In our game, 22% of the players were always honest even when it did not yield any tangible benefits, which suggests inherent honesty. For these players, lying carries infinite intrinsic costs, both direct, such as an aversion to deception, and indirect, such as guilt aversion (Battigalli and Dufwenberg, 2007; Battigalli et al., 2013), an aversion to violating social norms (Kajackaite and Gneezy, 2017), or a desire to maintain a positive social and/or self-image (Bénabou and Tirole, 2006). Occasional cheaters suffer the same direct and indirect cost of lying, but to a limited extent: they cheat when the benefits of dishonest behavior outweigh their finite costs of lying. Although representing only 19% of our sample, this type of player is the most likely to adjust their behavior based on observability and sanction features of our game. Finally, 59% of players fall into the "economic type" category, as defined in previous studies (Gibson et al., 2013; Kajackaite and Gneezy, 2017). These individuals exhibit a zero cost of lying and engage in cheating whenever it benefits them. Middlemen who consistently cheated in our settings thus earned an average gain of 6,753 FCFA, representing 38% more than the average gain of occasional cheaters (4,872 FCFA) and three times that of never cheaters (2,318 FCFA) (Table D2).

**Figure 3:** Distribution of individual cheating rates

*Source:* Authors' calculation.

**Note:** The overall individual cheating rate is the proportion of rounds in which individuals under-report the dice outcome relative to the proportion of losing rounds.

**Cheater profile determinants.** We explore the effects of players' characteristics on their cheating profile, i.e., whether they never cheat, sometimes cheat, or cheat consistently, using a multinomial logistic regression with the reference category being players who "never" cheat. Table 2 shows that few socio-demographic characteristics influence cheating behavior. We find that age is associated with a 10% reduction in the probability of being classified as "always cheaters" compared to "never cheaters" (column 2), which is in line with previous findings (Gerlach et al., 2019). While the level of education has no significant effect on the player's cheating profile, religion seems to play a major role in the decision to engage in dishonest behavior. Compared with Animists, we find that the probability of being an "occasional cheater" and an "always cheater", compared with being a "never cheater", was respectively reduced by 94% and 92% when the player was Muslim. Christian players, compared to Animists, are also 94% less likely to belong to the "always cheaters" category. These results align with field observations and qualitative discussions with players, as some cocoa middlemen who refused to take part in the actual game or indicated that they had never cheated during the game mainly mentioned religious considerations.

We further find that players' risk and time preferences influence cheating behavior. An inclination towards risk has a significant positive effect on the probability of being a "sometimes cheater" or an "always cheater", and stands as the main individual determinant of cheating. This finding can be expected and is consistent with the hypothesis that risk-lover individuals engage in more risky behavior

in order to maximize their payoffs. We further find a significant and negative relationship between the subjective discount factor and the probability of always cheating when facing a losing round, compared to never cheating. Conversely, this means that players who value the present (and therefore exhibit a low subjective discount factor) are more likely to cheat and belong to the category of always cheaters.

Cheating behavior could also be associated with various factors related to middlemen's activities, including their income as cocoa middlemen, their role as main middlemen, and the number of years of professional experience as cocoa middlemen. One might think that a middleman with more experience and a higher income would have less incentive to cheat to increase his potential earnings. However, our findings do not provide significant evidence supporting this hypothesis.

Whether the cheating behaviors measured in this study correlate with middlemen's real-life behavior remains uncertain. Without real-life traceability measures of dishonest behavior among middlemen, we cannot provide a definitive answer. However, middlemen who never cheated during the game are more likely to pay cocoa farmers by mobile money (23%) compared to those who always cheated (8%) (Table D3). Similarly, those in the 'never cheaters' category are significantly less likely to prefer cash to mobile money for confidentiality reasons (78%) compared to occasional cheaters (96%) and always cheaters (92%).

**Table 2:** Individual determinants of cheating

	Relative risk ratios	
	(1) Sometimes cheater	(2) Always cheater
<b>Socio-demographic characteristics</b>		
Age	0.940 (0.041)	0.900* (0.040)
Primary education	1.025 (1.140)	0.688 (0.672)
Middle school and higher	1.666 (2.019)	2.104 (2.108)
Ivorian	1.075 (1.217)	1.196 (1.383)
Christian	0.081 (0.118)	0.057* (0.078)
Muslim	0.059* (0.074)	0.077* (0.095)
<b>Middleman's activity characteristics</b>		
Income from cocoa	1.005 (0.131)	0.888 (0.106)
Main trader	0.330 (0.248)	0.640 (0.414)
Years of experience as cocoa middleman	1.029 (0.030)	1.044 (0.027)
<b>Risk and time preferences</b>		
Subjective discount factor	0.126 (0.138)	0.052** (0.053)
Risk-lover	8.033* (8.077)	8.594* (8.861)
Enumerators FE	Yes	Yes
Region FE	Yes	Yes
N	142	142
Pseudo R2	0.349	0.349

Source: Authors' calculation.

**Note:** Reported coefficients come from multinomial logistic regressions, using "Never cheater" as the reference category. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

## 6. Effect of observability and sanction among occasional cheaters

The second part of this paper seeks to investigate the effect of introducing observability and financial penalties on players' cheating behavior. We hypothesize that the likelihood of cheating decreases

under the following conditions: (1) the presence of an external observer (i.e., the experimenter) able to sometimes monitor players' dishonest behaviors, and (2) the potential of penalties if the experimenter detects cheating. Previous literature highlights that being observed may be enough to reduce players' cheating behavior, as they could be concerned about being seen as displaying socially undesirable behavior in the eyes of the experimenter (Bašić and Quercia, 2022; Fries et al., 2021). To our knowledge, no previous study tested the combined effect of observability and penalty.

## 6.1 Empirical specification

In the previous section, we observed that a non-negligible proportion of players never cheated to secure higher payoffs, even in the absence of observability and sanctions. On the other hand, a significant proportion of players always cheated, even in the presence of observability and sanctions, indicating that these players were not influenced by these game features, or were not sufficiently disincentive to alter their cheating behaviors. In this section, we are therefore interested in the sub-sample of players who exhibited heterogeneous cheating behavior across sessions, and how their behavior is influenced by the introduction of observability and financial penalties.

We perform the following conditional logistic regression with player fixed effects:

$$Cheating_{ir} = \beta_1 + \beta_2 Session2_i + \beta_3 Session3_i + \beta_4 \alpha_i + X_{ir} + \delta_i + \epsilon_i \quad (2.1)$$

where  $Cheating_{ir}$  is a dummy variable taking the value of 1 if the player  $i$  cheated during the round  $r$ , and 0 otherwise.  $Session2$  and  $Session3$  are dummies indicating the session being played. In some specifications, we also include  $\alpha$ , which is a vector of variables related to players' previous game experience.  $\delta_i$  stands for individual fixed-effects.  $X_{ir}$  is a set of round control variables, which include the round number and the lottery selected during the Holt and Laury first phase of the game, to respectively control for the probability of losing and the risk attitude. As a robustness check, we also provide results in the Appendix using a linear probability model (Table E1) and extending the same specification to the full sample (Table E2).

## 6.2 Findings

**Threat of observability and sanction.** We first explore whether the mere introduction of observability and observability coupled with penalties for cheating, inhibits dishonest behavior in our game. In session 2, players face the possibility of experimenters directly monitoring the true outcome of their die roll on tablets without penalty. After rolling the die on their tablets, players physically roll a die to determine whether the experimenter will observe their actual outcome. This risk is 50%, as the experimenter observes the player's reported result only if the second roll's outcome is even. Since not all individuals exhibit social-image concerns (resulting in being observed by a third party having no impact on their cheating behavior) (Huber et al., 2023), the combination of observation with a financial penalty could potentially alter the players' cost-benefit analysis of cheating (Thielmann and Hilbig, 2018). In session 3, players are thus confronted with the risk of getting a financial penalty if the experimenter observes a discrepancy between the number they reported and the actual outcome displayed on the tablet, following the same rules as in session 2.

Table 3 (column 1) presents findings indicating that introducing the mere risk of behavior monitoring decreases players' likelihood of cheating. Specifically, players are 41.5% less likely to cheat in session 2 than in session 1. This likelihood is further reduced when observability risk is combined with a penalty risk in the event of cheating, as they are 51.6% less likely to cheat in session 3 compared to session 1.

**Previous rounds history.** Considering the game's iterative nature with fifteen rounds, we argue that players' behavior and outcomes in previous rounds are likely to influence their decisions in subsequent rounds. We look at the player's game history in previous rounds, considering both previous losses and cases of cheating, to explore how these factors may influence cheating behavior in the current round.

The "previous loss rate" indicates the number of times a player has experienced losses relative to the total number of previous rounds, while "lost in previous rounds" is a dummy variable indicating whether the player had already lost at least once in previous rounds. A positive effect would suggest that players are likely to engage in cheating after a series of losing rounds, indicating a potential compensation effect to secure a win. Results in Table 3 (columns 2 and 3) reveal a positive but non-significant association between the previous loss history and the likelihood of cheating in the current round among occasional cheaters.

**Table 3: Effect of history loss, previous cheating, and game features on cheating**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Session 2	0.585* (0.158)	0.566* (0.158)	0.505* (0.177)	0.724 (0.221)	0.705 (0.240)					
Session 3	0.484** (0.132)	0.463** (0.135)	0.387* (0.160)	0.589 (0.179)	0.585 (0.202)	0.845 (0.233)	0.912 (0.289)	1.421 (0.460)	2.421* (0.930)	0.000 ( )
Previous loss rate		1.762 (2.200)								
Lost in previous rounds			1.336 (0.533)							
Previous cheat rate				0.566 (0.225)						
Cheated in previous rounds					0.683 (0.295)					
Previous observation rate						0.451 (0.326)				
Observed in previous rounds							0.723 (0.392)			
Previous cheating observation rate								0.058*** (0.046)		
Observed while cheating in previous rounds									0.113*** (0.060)	
Number of times players get a sanction										0.060*** (0.047)
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sessions	1, 2, 3	1, 2, 3	1, 2, 3	1, 2, 3	1, 2, 3	2, 3	2, 3	2, 3	2, 3	3
N	435	435	435	435	435	270	270	270	270	110
R2	0.136	0.137	0.138	0.141	0.138	0.110	0.106	0.166	0.182	0.331

Source: Authors' calculation.

Note: Reported coefficients are odd-ratios from conditional logistic regression. Controls include the selected lottery and round number. Standard errors in parenthesis. \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001.



Additionally, we consider the "previous cheat rate," which is the ratio of previous cheats to the number of times they faced a losing die, and a binary variable indicating whether the player has previously cheated in rounds. Although the relationship is not statistically significant, our results suggest a negative association between having cheated in previous rounds and the likelihood of cheating in the current round (Table 3, columns 4 and 5). Specifically, the probability of cheating decreases by 31.7% when the player has previously cheated at least once. This pattern could be explained by players feeling comfortable cheating occasionally but not frequently, perhaps fearing being perceived as regular cheaters by the experimenter or being concerned that cheating more often would increase their risk of getting caught.

**Previous observability and penalty.** We further hypothesize that previous game experiences, such as having been observed and having been observed while cheating, are likely to influence the player's cheating behavior in the current round. Therefore, we examine the influence of having been observed in previous rounds using the previous observation rate, which corresponds to the number of times a player has been observed relative to the number of rounds, and a dummy indicating whether the actual outcome of the players' die-roll has already been observed in previous rounds. Similarly, we look at the previous cheating observation rate, which is the number of times a player was observed cheating out of the number of times they cheated, and a dummy indicating whether the player has already been observed while cheating. When investigating the effect of having actually been observed by the experimenter in previous rounds (columns 6 and 7), whether they were observed cheating or reporting the true result, we find no significant impact on players' behavior in the current round. Although the relationship's direction aligns with expectations (i.e., less cheating when the player has been observed by the experimenter), it is not statistically significant. In contrast, we observe a negative and statistically significant relationship regarding the probability of cheating when players have been observed cheating in previous rounds (columns 8 and 9). Having been observed while cheating reduces the probability of cheating in the current round by 89%. When considering being observed cheating at least once in previous rounds, players appear to cheat more in session 3 than in session 2. Although not statistically significant, there is a shift in the relationship direction. We suggest that this might be explained by the fact that when players have already been observed cheating once, their concern about damaging their social image decreases, leading them to take any opportunity to increase their payoffs.

Finally, we explore the impact of having received a penalty during session 3, using the number of previous instances the player received a penalty (column 10). Financial penalties appear to decrease the likelihood of cheating when enforced. Specifically, our findings indicate that in the third session, the more penalties players have incurred in previous rounds, the less likely they are to cheat in the subsequent ones.

## 7. Conclusion

Pressure on cocoa-producing countries and companies to improve cocoa traceability has been growing steadily recently.<sup>8</sup> However, achieving traceability at the farmer level is a major challenge, given that the cocoa value chain in Côte d'Ivoire is made up of a large number of geographically dispersed players whose relationships are fraught with information asymmetry. Cocoa middlemen, the intermediaries linking cocoa farmers and cooperatives, play a key role in traceability as being the first to supply cocoa into the supply chain. However, middlemen are likely to engage in opportunistic behavior and misreport the nature of the cocoa they bring to the cooperative (whether certified or conventional), to increase their profits. Yet, middlemen are currently excluded from the monitoring and traceability mechanisms in place in the cocoa sector, despite their crucial role in the supply chain.

This paper relies on a lab-in-the-field involving an experimental cheating game with 151 cocoa middlemen operating in certified cocoa value chains, to investigate middlemen's dishonest behavior in situations of information asymmetry. This experimental approach, an adapted die-under-cup task, allows us to control the environment related to the decision to misbehave, thereby separating individuals' intrinsic (dis)honesty from other confounding motives. We argue that this approach provides an effective way to measure middlemen's propensity to engage in opportunistic behavior, given the challenges of using traditional observational or survey data, as these behaviors are hidden and subject to social desirability bias. Besides investigating whether middlemen do engage in opportunistic behavior and to what extent, we investigate the key determinants associated with such dishonest behavior. Our experimental design also provides an opportunity to assess whether middlemen behave

---

<sup>8</sup>Import restrictions on a number of commodities, including cocoa, are becoming stricter in consumer markets. For example, the European Union recently banned the import of commodities derived from deforestation (see the Deforestation Regulation (EU) 2023/1115, which came into effect on June 29th, 2023) and has agreed on a new regulation aimed at banning the import of products that have involved forced labor (see the Proposal for a regulation on prohibiting products made with forced labor (COM(2022)0453 – C9-0307/2022 – 2022/0269(COD))).

differently when they are observed and sanctioned for being caught cheating, mechanisms that could be implemented in real life as part of improved traceability monitoring.

We find that 78% of middlemen cheated at least once when they had a losing die, to maximize their payouts. Using individual data from each player over the 15 rounds, middlemen are classified into different categories based on their cheating behavior. We find that 59% of players consistently lie, while 19% do so occasionally. Remarkably, 22% of cocoa middlemen never lie, which is in line with previous experimental studies showing that some players consistently prioritize honesty, even when presented with opportunities to increase profits without consequences (Rustagi and Kroell, 2022; Kajackaite and Gneezy, 2017; Gerlach et al., 2019). Few individual characteristics, such as age, religion, and attitudes toward time and risk, influence cheating behavior. Further, we show that the simple threats of being observed and of getting a financial penalty for cheating induce a reduction in the probability of cheating among middlemen with heterogeneous cheating behaviors over the course of 15 rounds (i.e., the occasional cheaters). The increased costs associated with social image when there is a risk of being observed have thus a strong inhibiting effect, even when the observation has not actually taken place, and even in the absence of sanction. Being observed in previous rounds has no deterrent effect on cheating beyond the mere threat, except when players have been caught cheating and/or sanctioned for it. Financial penalties work, as the probability of cheating decreases considerably with the number of times players are sanctioned. However, whether this is due to a social sanctioning effect that discourages players from cheating after being caught once, concerns about further damaging their social image by repeatedly being observed cheating, or fears of increased penalties, remains an open question.

Whether monitoring systems should be extended to middlemen to reduce side-sourcing in the certified cocoa sector is a question that goes beyond the scope of this study. Our paper presents evidence of middlemen engaging in opportunistic behavior in a controlled environment, and presents some limitations. First, although representative of middlemen in the certified cocoa sector in Côte d'Ivoire, our sample is rather small, which may affect the statistical power of our results. One should also keep in mind that a few middlemen who attended the game's information sessions decided not to take part in the experiment, thus probably excluding from our sample the most risk-averse individuals. Second, while we provide evidence that observation and sanction mechanisms are effective in reducing dishonest behavior among certain types of players (the ones rational in the sense of Becker (1968), i.e., those adjusting their behavior toward honesty when the costs of cheating exceed the benefits), the actual proportion of middlemen falling into this category remains uncertain. Although this group

of "occasional cheaters" represents only 19% of the players in our experiment, suggesting that such monitoring and sanction mechanisms may be ineffective with most middlemen, it is likely that this proportion is inherent on game design. Indeed, the probability of being observed and the magnitude of financial penalties remain constant in the experiment. Yet, it is likely that a higher penalty or a greater likelihood of incurring penalties could have influenced the cost-benefit analysis for the "always cheaters" who might have fallen into the "occasional cheaters" category. Further, while labs-in-the-field allow for the unbiased measurement of cheating behavior without the influence of confounding factors, we acknowledge that predicting the real-life impact of monitoring and sanction mechanisms remains challenging. Various factors, including repeated interactions with farmers and cooperatives, may constrain middlemen's cheating behaviors in real-life scenarios. These dynamics introduce trust and reputational considerations which may influence middlemen's likelihood to cheat (Banerjee and Duflo, 2000; Macchiavello and Morjaria, 2015). In light of such complexities, it is likely that a greater proportion of "occasional cheaters", who are responsive to monitoring and sanctions, are found in real-life situations than what our experimental game may capture.

## References

- Abeler, J., Nosenzo, D., and Raymond, C. (2019). Preferences for Truth-Telling. *Econometrica*, 87(4):1115–1153. [54], [57], [61]
- Armand, A., Coutts, A., Vicente, P. C., and Vilela, I. (2023). Measuring corruption in the field using behavioral games. *Journal of Public Economics*, 218:104799. [57], [60]
- Banerjee, A. V. and Duflo, E. (2000). Reputation effects and the limits of contracting: A study of the Indian software industry. *The Quarterly Journal of Economics*, 115(3):989–1017. [79]
- Bašić, Z. and Quercia, S. (2022). The influence of self and social image concerns on lying. *Games and Economic Behavior*, 133:162–169. [54], [55], [73]
- Battigalli, P., Charness, G., and Dufwenberg, M. (2013). Deception: The role of guilt. *Journal of Economic Behavior & Organization*, 93:227–232. [69]
- Battigalli, P. and Dufwenberg, M. (2007). Guilt in games. *American Economic Review*, 97(2):170–176. [69]
- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of political economy*, 76(2):169–217. [54], [78]
- Bénabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *American economic review*, 96(5):1652–1678. [69]
- Bergquist, L. F. and Dinerstein, M. (2020). Competition and Entry in Agricultural Markets: Experimental Evidence from Kenya†. *American Economic Review*, 110(12):3705–3747. [57]
- Bergquist, L. F. and McIntosh, C. (2021). Search Cost, Intermediation, and Trade: Experimental Evidence from Ugandan Agricultural Markets. *CEGA Working Paper 173*, pages 1–58. [57]
- Bernard, T., Boutin, D., and Jouvin, M. (2024). Traceability along the value chain: the case of certified cocoa in Côte d'Ivoire. *Mimeo*. [54], [59]
- Bolton, G., Dimant, E., and Schmidt, U. (2021). Observability and social image: On the robustness and fragility of reciprocity. *Journal of Economic Behavior & Organization*, 191:946–964. [57]
- Casaburi, L. and Reed, T. (2022). Using Individual-Level Randomized Treatment to Learn about Market Structure. *American Economic Journal: Applied Economics*, 14(4):58–90. [57]
- Charness, G., Blanco-Jimenez, C., Ezquerra, L., and Rodriguez-Lara, I. (2019). Cheating, incentives, and money manipulation. *Experimental Economics*, 22(1):155–177. [55]
- Chatterjee, S. (2023). Market Power and Spatial Competition in Rural India. *The Quarterly Journal of Economics*, 138(3):1649–1711. [57]
- Dai, Z., Galeotti, F., and Villeval, M. C. (2018). Cheating in the lab predicts fraud in the field: An experiment in public transportation. *Management Science*, 64(3):1081–1100. [55]
- Fischbacher, U. and Föllmi-Heusi, F. (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, 11(3):525–547. [54], [55], [61]
- Fries, T., Gneezy, U., Kajackaite, A., and Parra, D. (2021). Observability and lying. *Journal of Economic Behavior & Organization*, 189:132–149. [55], [57], [73]

- Garbarino, E., Slonim, R., and Villeval, M. C. (2019). Loss aversion and lying behavior. *Journal of Economic Behavior and Organization*, 158:379–393. [55]
- Gerlach, P., Teodorescu, K., and Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological bulletin*, 145(1):1. [68], [70], [78]
- Gibson, R., Tanner, C., and Wagner, A. F. (2013). Preferences for truthfulness: Heterogeneity among and within individuals. *American Economic Review*, 103(1):532–548. [69]
- Gneezy, U., Kajackaite, A., and Sobel, J. (2018). Lying aversion and the size of the lie. *American Economic Review*, 108(2):419–453. [55], [57]
- Hanna, R. and Wang, S.-Y. (2017). Dishonesty and selection into public service: Evidence from india. *American Economic Journal: Economic Policy*, 9(3):262–290. [57]
- Harris, D., Borcan, O., Serra, D., Telli, H., Schettini, B., and Dercon, S. (2024). Proud to Belong: The Impact of Ethics Training on Police Officers in Ghana. [57], [60]
- Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655. [55], [61]
- Huber, C., Litsios, C., Nieper, A., and Promann, T. (2023). On social norms and observability in (dis) honest behavior. *Journal of Economic Behavior Organization*, 212:1086–1099. [54], [74]
- Huetz-Adams, F. and Fountain, A. C. (2022). 2022 Cocoa Barometer. Technical report, the Cocoa Barometer Consortium. [58]
- Imbens, G. W. and Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press. [91]
- Jiang, S. and Villeval, M. C. (2022). Dishonesty as a collective-risk social dilemma. *Available at SSRN*. [57]
- Kajackaite, A. and Gneezy, U. (2017). Incentives and cheating. *Games and Economic Behavior*, 102:433–444. [55], [57], [69], [78]
- Khalmetski, K. and Sliwka, D. (2019). Disguising lies - Image Concerns and Partial Lying in Cheating Games. *American Economic Journal: Microeconomics*, 11(4):79–110. [54]
- Macchiavello, R. and Morjaria, A. (2015). The value of relationships: evidence from a supply shock to kenyan rose exports. *American Economic Review*, 105(9):2911–2945. [79]
- Mitra, S., Mookherjee, D., Torero, M., and Visaria, S. (2018). Asymmetric information and middleman margins: An experiment with Indian potato farmers. *Review of Economics and Statistics*, 100(1):1–13. [54], [60]
- Renier, C., Vandromme, M., Meyfroidt, P., Ribeiro, V., Kalischek, N., and Zu Ermgassen, E. K. (2023). Transparency, traceability and deforestation in the ivorian cocoa supply chain. *Environmental Research Letters*, 18(2):024030. [57], [58]
- Rubinstein, A. and Wolinsky, A. (1987). Middlemen. *The Quarterly Journal of Economics*, 102(3):581–593. [57]
- Rustagi, D. and Kroell, M. (2022). Measuring honesty and explaining adulteration in naturally occurring markets. *Journal of Development Economics*, 156:102819. [54], [57], [59], [60], [61], [68], [78]
- Stoop, P., Ramanan, N., Geens, H., Lambrecht, A., and Dekeister, S. (2021). Technical brief on cocoa traceability in West and Central Africa. [57], [59], [60]

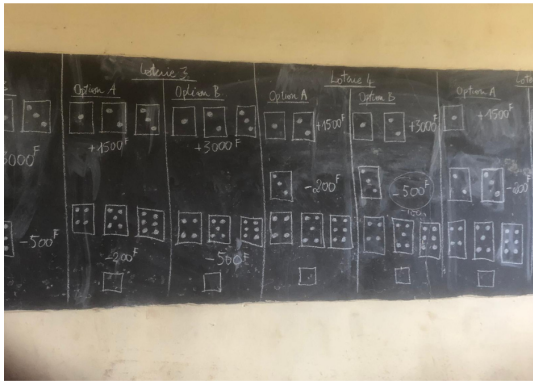
Thielmann, I. and Hilbig, B. E. (2018). Daring dishonesty: On the role of sanctions for (un) ethical behavior. *Journal of Experimental Social Psychology*, 79:71–77. [55], [74]

World Bank (2019). Au pays du cacao comment transformer la Côte d'Ivoire. Technical report, The World Bank. [58]

## Appendix

### Appendix A : Lab-in-the-field experiment information

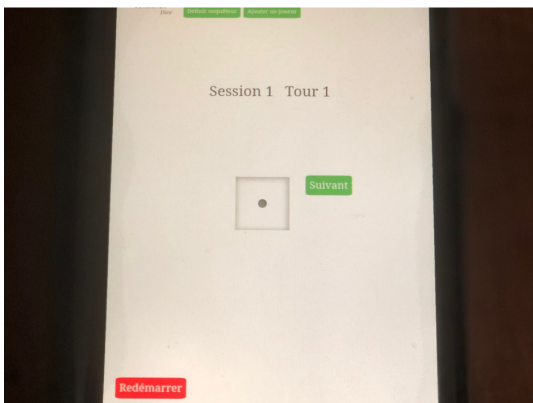
**Figure A1:** Lab-in-the-field experiment setting and tools



Visual aids displaying the amounts with each decision



Typical session



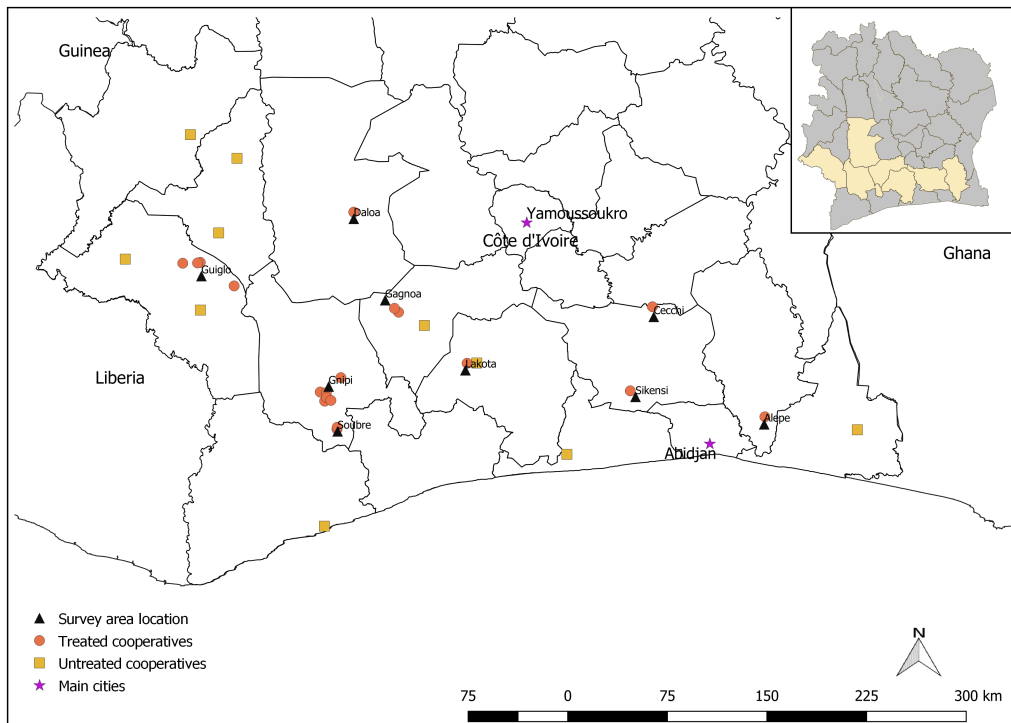
Game for player



Observation and penalty tools



Figure A2: Survey areas



Source: Authors.

## Experimenter sheet

Player n°

### Lottery choices

Report here the options chosen by the player.

	Option A	Option B
Lottery 1	<input type="checkbox"/>	<input type="checkbox"/>
Lottery 2	<input type="checkbox"/>	<input type="checkbox"/>
Lottery 3	<input type="checkbox"/>	<input type="checkbox"/>
Lottery 4	<input type="checkbox"/>	<input type="checkbox"/>
Lottery 5	<input type="checkbox"/>	<input type="checkbox"/>

Double-check with the player to ensure there are no mistakes before proceeding to the next part.

### Session 1 : Basic die roll

The player rolls a die on the tablet. Record the number declared by the player here. Repeat the procedure for each round.

Round number	Outcome declared by the player
Round 1	
Round 2	
Round 3	
Round 4	
Round 5	

The player will draw the round number selected for payment.

The round selected for payment is round number:

For this session, the middleman lost/won:


## Session 2 : Die roll with observation risk

The player rolls a die on the tablet. Record the number declared by the player here. Repeat the procedure for each round.

The player will now roll a physical die: if the outcome is an even number (2, 4, or 6), move to observe the roll-die outcome on the player's screen and indicate the observed number in the second column. If the roll-die outcome is an odd number (1, 3, or 5), do not move and make a cross in the corresponding box.

Round number	Outcome declared by the player	Outcome observed by the enumerator (if 2, 4, 6)
Round 1		
Round 2		
Round 3		
Round 4		
Round 5		

The player will draw the round number selected for payment.

The round selected for payment is round number:


For this session, the middleman lost/won:

## Session 3 : Die roll with observation and penalty risks

The player rolls a die on the tablet. Record the number declared by the player here. Repeat the procedure for each round.

The player will now roll a physical die: if the outcome is an even number (2, 4, or 6), move to observe the roll-die outcome on the player's screen and indicate the observed number in the second column. If the roll-die outcome is an odd number (1, 3, or 5), do not move and make a cross in the corresponding box.

Repeat the procedure for each round.

Round number	Outcome declared by the player	Outcome observed by the enumerator (if 2, 4, 6)
Round 1		
Round 2		
Round 3		
Round 4		
Round 5		

The player will draw the round number selected for payment.

The round selected for payment is round number:


For this session, the middleman lost/won:

## CONSENT FORM

### To be read to the respondent before the experimental game

The experimental game will mainly focus on your risk attitude. No specific skills are required for participating in this game. Your game choices and payouts will remain confidential. The experiment will be divided into 4 parts. In the first part, you will choose the lotteries you prefer to play. There are no correct or incorrect choices; you should make your decisions based on your best judgment. In the second, third, and fourth parts, you will play the game based on the lotteries you selected by rolling a die. A test game will be conducted to explain the rules, during which you can ask any questions you may have. Instructions will be reiterated at the beginning of each session.

An experimenter will be present to record your lottery choices in the first part and the results of your die rolls in the second, third, and fourth parts of the game. At the end of each game session, one round will be randomly selected for payment. At the end of the experiment, your payouts will be paid on your Mobile Money account. Your payouts are contingent upon the lotteries you choose in the first part of the experiment, and the random outcomes of your die rolls. You have already received 6,000 FCFA for your participation in the survey. By participating in this game, you can earn a maximum of an additional 9,000 FCFA. Nevertheless, there is a possibility that you may incur losses; in the worst-case scenario, you could lose up to 6,000 FCFA.

The information you provide will help us gain a better comprehension of cocoa middlemen activity and, consequently, a more comprehensive understanding of the cocoa value chain in Côte d'Ivoire. The information we collect will remain confidential and will under no circumstances be disclosed to third parties or to the cooperative. The data collected will be encrypted and anonymised by researchers from the University of Bordeaux, within the strict limits of research activities and with due respect for individual privacy. The data and encryption key will not be communicated to any third parties. Survey findings will be published in academic journals.

Your participation in this experiment is entirely voluntary. You have the option to decline participation, and you can stop the experiment at any point without any consequences. If you have questions about the study at any time, you may reach out for clarifications.

In compliance with applicable laws and regulations, you have the right to access, rectify, obtain a copy of, or delete your data, as well as to restrict its processing or object to it. You can use these rights by contacting us via email at the following address: [REDACTED]@u-bordeaux.fr.

**CONSENT FORM (2 copies, one for the participant, one for the experimenter)**

*Note: This consent form was originally written in French and will be provided to study participants after translation into the local language.*

I, the undersigned \_\_\_\_\_ (the participant), confirm that:

- I have been informed of the above details, which have been thoroughly explained to me.
- I have a clear understanding of this information.
- I have had the opportunity to seek clarification and received satisfactory answers to all my questions.
- My participation in the experiment is entirely voluntary.
- I have received a copy of this informed consent form.
- I am aware that I can contact one of the study's coordinator [REDACTED] if I have any questions about the study or if I wish to stop the experiment.

Date /\_\_/\_/\_\_/\_/







Location \_\_\_\_\_







Player signature
------------------





Signature of the person who obtained consent
--

## Player n°

Which of these two options do you prefer?  
Tick the one you prefer and let the enumerator know.

Lottery 1	
Option A	Option B
 	 
1 500 FCFA	3 000 FCFA
	
-1 000 FCFA	-2 000 FCFA
<input type="checkbox"/> Option A	<input type="checkbox"/> Option B

Lottery 2	
Option A	Option B
 	 
1 500 FCFA	3 000 FCFA
	
-1 000 FCFA	-2 000 FCFA
<input type="checkbox"/> Option A	<input type="checkbox"/> Option B

Lottery 3	
Option A	Option B
 	 
1 500 FCFA	3 000 FCFA
-1 000 FCFA	-2 000 FCFA
<input type="checkbox"/> Option A	<input type="checkbox"/> Option B

## Lottery 4

Option A



1 500 FCFA



-1 000 FCFA



Option A

Option B



3 000 FCFA



-2 000 FCFA



Option B

## Lottery 5

Option A



1 500 FCFA



-1 000 FCFA



Option A

Option B



3 000 FCFA



-2 000 FCFA



Option B

## Appendix B: Test for selection bias in game participation

Table B1 aims to detect potential self-selection biases between middlemen who agree to participate in the game (151 individuals) and those who opted out (64 individuals). We compare the two groups' key demographic, work-related, and time-preference variables. For each variable, the table presents the number of observations (N), mean value, and standard deviation for both groups. Subsequent columns highlight the raw difference in means between the two groups, the normalized difference according to Imbens and Rubin (2015) for assessing the magnitude of this difference independent of sample size,<sup>9</sup> and a p-value to test the statistical significance of the observed differences.

On average, those who refused to participate are older than those who accepted, with a normalized difference of 0.45 and a p-value of 0.04, and less educated, with a normalized difference of -0.27 and a p-value of 0.04. There is no significant difference in terms of marital status and religion. We cannot find any difference in terms of work-related characteristics and time preferences, as all the variables exhibit low normalized differences and high p-values. However, we do find a significant difference between participants and non-participants in terms of risk attitudes. Non-participants are more likely to be risk-averse, positively correlated with never cheating in our regressions. Thus, we are likely to overestimate the proportion of players cheating consistently. Therefore, our analysis results are a lower bound of honest behavior.

---

<sup>9</sup>Normalized difference is used to assess the balance of observed covariates between two distinct groups. The advantage of using normalized differences over traditional p-values from statistical tests (such as t-tests) is that sample size does not affect normalized differences. Traditional tests might fail to detect imbalances in small samples, like ours, due to low statistical power.



**Table B1:** Balancing test comparing individuals who accepted vs. those who declined participation in the game

	(1)			(2)			(2) - (1)		
	Accepted to participate	Refused to participate	Difference	Accepted to participate	Refused to participate	Difference	Raw diff.	Norm. diff.	P-value
	N	Mean	St. Dev.	N	Mean	St. Dev.	Raw diff.	Norm. diff.	P-value
<b>Demographics</b>									
Age	151	39.40	9.16	64	43.66	9.71	4.26	0.45	0.04
Married	151	0.92	0.27	64	0.97	0.17	0.05	0.22	0.50
Educated	151	0.83	0.37	64	0.72	0.45	-0.11	-0.27	0.04
Christian	151	0.46	0.50	64	0.33	0.47	-0.13	-0.27	0.10
Muslim	151	0.40	0.49	64	0.48	0.50	0.08	0.16	0.27
<b>Work</b>									
Years of experience as a trader	151	21.58	15.07	64	20.59	15.92	-0.99	-0.06	0.23
Worked in the cocoa sector before being a trader	151	0.90	0.30	64	0.87	0.33	-0.03	-0.06	0.08
Receive help from coxeurs	151	0.40	0.49	64	0.36	0.48	-0.04	-0.08	0.65
Delivered cocoa volume (tons)	151	89.51	95.92	64	86.65	111.40	-2.86	-0.03	0.63
Overall declared income (ln)	133	13.85	1.01	51	13.78	1.06	-0.07	-0.07	0.16
Always experience cash shortages	151	0.30	0.46	64	0.33	0.47	0.03	0.06	0.95
Provide credit to farmers	151	0.92	0.27	64	0.91	0.29	-0.01	-0.04	0.46
Own storage facility	149	0.41	0.49	64	0.53	0.50	0.12	0.24	0.10
<b>Time and risk preferences</b>									
Subjective discount factor	151	0.59	0.32	64	0.58	0.33	-0.01	-0.03	0.59
Risk-lover	151	0.46	0.50	23	0.09	0.29	-0.37	-0.91	0.00

Source: Authors' calculation.

**Note:** Norm. Diff. stands for normalized differences and were assessed using t-tests following the method of Imbens and Rubin (2015). Values reported for the "risk-adverse" variable for players who refused to take part in the game are computed on the sub-sample of players (n=23) who agreed to take part in the test but refused to play the game.

## Appendix C: Descriptive statistics

**Table C1:** Socio-economic characteristics of traders

	Obs.	Mean (%)	Std. Dev.
<b>Socio-demographic characteristics</b>			
<i>Trader's age</i>			
Age	151	39.39	9.16
35 years or less	151	35.10	
Between 36 and 45 years	151	41.72	
More than 45 years	151	23.18	
<i>Marital status</i>			
Married/in-relationship	151	92.05	
Widower	151	1.32	
Single	151	6.62	
<i>Education level</i>			
No education	151	19.21	
Primary education	151	27.15	
Middle school	151	24.50	
High school or higher	151	29.14	
<i>Household size</i>			
Number of adults	151	4.48	3.48
Number of children	151	4.40	2.60
<i>Religion</i>			
Christian	147	46.94	
Muslim	147	41.50	
Animist	147	11.56	
<b>Risk and time preferences</b>			
Risk-lovers	151	38.41	
Subjective discount factor	151	0.59	0.32
<b>Working characteristics</b>			
# of years since trader started working for current cooperative	151	4.30	2.54
# of years of experience as cocoa trader	151	21.58	15.07
Had worked in the cocoa sector before its current position	145	93.79	
# of years of experience in the cocoa sector	151	24.42	8.20
Trader is also a cocoa farmer	151	90.07	
Cocoa trader as main activity	151	39.73	

Source: Authors' calculation.

**Table C2:** Characteristics of cocoa traders' activities

	Obs.	Mean (%)	Std. Dev.
<b>Business size</b>			
Number of farmers	151	113.67	85.84
Cocoa delivered during 20-21 cocoa season (in tons)	151	85.57	84.33
Use of assistants (coxeurs)	151	40.40	
<b>Performance</b>			
No objective set by the cooperative	151	49.01	
Failed to meet objectives	77	89.61	
Met or exceeded objectives	77	10.39	
<b>Income</b>			
Commission per kilos during main crop (CFA)	151	34.07	26.94
Income from cocoa trader's activity (CFA 1 000)	146	1,430	1,362
<b>Costs</b>			
Fuel	151	88.08	
Vehicle leasing or purchase	151	28.48	
Vehicle maintenance	151	17.88	
Road tolls	151	66.89	
Driver salaries	151	66.22	
Coxeurs salaries	151	32.45	
Average costs (CFA 1 000)	146	938.47	990.60
<b>Challenges in buying cocoa</b>			
Faced difficulties in the last 12 months	151	78.81	
Lack of operational means of transport	119	46.22	
Poor road conditions	119	76.47	
Farm-gate price fluctuations	119	70.59	
Lack of money to buy cocoa	119	30.25	
<b>Cash shortage</b>			
The cooperative does not give cash advances	151	29.80	
Always cash shortage	151	30.46	
Often cash shortage	151	25.83	
Sometimes cash shortage	151	0.66	
Rarely cash shortage	151	1.99	
<b>Cocoa storage</b>			
Own/rent a warehouse to stock cocoa	151	41.72	
Use the cooperative's warehouse	151	19.21	
Use storage facilities provided by the cooperative (in communities)	151	39.07	

Source: Authors' calculation.

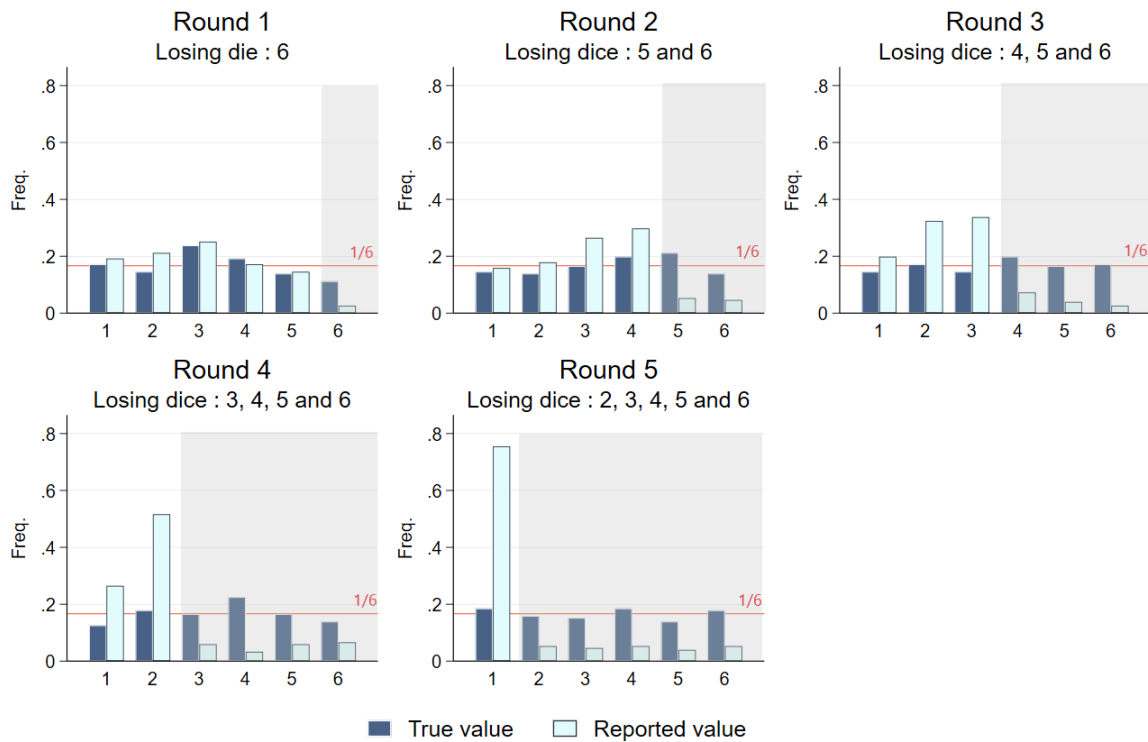
**Table C3:** Traders' relationship with farmers

	Obs.	Mean (%)
<b>Loans</b>		
Provides loans to farmers	151	92.05
<b>Proportion of farmers to whom the trader lends money</b>		
0-25%	139	35.25
More than 25% and 50%	139	47.48
More than 50% and 75%	139	11.51
More than 75%	139	5.76
<b>Reasons for lending money to farmers</b>		
School fees	139	89.21
Agricultural input purchases	139	55.39
Ceremonial purposes (e.g. funeral)	139	46.76
Medical costs	139	85.61
Food purchases	139	28.06
<b>Mobile money</b>		
Trader sometimes pays farmers by mobile money	141	12.06

Source: Authors' calculation.

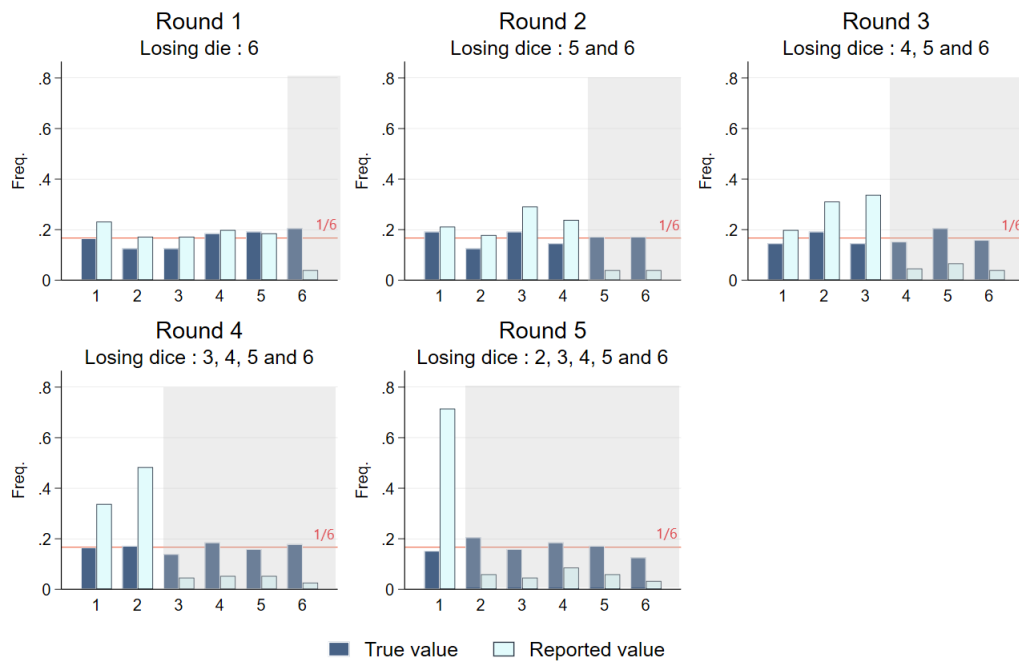
## Appendix D: Additional experimental results

**Figure D1:** Distribution of true die roll outcomes and outcomes declared by players (Session 2)



Source: Authors' calculation.

**Note:** The red horizontal line indicates the theoretical distribution of a random draw. This graph shows the distribution of actual and reported roll-dice outcomes for session 2, where players face the risk of observability.

**Figure D2:** Distribution of true die roll outcomes and outcomes declared by players (Session 3)

Source: Authors' calculation.

**Note:** The red horizontal line indicates the theoretical distribution of a random draw. This graph shows the distribution of actual and reported roll-dice outcomes for session 3, where players face the risk of observability and financial penalty in case of cheating.

**Table D1:** Effect of the probability of losing on unconditioned cheating, using linear probability model (LPM)

	(1)	(2)
Round 2	0.113***	
$P(\text{loss})=2/6^a$	(3.65)	
Round 3	0.212***	
$P(\text{loss})=3/6^a$	(6.87)	
Round 4	0.327***	
$P(\text{loss})=4/6^a$	(10.60)	
Round 5	0.419***	
$P(\text{loss})=5/6^a$	(13.60)	
Theoretical losing probability <sup>b</sup>		0.632*** (15.28)
N	2,265	2,265
R2	0.094	0.093

Source: Authors' calculation.

**Note:** Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

<sup>a</sup> : Reference category : Round 1, where  $P(\text{loss})=1/6$ .

<sup>b</sup> : The theoretical losing probability is 0.167 for round 1, 0.333 for round 2, 0.500 for round 3, 0.667 for round 4, and 0.833 for round 5.

**Table D2:** Payoffs distribution according to players' cheating profile

Payoffs (FCFA)	Never cheaters (%)	Occasional cheaters (%)	Always cheaters (%)	Total (%)
-4,000	3.03	0.00	0.00	0.66
-3,000	3.03	0.00	0.00	0.66
-1,500	3.03	3.45	0.00	1.32
-1,000	18.18	6.90	0.00	5.30
-500	3.03	3.45	0.00	1.32
0	9.09	0.00	0.00	1.99
1,000	6.06	0.00	4.49	3.97
2,000	6.06	0.00	0.00	1.32
2,500	0.00	13.79	4.49	5.30
3,500	6.06	6.90	0.00	2.65
4,000	18.38	17.24	12.36	14.57
4,500	0.00	10.34	7.87	6.62
5,000	12.12	3.45	0.00	3.31
6,000	00.00	6.90	11.24	7.95
7,500	3.03	0.00	19.10	11.92
9,000	9.09	27.59	40.45	31.13
Total	100.00	100.00	100.00	100.00

Source: Authors' calculation.

**Table D3:** Differences in payment modes with farmers by players' cheating profiles

Variable	(1) Never cheaters		(2) Sometimes cheaters		(3) Always cheaters		T-test P-value
	N	Mean/SE	N	Mean/SE	N	Mean/SE	
Pay farmers through mobile money	30	0.233 (0.079)	27	0.111 (0.062)	84	0.083 (0.030)	0.032**
Prefer cash for confidentiality reasons	23	0.783 (0.088)	24	0.958 (0.042)	77	0.922 (0.031)	0.062*

Source: Authors' calculation.

Note: The value displayed for t-tests are p-values. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent critical level.



## **Appendix E: Robustness checks**

Table E1: Linear Probability Model (LPM) on the subsample of occasional cheaters

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Session 2	-0.103* (0.052)	-0.108* (0.054)	-0.130* (0.065)	-0.071 (0.060)	-0.089 (0.066)					
Session 3	-0.138** (0.052)	- 0.144** (0.054)	-0.176* (0.077)	-0.107 (0.059)	-0.122 (0.069)	-0.025 (0.052)	-0.015 (0.062)	0.054 (0.056)	0.126* (0.063)	0.000 (.)
Previous loss rate		0.068 (0.207)								
Loss in previous rounds			0.050 (0.076)							
Previous cheat rate				-0.083 (0.077)						
Cheated in previous rounds					-0.028 (0.081)					
Previous observation rate						-0.126 (0.124)				
Observed in previous rounds							-0.055 (0.097)			
Previous cheating observation rate								0.433*** (0.124)		
Observed while cheating in previous rounds									- 0.358*** (0.088)	
Number of times players get a sanction										- 0.376*** (0.111)
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sessions	1, 2, 3	1, 2, 3	1, 2, 3	1, 2, 3	1, 2, 3	2, 3	2, 3	2, 3	2, 3	3
N	435	435	435	435	435	270	270	270	270	110
R2 (within)	0.129	0.129	0.130	0.131	0.129	0.095	0.092	0.133	0.147	0.192

Source: Authors' calculation.

Note: Controls include the selected lottery and round number. Standard errors in parenthesis. \* p &lt; 0.05, \*\* p &lt; 0.01, \*\*\* p &lt; 0.001.

**Table E2: Linear Probability Model (LPM) on overall sample**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Session 2	-0.023 (0.021)	-0.012 (0.022)	0.005 (0.023)	-0.023 (0.024)	-0.023 (0.025)					
Session 3	-0.011 (0.021)	0.001 (0.022)	0.022 (0.023)	-0.011 (0.024)	-0.011 (0.025)	0.009 (0.022)	-0.002 (0.026)	0.005 (0.023)	0.039 (0.026)	0.000 (.)
Previous loss rate		-0.152 (0.090)								
Loss in previous rounds			-0.117** (0.037)							
Previous cheat rate				0.000 (0.031)						
Cheated in previous rounds					0.002 (0.034)					
Previous observation rate						0.026 (0.051)				
Observed in previous rounds							0.038 (0.039)			
Previous cheating observation rate								0.034 (0.050)		
Observed while cheating in previous rounds									-0.063 (0.037)	
Number of times players get a sanction										- 0.153*** (0.042)
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample: Sessions kept	1, 2, 3	1, 2, 3	1, 2, 3	1, 2, 3	1, 2, 3	2, 3	2, 3	2, 3	2, 3	3
N	2,265	2,265	2,265	2,265	2,265	1,510	1,510	1,510	1,510	755
R2 (within)	0.125	0.126	0.129	0.125	0.125	0.123	0.123	0.123	0.124	0.145

Source: Authors' calculation.

Note: Controls include the selected lottery and round number. Standard errors in parenthesis. \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001.





---

# TRACEABILITY ALONG THE VALUE CHAIN: THE CASE OF CERTIFIED COCOA IN CÔTE D'IVOIRE

*Co-authored with Tanguy Bernard and Delphine Boutin.*

## **Abstract.**

This study investigates the traceability of environmentally and socially sustainable (ESS) certified cocoa in Côte d'Ivoire, the world's leading cocoa-producing country. Despite ESS certification schemes aimed at ensuring sustainable production practices, significant information asymmetries persist within the value chain, leaving room for opportunistic behaviors. Side-sourcing, i.e. when conventional cocoa is labeled as certified, represents a potential issue. We explore reporting discrepancies regarding transactions of allegedly certified cocoa, using survey data from farmers and middlemen, and administrative data provided by cooperatives. We find large discrepancies between volumes declared by farmers and cooperatives across all cooperatives in our sample, indicating widespread reliance on side-sourcing. We further assess the impact of a new digital traceability system (DTS), implemented through a randomized controlled trial, designed to reduce opacity in transactions between farmers and middlemen (and hence, cooperatives). Preliminary findings show low adoption of the DTS, and suggest that it impacts both treated cooperatives' sourcing and misreporting strategies. Overall, the findings revealed strong resistance to such a traceability tool among supply chain actors, and that its implementation does not lead to a reduction in side-sourcing, but rather to modified data misreporting practices.

## 1. Introduction

In high-income countries, import regulations increasingly request evidence that agricultural commodities entering their markets are produced under environmentally and socially sustainable conditions (hereafter ESS).<sup>1</sup> Because ESS attributes cannot be inferred from the product itself, as it is a characteristic invisible to the naked eye, importers must ensure that commodities are traceable to their original producers and that they genuinely monitor, incentivize, and possibly support the latter's compliance with ESS conditions of production.

Where value chains are vertically integrated from production to exports, compliance with ESS regulations is a matter of intra-firm management and organization. Examples of such value chains include Ethiopian cut-flowers or Senegalese cherry tomatoes (Mengistie et al., 2017; Meemken et al., 2021). Many agricultural value chains in low-income countries are however, long and complex, involving several layers of actors responsible for collecting and gradually aggregating crops originally produced by numerous smallholder farmers. Some "certified" producers receive support, are to some extent monitored, and obtain a price premium for their adherence to ESS production practices. However, information asymmetries throughout the value chains provide room for opportunistic side-sourcing behavior, wherein (part of) the products sold as "certified" may originate from non-certified farmers. This situation portrays the cocoa value chain in Côte d'Ivoire, which accounts for 40% of global annual production and where 60% of farmers are now linked to certification standards (World Bank, 2019). Cocoa stands as a key commodity targeted by the 2023 European Deforestation Regulation, thus increasing pressure on the cocoa sector to achieve traceability at the farmer level.

In this paper, we study the extent to which ESS-certified cocoa in Côte d'Ivoire genuinely originates from ESS-certified farmers. Although side-sourcing is well-known to most actors in the cocoa industry (Stoop et al., 2021), its magnitude remains uncertain due to the lack of appropriate data and to challenges in detecting and measuring such hidden behaviors aimed at mislabeling agricultural goods with credence characteristics (Boutin et al., 2024; Rustagi and Kroell, 2022). As a result, the proportion of cocoa reaching high-income markets that is produced following ESS production practices is unclear. In partnership with a global top 5 cocoa exporter, we gathered an unusually comprehensive set of

---

<sup>1</sup>On April 19th, 2023 for instance, the European Parliament enacted a new law prohibiting the import of a range of commodities, including cocoa, if associated with deforestation or forest degradation – see the Deforestation Regulation (EU) 2023/1115, which took effect on June 29th, 2023.

survey-based and administrative data, enabling us to compare information on cocoa transactions reported by farmers, rural middlemen who purchase cocoa on behalf of cooperatives, and cooperatives themselves. We infer the magnitude of side-sourcing by analyzing reporting discrepancies in transaction data between these different value chain players. We find very large and systematic discrepancies between the volume of cocoa reported by surveyed farmers and the corresponding figures recorded in cooperatives' registers, ranging between 123% to 870% of the farmers' reports in our sample of cooperatives. In contrast, we find no meaningful differences between farmers' reports and those obtained from middlemen. While we cannot entirely rule out the possibility of opportunistic behaviors by farmers and middlemen, our evidence suggests significant data manipulation by cooperatives, as indicated by inconsistencies across two versions of their traceability registers, the mid- and end-of-season ones, without any relevant justification for the observed changes.

Next, we assess the impact of a newly introduced digital traceability system (DTS) on cooperatives sourcing and misreporting strategies. This system assigns each "certified" farmer a unique ID card, which middlemen must flash using a dedicated mobile application before entering transaction details, such as date and volume. The resulting data are then accessible to middlemen, cooperatives, and the exporting company, thereby limiting room for discrepancies across these entities. Our analysis relies on a purposefully designed randomized controlled trial, wherein a random subset of sections (and thus, middlemen) was selected (treated) from a random subset of (treated) cooperatives to use the system in the first season of its implementation. Within these treated cooperatives, we further varied the share of sections selected to use the DTS: either one-third or two-thirds. Many middlemen and cooperatives expressed reluctance to use the DTS, resulting in a rather low take-up rate. While some cooperatives have digitized up to 58% of their certified cocoa purchases in the traceability tool, other treated cooperatives have barely digitized any.

Finally, we estimate the intent-to-treat effects of DTS on cooperatives sourcing and misreporting strategies. First, we show that DTS implementation has led to a shift in cocoa sourcing at the cooperative level, with middlemen shifting from farmers in treated sections to purchasing cocoa from farmers located in control ones. Second, treated cooperatives report more volume and number of transactions in their traceability registers compared to control cooperatives, particularly at the beginning of the season (October/November), which is in line with field observations that this period is most prone to side-sourcing activities. Finally, we compare transactions reported in the first and final traceability registers at the beginning of the season (October-December), which should remain unchanged across



both traceability registers, to assess whether cooperatives make fraudulent modifications to transaction data during the season. We find that DTS implementation increases data manipulation of traceability registers at cooperative-level, but cooperatives engage in different misreporting strategies depending on treatment intensity. One-third treated cooperatives modify less the transactions initially reported in their first traceability registers, resulting in more transactions being matched between the first and final traceability registers, as their data modification strategy relies mainly on adding new (potentially fraudulent) transactions. Two-thirds treated cooperatives add a moderate number of new transactions between their first and final traceability registers, but rather modify existing transactions reported in the first traceability register to adjust their volumes between mid- and end-season traceability registers. This leads to a lower proportion of transactions matching between their first and final registers.

This study contributes to the literature on agricultural certification norms and standards. Previous research has mainly focused on the analysis of actors, processes, and power dynamics in the development and implementation of these standards (Lambin and Thorlakson, 2018; Oya et al., 2018; Ponte and Gibbon, 2005; Taylor, 2005), or on their impacts on farmers (DeFries et al., 2017; Meemken, 2020; Sellare et al., 2020; Akoyi and Maertens, 2018). To the best of our knowledge, no research has yet investigated whether these certification schemes actually guarantee that labeled goods comply with certification criteria, nor the existence of opportunistic behaviors from certified value chain actors induced by information asymmetries regarding credence characteristics of goods. Previous research highlighted side-selling issues in the context of agricultural cooperatives, where farmers sold part of their production to buyers other than the cooperative they are committed to (Hao et al., 2018; Shumeta et al., 2018; Wollni and Fischer, 2015). However, side-sourcing, i.e. the fact that conventional cocoa beans are sold as certified, is another source of contract failure at every supply chain stage in the context of certified cooperatives, and remains under-researched. Thus, the first contribution of this paper is to provide firsthand evidence of the existence and magnitude of side-sourcing in the certified cocoa value chain of the world's first-producing country.

In detecting fraudulent behaviors in the most opaque segment of the cocoa supply chain (Stoop et al., 2021), our work aligns with recent empirical studies showing that agricultural intermediaries tend to exploit market failures, such as low competition in the market or price information asymmetries, to increase their profit margins (Bergquist and Dinerstein, 2020; Bergquist et al., 2021; Casaburi and Reed, 2022). We provide evidence that intermediaries in the cocoa value chain (cooperatives and, to a lesser extent, middlemen) engage in dishonest behavior on a large scale, exploiting information asymmetries

on credence characteristics. More broadly, our research contributes to the literature aiming at detecting and measuring fraudulent behavior among agricultural value chain players (Boutin et al., 2024; Rustagi and Kroell, 2022), and on the use of both survey and administrative data to detect corporate fraud (Almunia et al., 2024) in developing countries.

Finally, this paper contributes to the literature on traceability systems in agricultural value chains. Many initiatives aimed at increasing information transparency along the supply chain, and thus traceability, have been introduced in recent years (Gardner et al., 2019). New digital technologies are considered to facilitate transparency by providing reliable, secure, and accurate information about the supply chain and sustainable practices (Astill et al., 2019; Bai et al., 2022; Gualandris et al., 2021). In the cocoa sector, digital traceability tools are becoming increasingly popular. They are receiving growing attention from industry players, as they are considered one of the main ways of improving cocoa traceability (Nitidae and EFI, 2021; Stoop et al., 2021). This paper provides empirical evidence of supply chain actors' adherence to such tools and their impact on the behavior of upstream supply chain players.

This paper is organized as follows. Section II provides contextual information on the certified cocoa supply chain in Côte d'Ivoire and Section III describes the data. Section IV examines the extent of side-sourcing in the ESS-certified cocoa value chain in Côte d'Ivoire. Section V presents the randomized controlled trial design implemented to assess the impact of the DTS on supply chain actors' behaviors. Section VI discusses compliance and resistance to the DTS and Section VII presents the effects of the DTS implementation on cooperatives sourcing and misreporting strategies. Section VIII concludes.

## **2. Context**

Côte d'Ivoire stands as the world's leading cocoa producer, supplying about 2 million tonnes each year, or 40% of the world's annual production (World Bank, 2019). Its capacity to respond to the new regulations of main importing countries is mediated by the organization of its domestic value chains and the possible implementation of traceability systems, which we describe below.

## 2.1 Cocoa value chain in Côte d'Ivoire

Ivorian cocoa is produced by an estimated 600,000 small-scale farmers (International Cocoa Organization, 2024). As described in Table 1, the average cocoa farmer cultivates less than 3 hectares of cocoa, yielding less than 600 kilos of dried beans per hectare, from which the farmer earns an average annual income per hectare of 440,920 FCFA - roughly equivalent to \$ 707 (PPP).

**Table 1:** Descriptive statistics of cocoa value chain actors

	25th percentile	50th percentile	75th percentile	Mean
<b>Cooperatives</b>				
Number of farmers	886	1,357	1,781	1,392
Number of sections	5	8.5	12	9
<b>Sections</b>				
Number of farmers	67	127	202	149
<b>Farmers</b>				
Age	34	40	49	42
Household size	4	5	7	6
Cocoa farm size (ha)	1.46	2.21	3.33	2.72
Yields	296.48	485.42	771.05	588.85
Annual income/hectare	221,747	363,092	577,002	440,920
Distance from farm to road (km)	1.38	2.92	4.91	3.54
Distance from farm to cooperative (km)	11.20	18.93	27.55	20.80

Source: Authors' calculation.

**Note:** Statistics reported for cocoa farmers come from farmer surveys conducted among 4,458 cocoa farmers in 2019. Statistics reported for cooperatives and sections come from data collected in the 2022-2023 cocoa season, on a sample of 22 cooperatives and 208 sections.

Cocoa value chains in Côte d'Ivoire are long, complex, and strictly regulated. The vast majority of cocoa producers operate on a small scale. Small farmers, particularly those in remote areas, may face constraints in accessing inputs and sales markets. Cooperatives can play a significant role by providing farmers with a shared space for capacity development, joint acquisition of inputs, and negotiation power for better market conditions and prices (Bymolt et al., 2018). Cooperatives typically aggregate beans from hundreds of farmers, spread across various geographical areas called sections. In each section, middlemen ("*délégués de section*") working on behalf of the cooperative are responsible for purchasing cocoa beans from about 150 farmers, for which they usually receive an advance from the cooperative, and transport the beans to the cooperative's warehouse, located up to 30 kilometers away (Table 1). Ivorian cocoa cooperatives are integrated into an export/certification system, which creates some opacity in the cooperatives' relationships with their members and operational shortcomings (Renier et al., 2023).

Studies highlight issues such as poor management, lack of financial resources, and the inability to effectively support farmers (Bymolt et al., 2018).

Although exporters and local processors can provide training, inputs, and other services directly to farmers, they cannot purchase beans at the farm gate. Buyers can only procure cocoa from a dense network of some 3,000 cooperatives downstream in charge of collecting the requested volume on their behalf (International Cocoa Organization, 2024).<sup>2</sup> Sales are also temporally regulated, with purchases officially authorized from October 1st onward. Due to the multiplicity of actors, what the sector refers to as "direct sourcing" is traceable up to the first buyer, typically cooperatives, but not necessarily to the middlemen or the farmers.

## 2.2 Conventional vs ESS-certified cocoa

In Côte d'Ivoire, the cocoa sector encompasses both conventional and ESS (Environmentally and Socially Sustainable) certified cocoa production. Conventional cocoa is produced with standard farming practices, typically involving the use of chemical fertilizers and pesticides, and the labor practices may not always adhere to fair labor standards. In Côte d'Ivoire, cocoa farming is estimated to account for 37% of forest loss in protected areas (Kalischek et al., 2023) and to involve 790,000 child laborers in 2019 (Sadhu et al., 2020). In contrast, ESS-certified cocoa is produced under stringent guidelines that ensure both environmental sustainability and social responsibility. Certification programs such as Fair Trade, Rainforest Alliance, and UTZ are examples that set criteria for ESS certification. These programs emphasize the reduction of chemical usage, conservation of biodiversity, and the promotion of sustainable farming techniques. Additionally, they ensure that farmers receive fair wages, called certification premiums, and work under safe conditions. With these certification schemes, exporting companies seek to respond to the growing demand of high-income countries' consumers regarding the ethical standards of their purchases. Sixty percent of Ivorian cocoa farmers are "certified" producers (World Bank, 2019).

However, whether cocoa is produced according to ESS or conventional practices is indistinguishable from the beans themselves. As certified and conventional farmers co-exist in all cocoa-producing areas of the country, the broad geographic origin of the cocoa is also insufficient to ensure that it satisfies ESS

---

<sup>2</sup>See Ordinance n°2011-481 of December 28th, 2011, and Decree n°2012-1008 of October 17th, 2012.

production practices. Thus, the only way to ensure that cocoa beans were indeed produced under ESS production practices is to trace their origin all the way to an ESS-certified farmer.

### 2.3 Current traceability system

Cooperatives who purchase cocoa from certified farmers in response to their clients' orders are required by certifiers and clients to keep a traceability register, which records all transactions made with certified cocoa farmers.<sup>3</sup> Information reported in the traceability register is collected by middlemen. For each transaction, middlemen fill in a receipt that includes characteristics of the transaction, such as the farmer's unique identification code, the volume (in kg) purchased, the date of purchase, the price paid to the farmer, and the type of cocoa (certified or conventional). Farmers keep a copy of this receipt, which provides evidence of selling certified cocoa and allows them to claim their ESS-related premium at the end of the cocoa season. Cooperatives enter all receipt information in their electronic traceability registers, a copy of which is shared with their client and may further be requested by auditors from the national certification bodies.<sup>4</sup>

This traceability system is considered insufficient by many observers to guarantee that all certified cocoa does originate from ESS-certified farmers, and that conventional cocoa does not enter the value chain as certified one. In particular, concerns arise regarding actors at all levels of the value chains are truly incentivized to comply with the requested collection of information and/or to verify that no conventional cocoa beans are aggregated and sold as certified ones:

- Certified farmers may procure beans from non-certified neighbors and claim they come from their own farm;
- Middlemen in charge of sourcing ESS cocoa beans may aggregate beans from both certified and conventional farmers, and pretend they all come from certified ones;
- Cooperatives may register transactions with conventional farmers as originating from certified ones in their traceability registers;

---

<sup>3</sup>This has recently become an obligation requested by the Conseil Café Cacao (CCC), the cocoa regulatory organization in Côte d'Ivoire, as part of the ARS-1000 regional standard on sustainable cocoa, and in alignment with the requirements of the EU Deforestation Regulation (EUDR)

<sup>4</sup>Certification bodies are also requested to conduct audits at the farm level to assess whether certified farmers comply with the ban on child labor. However, not all farmers can be monitored due to the high costs and inefficiency of individual verification (audits are planned in advance and involve visiting only a few farms per cooperative), raising questions about the validity of the information they collect.

- Exporters may turn a blind eye to the possibility of such practices, especially when global demand for ESS-certified cocoa exceeds the production capacity of ESS-certified farmers.

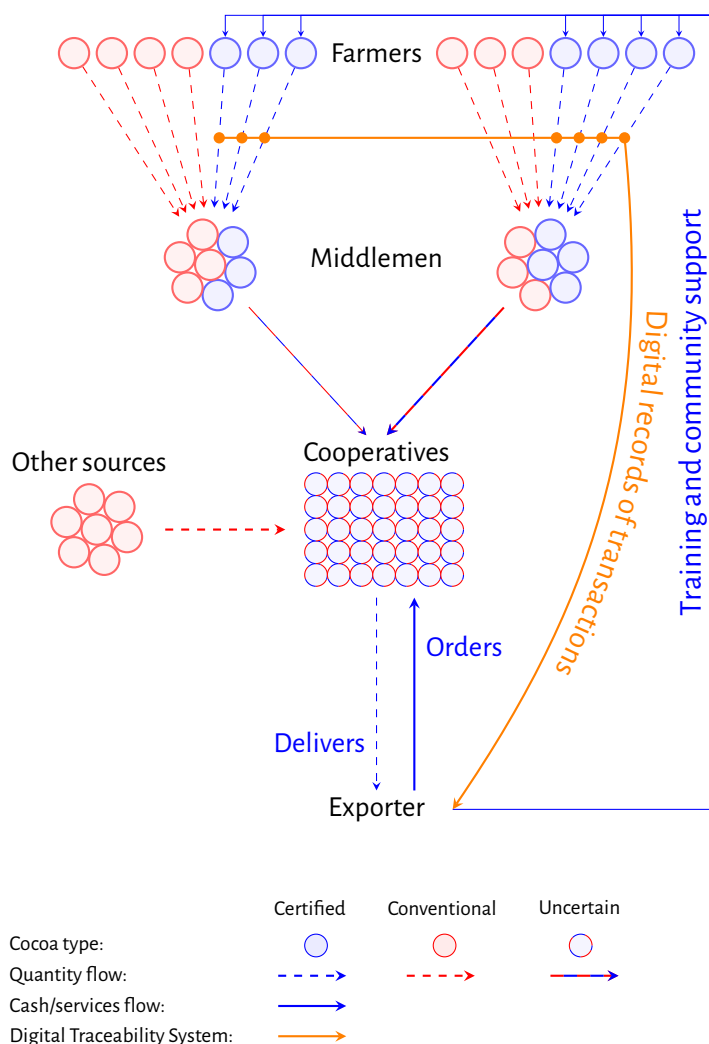
Overall, the system is such that all actors downstream in the value chain indirectly benefit from side-sourcing activities of their upstream suppliers as the procurement of larger volumes of ESS-certified beans increases their expected revenues. Thus, cocoa value chain actors face (very) limited incentives to monitor the effective origin of the beans. This is further compounded by the fact that thorough verification of transactions in cooperatives' registers is beyond the capacity of existing certification bodies, and the probability of sanctions is correspondingly limited.

Thus, the current organization of the cocoa value chain makes it possible for beans from certified cocoa farmers to be mixed with those produced by conventional cocoa farmers at several levels of the value chain. However, there is limited reliable information on the overall magnitude of side-sourcing practices and their distribution across the actors of the cocoa value chain. The current organization is depicted in Figure 1, showing how conventional beans (red in the figure) can be mixed with conventional (blue) ones, at various levels of the value chain.

## 2.4 A new digital traceability application

Recently, a global top 5 cocoa trading company operating in Côte d'Ivoire acquired a mobile application designed to increase the traceability of cocoa transactions across upstream actors in the cocoa supply chain (farmers, middlemen, and cooperatives). This application aims to directly collect information that cocoa middlemen have traditionally filled on paper when purchasing cocoa from farmers. In this Digital Traceability System (DTS), each farmer profile is linked to a unique identification (ID) card with a QR code, and kept by the corresponding farmer. When purchasing cocoa, middlemen must scan the QR code on farmers' ID cards. This action automatically displays the farmer's profile, where they can record the information on the transaction that they used to collect on paper receipts (date of purchase, type of cocoa -certified or conventional- quantity, and price). By collecting this information directly in the mobile application, the exporter expects to reduce farmers', middlemen', and cooperatives' capacities to supply conventional cocoa beans as if they were produced by ESS-certified farmers. This system is described as the orange information flow in Figure 1.

**Figure 1:** Products, services, and information flows in cocoa value chains



Source: Authors.

### 3. Data

Although widespread, side-sourcing is known by all actors in the cocoa value chain to be forbidden and subject (albeit with limited probability) to sanctions. Direct measures of side-sourcing through interviews of actors are prone to social desirability bias. While they could provide lower-bound estimates of the importance of this issue, the (potentially large) magnitude of the downward bias would remain unknown. An alternative approach is to reveal side-sourcing through comparisons of information on a given transaction as reported by the different actors of the value chain. Accordingly, side-sourcing would be revealed by the existence of meaningful and systematic differences across these reports. To this end, we gathered a uniquely rich set of data covering 22 out of the 29 cooperatives from which our partner cocoa trading company sources its ESS-certified cocoa beans. For the 2022/2023 season

in particular, we are able to compare information on a single transaction from four different sources: (1) traceability registers provided by cooperatives, (2) a farmer survey, (3) a trader survey, and (4) data reported by middlemen in the digital traceability application described in Section 2. We also accessed the cooperatives' traceability registers for the previous two cocoa seasons. These data sources and their time coverage are summarized in Table 2 and further described below.

**Table 2:** Cocoa transactions: sources of information

	Month of transaction					
	October	November	December	January	February	March
<b>2022/2023</b>						
Initial cooperative register	✓	✓	✓	✓		
Final cooperative register	✓	✓	✓	✓	✓	✓
Farmer survey	✓	✓	✓			
Middleman survey	✓	✓	✓			
Digital traceability application	✓	✓	✓	✓	✓	✓
<b>2021/2022</b>						
Final cooperative register	✓	✓	✓	✓	✓	✓
<b>2020/2021</b>						
Final cooperative register	✓	✓	✓	✓	✓	✓

*Source:* Authors.

**Note:** The table shows the various sources of information available on transactions conducted during the main cocoa season, from October to March.

**Cooperatives' traceability registers.** Each year, cooperatives are required to share their traceability registers with their clients as supportive evidence that the cocoa beans originate from farmers trained in adopting ESS practices. For each cooperative in our sample, we accessed the traceability registers for the 2020-2021, 2021-2022, and 2022-2023 cocoa seasons. The registers provide information on each transaction made during the season, specifying the volume of cocoa purchased, the date, the section where the transaction took place, and the farmer's name. Together, the end-of-season registers for these three seasons account for over 620,000 individual transactions.

For the 2022-2023 season, we further accessed two distinct versions of the traceability register. The first versions were shared with the exporting company at mid-season, and contained information on 116,765 transactions occurring between October and December. The second version was shared at the end of the season and included information on all transactions made during the season – those conducted in the months covered by the mid-season register and those conducted in the subsequent months – for a total of 278,793 transactions. In these final registers, there are 128,638 transactions corresponding to the



period covered by the first traceability register. These transactions constitute our main sample for impact analyses.

**Farmers and middlemen surveys.** In the 2022-2023 season, we designed and implemented two phone surveys between October 24 and December 11, capturing the peak of the main cocoa season. Within each cooperative, we randomly selected 8 farmers per section from the list of ESS-certified farmers kept by each cooperative, for a total of 1,652 farmers. At the same time, we also surveyed the 307 middlemen working for our sampled cooperatives.

The farmer survey focused on cocoa sales made by farmers in the three weeks preceding the phone interview. For farmers reporting at least one cocoa sale during this recall period, we collected information on the transaction date, the corresponding volume of cocoa beans (in kilograms and bags), the transaction value, the middleman's identity, and whether the middleman was using the digital traceability application during the transaction. We also collected socio-demographic information on the farmers and their relationship with middlemen.

The middlemen survey was structured into two main sections. First, middlemen were asked about their cocoa purchases with the surveyed farmers in their section over the past three weeks. To prevent any exchange of information between farmers and middlemen, both farmer and middlemen surveys were conducted simultaneously, typically on the same day. Second, middlemen were asked about their familiarity with the digital traceability application, whether they had received the associated tablet and training for its utilization, as well as their overall proficiency and usage patterns (such as frequency of use, reasons for non-use, perceptions of benefits and drawbacks, etc.).

The two surveys resulted in a total of 545 certified cocoa sales reported by farmers and 481 certified cocoa purchases reported by middlemen over the three weeks preceding the survey calls. The surveys were conducted in two phases: (1) from October 24 to November 13, covering all middlemen and half of the selected farmers across the 22 cooperatives; (2) from November 21 to December 11, covering all middlemen again, and the remaining half of the sampled farmers. This two-stage approach was adopted to ensure that farmers and middlemen from different cooperatives were surveyed within a similar time frame during the season. Starting in the last week of October, the survey covered transactions from early October, at the beginning of the main cocoa season, until mid-December.

**Traceability application data.** Only a subset of middlemen were targeted to use the traceability application during the main cocoa season (see Section 5). Overall, 9,634 transactions were reported on the mobile traceability application from October 1st to March 31st, 2023, representing a total of 2,827 tons of cocoa.

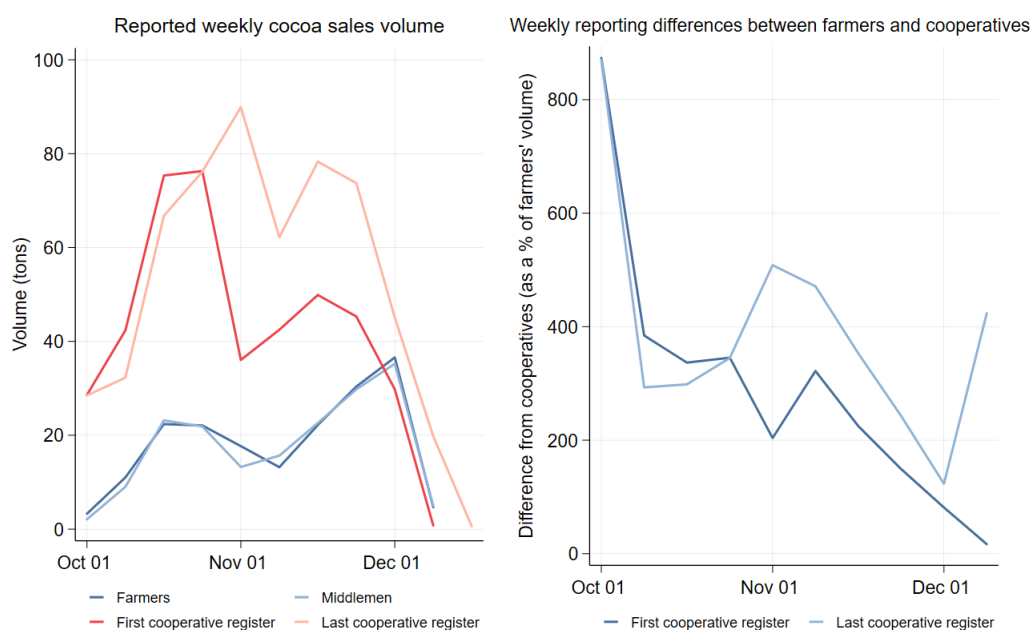
## 4. Side-sourcing in ESS-certified cocoa value chains

In this section, we use data from the multiple sources of transaction-level information described in Section 3, to estimate the extent of side-sourcing in ESS-certified cocoa value chains.

### 4.1 Reporting discrepancies across farmers, middlemen, and cooperatives

In the absence of side-sourcing, one should not observe meaningful differences between the volumes of cocoa sold as reported by farmers during their recall period, and those reported by middlemen and cooperatives for the same farmers and period. Overall, volumes reported by middlemen (178 tonnes) are fairly similar to those reported by farmers (183.5 tonnes) (Figure 2). In contrast, 573 tonnes were recorded in cooperatives' registers for the same period and these same farmers, a more than three-fold increase. While some discrepancies can be explained by typographical or measurement errors (such as variations in cocoa weight upon reaching the cooperative), these are insufficient to account for the considerable differences we observe. Recall biases by farmers are also unlikely given the short recall period of three weeks, during which farmers usually made only one to two transactions. These biases are also incompatible with the limited differences found between farmers' and middlemen's reports.

The indication of side-sourcing is further supported by the inconsistency of over-reporting. In Figure 2, we explore the time dynamic of reporting discrepancies by plotting weekly volumes of cocoa sold according to each source (left) and as a percentage of farmers' reported volume (right). Accordingly, cooperatives' registers show weekly volumes exceeding farmers' reports by some 30 to 90 tones per week, representing 123% to 870% of farmers' reports (Figure 2). Interestingly, while volume discrepancies are highest in November, their size relative to farmers' reports is highest in the early days of the season. This aligns with informal interviews with value chain actors, suggesting that cooperatives are often able to procure cocoa at low prices ahead of the season from farmers facing temporary liquidity constraints, and later record these transactions as occurring at the onset of the season. Overall, the

**Figure 2:** Weekly reporting differences between farmers, middlemen, and cooperatives

Source: Authors' calculation.

**Note:** *Left panel:* aggregate weekly volume of cocoa sold by farmers in the three weeks preceding the survey, as declared by farmers, middlemen, and cooperatives (in both the initial and final traceability registers), respectively. *Right panel:* relative difference between farmers' and cooperatives' reported volumes in the three weeks preceding the survey. The cooperative reported volumes are expressed as a percentage of the volume reported by farmers.

magnitude of reporting discrepancies between value chain actors, and in particular between farmers and middlemen on one side, and cooperatives on the other, suggests a large amount of conventional cocoa entering the value chain as if it were certified. Figure 2 further suggests important differences in volumes reported in cooperatives' mid-season and final registers, for the same farmers and the same period, suggesting significant editing of the registers' data which we further discuss below.

## 4.2 Who engages in side-sourcing?

If reporting discrepancies suggest the existence of side-sourcing, it remains unclear whether all actors or only some within the value chain engage in this practice. As discussed in Section 2.3, side-sourcing can indeed occur at any stage in the cocoa value chain, whether at the farmer, middleman, or cooperative level.

**Side-sourcing by farmers.** The fact that cooperatives report higher volumes sold by farmers than what farmers themselves indicate does not preclude the latter from engaging in side-sourcing. ESS-certified farmers can procure conventional cocoa from fellow farmers and sell it to middlemen as if it came from their own (certified) farm. We assess the likelihood of this behavior by comparing the volume that farmers reported selling during the survey's recall period (covering three weeks between October and December), to their predicted production. The latter is estimated based on their cocoa farm size and the average cocoa yields over the October-December period (corresponding to the first part of the main cocoa season). We use a yield of 0.5 tonnes per hectare as the lower bound and 0.6 tonnes per hectare as the upper bound (Kozicka et al., 2018). Cocoa farm size is calculated using both geospatial information (available for 1,308 farmers out of the 1,652 surveyed) and declared farm size available for all 1,652 surveyed farmers. In Table 3, we compare these estimates to farmers' reports and assess whether they systematically report higher sales than the theoretically feasible production of their farm. Our results indicate that only between 5.88% and 4.97% of the farmers have reported transaction sizes that exceed the expected theoretical volume given their farm size.

**Table 3:** Farmers' average theoretical and observed yields during the first part of the main cocoa season (October-December)

	Mapped farm size	Declared farm size
<b>Theoretical yields per farmer</b>		
Using the upper bound of yields (0.6/ha)	685.5	923.5
Using the lower bound of yields (0.5/ha)	570.8	768.7
<b>Observed yields per farmer</b>	225.5	134
<b>% of farmers having an observed yield higher of theoretical one</b>		
Using the upper bound of yields	5.88%	0%
Using the lower bound of yields	4.97%	0%
N	1,308	1,652

*Source:* Authors' calculation.

**Note:** Yields are calculated for the first part of the main cocoa season (October to December), based on farmers' sales reports in the farmer survey. We use a yield of 0.5 tonnes per hectare as the lower bound and 0.6 tonnes per hectare as the upper bound (Kozicka et al., 2018). Main crop production is considered to represent around two-thirds of the year's cocoa production, with October to December accounting for 69% of the main season (according to 2020-2021 cooperative traceability registers).

These results align with our field observations indicating that side-sourcing by farmers is unlikely. ESS certification also leads to exporters' support of community development projects, so it is usually the case that neighbors of ESS-certified farmers are themselves certified. Given that cocoa-producing communities are generally remote and far from the main roads, it is also logistically difficult and costly for individual farmers to collect conventional cocoa from distant farmers. Overall, while indirect, the

evidence at hand suggests that, if it exists, side-sourcing by farmers is limited in magnitude. In what follows, we consider farmers' own reports on the cocoa they sold to only correspond to their own ESS-certified cocoa production, and thus as a credible benchmark for comparison with middlemen and cooperatives data.<sup>5</sup>

**Side-sourcing by middlemen and/or cooperatives.** Figure 2 shows large differences between farmers' reported transactions and the corresponding volumes recorded in cooperatives' traceability registers, indicating significant issues of side-sourcing. As observed, there are only limited differences between transaction volumes declared by farmers and those declared by middlemen. However, similar reports between farmers and middlemen do not necessarily exclude middlemen from engaging in side-sourcing. Indeed, we only observed transactions from our survey, which were self-declared by the middlemen over the last three weeks and for specific farmers. Therefore, we are unable to observe transactions that middlemen actually report to cooperatives, including those they may have conducted with non-certified farmers. Consequently, it is not possible to determine how much of the over-declaration of volumes found in cooperatives' traceability registers (Figure 2) is due to cooperatives' over-reporting on transactions declared by middlemen, and/or how much is due to middlemen themselves forging receipts of their transactions with ESS-certified farmers.

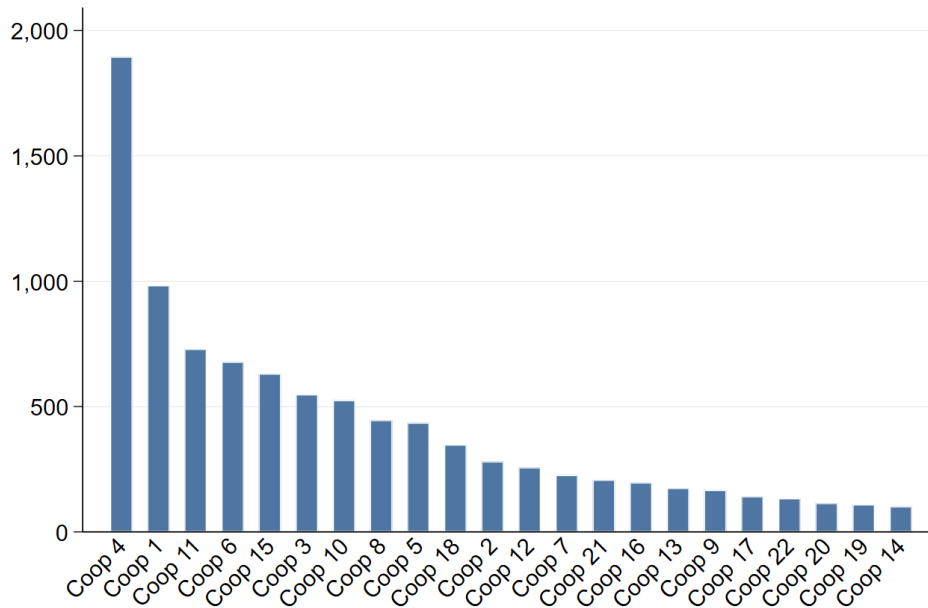
We next investigate whether over-reporting of volume is limited to some cooperatives or some sections, or if it is distributed across all of them. In Figure 3, we present the level of discrepancy between farmers' and cooperatives' reports in each cooperative. Accordingly, our data points to a systemic issue, with the smallest differences reaching 101% in one of the cooperatives, effectively doubling the volume of cocoa reported by farmers during the same period. The magnitude of these disparities however varies significantly from one cooperative to another. Delving one step further at the section level, we find excess volumes reported for 85% of them (see first bar column in Figure 5 below).

We investigate several possibilities for such discrepancy across cooperatives and sections, as compared to farmers' reports. One potential explanation may relate to the geographical location of cooperatives. In some regions, weather events may cause cooperatives temporary difficulties in sourcing ESS-certified cocoa to comply with the volume ordered by their clients, which is usually determined based on previous years' yields. Alternatively, in more remote areas where the official cocoa price is difficult to enforce,

---

<sup>5</sup>If anything, under-estimating farmers' side-sourcing would mean that we under-evaluate the extent of side-sourcing at the value chain level.

**Figure 3:** Volume reported by cooperatives for surveyed farmers during the recall period (as a % of volume reported by farmers)



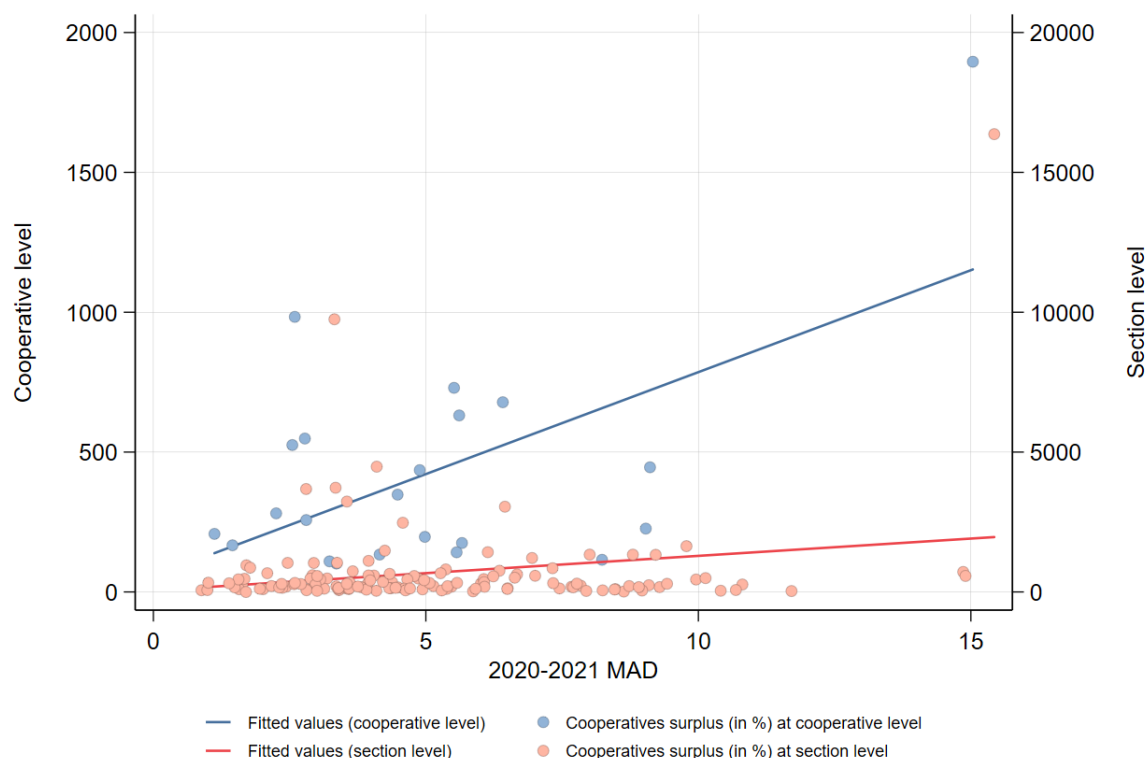
Source: Authors.

**Note:** The figure shows the reporting differences between cooperatives and farmers in terms of volume sold by surveyed farmers during their recall period. Thus, cooperative number 20 reported a volume of cocoa purchased from surveyed farmers during their recall period 101% higher than that declared by the farmers themselves.

cooperatives may be further encouraged to source their cocoa from areas that do not cover their certified farmers. However, we find no evidence that the geographical location of cooperatives helps predict the differences in reporting between farmers and cooperatives, as illustrated in Figure A1 in the Appendix.

Another explanation may relate to heterogeneous opportunistic behaviors across cooperative management teams and/or middlemen. With direct measurements of these behaviors unavailable, we apply a Benford Law Conformity (BLC) test to the cooperatives' registers from the 2020-2021 cocoa season and correlate the results with reporting discrepancies in the 2022-2023 season. The BLC test states that the distribution of first digits in standard numerical datasets follows a specific distribution, and a statistical series' Mean Absolute Deviation (MAD) from this distribution provides a reliable indication of the extent of data manipulation (Benford, 1938). The MAD from Benford's Law is increasingly used to detect potential anomalies in data and is often employed in fraud detection (Sambridge and Jackson, 2020; Kaiser, 2019; Demir and Javorcik, 2020). We compute the MAD based on transaction volumes reported by cooperatives in their traceability registers for the 2020-2021 cocoa season. We find high MADs for most cooperatives and for most sections, with an average of 5, while a MAD close to 0 suggests no data fraud, indicative of significant data manipulation for most cooperatives.

**Figure 4:** Correlation between differences in volume declared by cooperatives compared with farmers and 2020-2021 MAD



Source: Authors' calculation.

**Note:** The graph shows the correlation between differences in volume reported (in tonnes) by surveyed farmers and cooperatives, and the Mean Absolute Deviation from Benford's Law (MAD), both at section-level (orange line, right axis) and cooperative-level (blue line, left axis).

In Figure 4, we correlate cooperative-level and section-level MAD with the level of reporting differences between cooperative registers and farmers' own reports in our survey. We find a significant and positive correlation between the two series, indicating that the largest reporting differences in 2022-2023 are found in cooperatives whose 2020-2021 traceability registers were also the most suspicious according to the BLC test.<sup>6</sup>

### 4.3 Strategic misreporting

**Strategic misreporting by cooperatives and/or middlemen.** As mentioned in Section 2.3, inspecting cooperatives' traceability registers is the main auditing approach to assess that cocoa comes

<sup>6</sup>As can be seen on the figure, the slope of the correlation is in part driven by exceptionally high MADs and over-reporting in at least one cooperative and one section (top-right observations on the graph).

from certified farmers. Key audit variables are volume per farmer and volume per transaction, enabling the detection of abnormal figures. As audits have become more common in recent years, we find suggestive evidence that cooperatives and/or middlemen have also adapted their reporting strategy.

First, using cooperatives' registers over the past three seasons, we present in Table 4 the averages of the main key audit variables at the section level and their evolution over time. In the first two rows, we find that the average number of farmers per section and the average volume collected per farmer remained stable over the three years. The last row of Table 4 also suggests that the transaction volumes recorded in the cooperatives' registers evolve over the years, to align more closely with Benford's law distribution. However, this does not indicate a reduction in fraudulent practices but rather that data manipulation has become less detectable over time. Indeed, only transactions with an atypically high volume can be easily detected by audits. Instead, we find that the number of transactions per farmer has abnormally increased, almost doubling in three years, resulting in a significantly lower volume of cocoa per transaction (less than half) between the 2020-2021 and the 2022-2023 cocoa seasons. Some cooperatives may have made logistical changes in the way they purchase and collect cocoa (leading to smaller-sized transactions), but we have not found evidence of such significant changes during our field interviews. This significant increase in the number of transactions per farmer, which reduces the associated volume per transaction but not the average volume per farmer, suggests a shift in fraudulent practices towards an increase in the number of (possibly fictitious) transactions per farmer, compensated by a reduction in the volume per transaction.

Second, section-level differences between farmers' survey data and cooperatives' traceability registers also point to (mis)reporting strategies of cooperatives and/or middlemen to integrate additional volumes of cocoa in their reports without raising auditors' suspicions. As summarized in Figure 5, we find transactions attributed to farmers who did not report selling cocoa during the recall period when surveyed (but are indeed on the list of certified cooperative members) in 85% of the sections (second bar column). Out of the 1,652 surveyed farmers, 26.7% (27% in the middlemen survey) reported at least one cocoa transaction in the previous three weeks. For the same farmers and the same period, the final cooperatives' registers indicate that 78% of these farmers had sold cocoa at least once. This first reporting strategy accounts for 73% of the discrepancy in the number of transactions at the section level. The remaining 27% variation is explained by cooperatives adding transactions not declared by farmers who



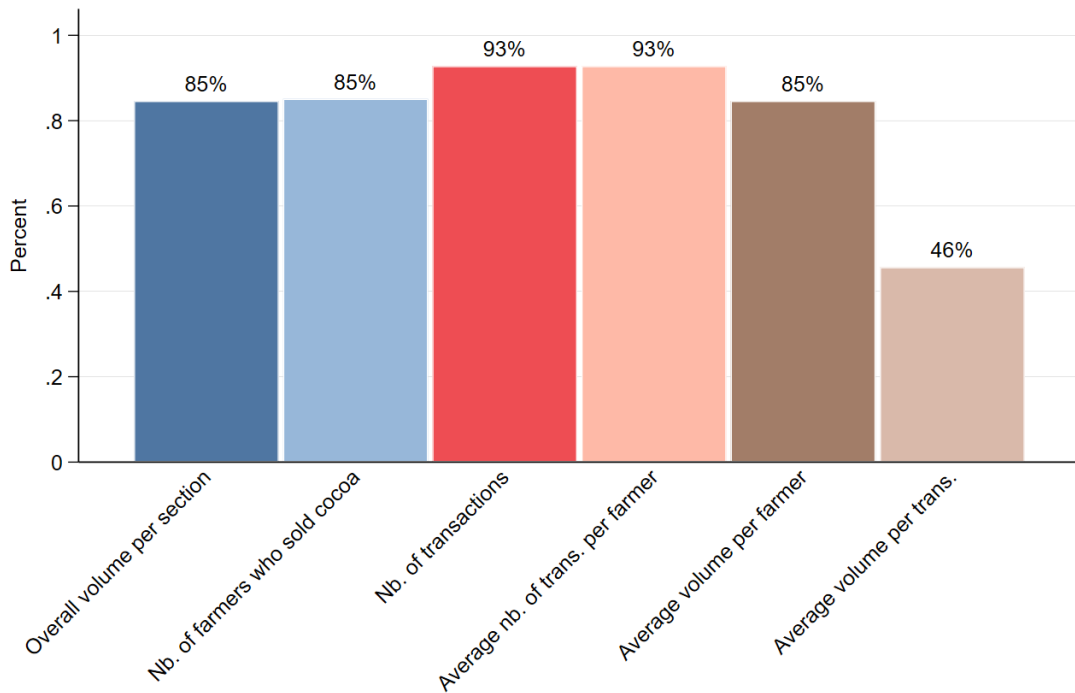
**Table 4:** Descriptive characteristics of sections across years, for the main cocoa season

	(1) 2020-2021		(2) 2021-2022		(3) 2022-2023		T-test P-value		
	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)	(2)-(3)
Average number of farmers per section	184	157.000 (12.074)	197	152.579 (12.693)	208	150.822 (14.092)	0.499	0.461	0.688
Average volume of cocoa bought per farmer	184	1864.092 (141.415)	197	1773.984 (143.729)	208	1743.979 (176.479)	0.290	0.312	0.732
Average volume of cocoa per transaction	184	375.737 (40.826)	197	300.375 (23.915)	208	204.483 (22.481)	0.035**	0.000***	0.001***
Average number of trans. per farmer	184	5.835 (0.598)	197	6.458 (0.653)	208	9.352 (0.757)	0.377	0.000***	0.001***
MAD	184	5.293 (0.570)	197	4.304 (0.360)	208	2.392 (0.298)	0.065*	0.000***	0.000***

Source: Authors' calculation.

Note: The values displayed for t-tests are p-values. Standard errors are clustered at the cooperative level. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent critical level. MAD stands for Mean Absolute Deviation from Benford's Law.

**Figure 5:** Proportion of sections for which a key indicator is higher in cooperatives' traceability registers than in farmers' reporting



Source: Authors' calculation.

**Note:** The figures reported are calculated from farmer survey data compared with section-level data obtained from the cooperatives' final 2022-2023 traceability registers for the same farmers and the same period.

reported having sold cocoa at least once during the survey.<sup>7</sup> The increase in the number of transactions in cooperative registers mechanically leads to an increase in the average number of transactions per farmer in 93% of sections. This strategy leads to an increase in the average volume sold per farmer for 85% of sections, with an average surplus of 301 kilos per farmer compared to their own declarations.

**Data manipulation at cooperative-level.** Our results thus far are not fully conclusive regarding cooperatives and/or middlemen's involvement in the large-scale and deliberate misreporting observed in our data. As discussed in Section 2.3, all actors downstream in the value chain, whether middlemen, cooperatives, or exporters, can indirectly benefit from side-sourcing operations undertaken by their upstream suppliers. Under these conditions, and given the scale of side-sourcing identified, it is unlikely that cooperatives would be entirely unaware of the side-sourcing activities of their middlemen if and where they exist. The reverse is not true: cooperatives can more easily aggregate conventional cocoa beans without middlemen's knowledge, and benefit from it.

<sup>7</sup>In the end, it appears that cooperatives' registers reported one additional transaction of each farmer over the considered 3 weeks period, irrespective of whether the farmer had effectively sold any cocoa during that period.

To assess whether cooperatives are directly engaged in side-sourcing activities, we compare the mid-season version of their traceability registers (sent around November/December) and the end-season one (sent in May/June). We report in Table 5 summary statistics of key differences between the two versions. For the same period covered by the initial register, there should not be a valid reason (other than a data entry error subsequently corrected) for the volume of cocoa, the date, or even the very existence of a transaction to be different between the mid-season and the final versions of the registers. Our results suggest otherwise, with significant increases in the number of supplying farmers and the number of corresponding transactions, albeit with a more stable (and slightly lower) overall volume. Taken together, these changes imply that the average volume per transaction is reduced by close to 13.5% in the final traceability register, making it less likely to be rejected by eventual auditors. Thus, while one cannot entirely rule out that middlemen do engage in side-sourcing of cocoa and related misreporting of their transactions, our results are more conclusive regarding cooperatives, suggesting that they do take an active part in side-sourcing activities and/or related misreporting.

**Table 5:** Differences in key indicators between first and final traceability registers of cooperatives during the 2022-2023 cocoa season

	Initial traceability register	Final traceability register	Variation
Overall volume (tons)	25,480	24,777	-2.76%
Number of farmers	28,127	29,236	+3.94%
Number of transactions	114,596	128,630	+12.24%
Average number of trans. per farmer	8.41	8.16	-2.97%
Average volume per trans.	222.35	192.63	-13.37%
Average volume per farmer	905.90	847.50	-6.45%
% of receipt numbers that are duplicates	9.41%	6.48%	-31.14%

*Source:* Authors' calculation.

**Note:** This table presents key characteristics of first and final traceability registers provided by cooperatives for the 2022-2023 cocoa season, as well as the variation in each indicator between the two registers.

## 5. Assessing the impact of a new digital traceability system: a randomized controlled trial

The new Digital Traceability System (DTS) described in Section 2.4 was initially set to be introduced in the 2021-2022 season in a subset of cooperatives but was postponed to the following year for technical reasons. In partnership with the cocoa exporter, we leverage the partial introduction planned for the

2022-2023 season to assess its effectiveness through a randomized allocation of the device to some cooperatives and sections.

## 5.1 Design

We rely on a two-level randomization design intended to (partially) identify the full effect of the DTS on side-sourcing and related misreporting, through comparisons of volumes reported in three distinct groups (where the superscripts  $c$  and  $s$  indicate cooperative-level and section-level treatment assignment, respectively):

$C^c$  – Control group: the cooperative was not to use the DTS in the 2022-2023 cocoa season. Its middlemen were neither trained nor equipped with DTS tablets. We label as  $C_0^s$  the sections that belong to these cooperatives, with the subscript 0 indicating their "pure control" treatment status.

$T^c$  – Treatment group: the cooperative was to use the DTS in the 2022-2023 cocoa season. Only a subset of sections from each of these cooperatives had their middlemen trained and equipped with DTS tablets, which we label  $T_\theta^s$ , while the remaining sections within treated cooperatives are control sections labeled as  $C_\theta^s$ . Here,  $\theta$  further indicates treatment intensity at cooperative-level, which we describe below.

This design intends to identify the full effect of the DTS on side-sourcing and related misreporting. Direct comparisons of treated sections ( $T_\theta^s$ ) with pure control ones ( $C_0^s$ ) indicate changes in misreporting by middlemen and/or cooperatives, but not necessarily changes in side-sourcing activities. As discussed in Section 4, cooperatives can alter their traceability registers, reporting transactions across all sections. Accordingly, cooperatives engaging in side-sourcing activities can choose to limit fraudulent reporting of the corresponding volumes on their treated sections ( $T_\theta^s$ ) and instead allocate these extra volumes to control ones ( $C_\theta^s$ ). Reduced side-sourcing in response to DTS simultaneously implies evidence of lower reported volumes in treatment sections ( $T_\theta^s$ ) compared to pure control ( $C_0^s$ ) ones, and equivalent (or lower in case of a cooperative-level change in side-sourcing) volumes in control sections of treated cooperatives ( $C_\theta^s$ ) compared to pure control ( $C_0^s$ ) ones.

Let  $y$  indicate the volume of certified beans reported in the cooperative register for a given section,  $y^*$  the volume effectively sourced from certified farmers, and  $\epsilon$  the volume of conventional beans. In sections of cooperatives where the DTS was not promoted in the 2022-2023 season, the expected

volume reported at section-level is  $E(y|C_0^s) = E(y^*) + E(\epsilon)$ . In comparison, in sections where the DTS is effectively used and enables full traceability of the cocoa beans, the reported volumes become  $E(y|T_\theta^s) = E(y^*) \leq E(y|C_0^s)$ , with  $\theta \in [0, 1]$  capturing the share of treated sections within the treated cooperatives. The expected volumes reported from control sections within these cooperatives is then a positive function of  $\theta$ , bounded by:

Lower bound:  $E(y|C_\theta^s) = E(y^*) + E(\epsilon) = E(y|C_0^s)$  if none of the conventional cocoa sourced in treated section is now reported in control ones, indicative that side-sourcing is mostly done by middlemen – as cooperatives would be able to account for it in non-treated sections if needed.

Upper bound:  $E(y|C_\theta^s) = E(y^*) + (1 + \theta)E(\epsilon) \geq E(y|C_0^s)$  if conventional cocoa that was to be reported in treated sections is now entirely distributed across control ones, indicative of side-sourcing being done (mostly) by cooperatives.

DTS may, however, only partially cover the transactions reported from treated sections. For instance, technical difficulties may imply that middlemen could only use the DTS some time after the beginning of the season. Let  $\delta \in [0; 1]$  measure the share of the season where DTS could effectively be used. Assuming that farmers sales are uniformly distributed throughout the season, the expected volume reported in treated sections now becomes  $E(y|T_\theta^s) = E(y^*) + (1 - \delta)E(\epsilon) \leq E(y|C^c)$ . Volumes in control sections of treated cooperatives now depend positively on both  $\theta$  and  $\delta$  and are bounded by:

Lower bound:  $E(y|C_\theta^s) = E(y^*) + E(\epsilon) = E(y|C^c)$

Upper bound:  $E(y|C_\theta^s) = E(y^*) + \delta\theta E(\epsilon) \geq E(y|C^c)$

Of the 25 cooperatives from which the cocoa exporting company sourced ESS-certified beans, 22 are part of the present study and were randomly allocated to one of three groups.<sup>8</sup> Treated cooperatives were further randomly allocated to one of two groups, representing two different levels of  $\theta$ :

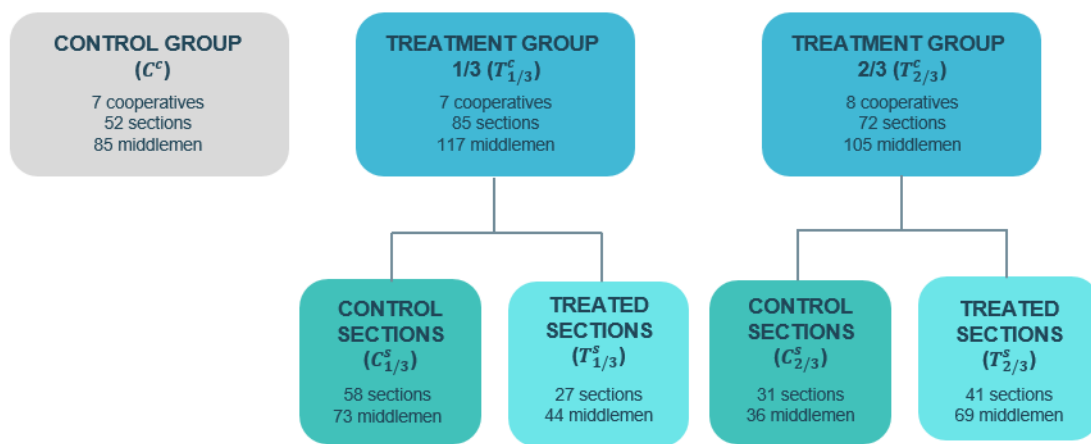
$T_{\frac{1}{3}}$  – Low-intensity treatment group: one-third of the cooperative's sections were to have their middlemen trained and equipped with DTS tablets. As the section level, let  $T_{\frac{1}{3}}^s$  and  $C_{\frac{1}{3}}^s$  be indicative of the treated and control sections in these cooperatives.

<sup>8</sup>The remaining 3 were not included for budget constraints.

$T_{\frac{2}{3}}^2$  – High-intensity treatment group: two-thirds of the cooperative’s sections were to have their middlemen trained and equipped with DTS tablets. As above, let  $T_{\frac{2}{3}}^s$  and  $C_{\frac{2}{3}}^s$  indicate the treatment status of the corresponding sections.

The resulting treatment allocation is summarized in Figure 6, where 52 sections are included in the control group ( $C^c$ ), 85 sections in the first treatment group ( $T_{\frac{1}{3}}^c$ ), and 72 sections in the second treatment group ( $T_{\frac{2}{3}}^c$ ). Overall, 68 sections were selected to use the traceability mobile application, covering 113 cocoa middlemen.<sup>9</sup>

**Figure 6:** Control and treatment groups at the cooperative and section level



Source: Authors.

**Note:** The figure shows the different treatment arms at cooperative- and section-level. Each cooperative and section was randomly allocated to the treatment or control group.

## 5.2 Experimental integrity

**Balance and attrition.** In Table 6, we report a series of simple difference-in-means to ensure that the random allocation of treatment across cooperatives and sections resulted in groups with comparable average characteristics. We rely on pre-intervention data from 2020-2021, obtained from the cooperatives’ registers. Panel A compares the characteristics of cooperatives allocated to either  $C^c$ ,  $T_{\frac{1}{3}}^c$ , or  $T_{\frac{2}{3}}^c$  groups. Although limited statistical power due to the small number of cooperatives, we do not find any significant differences across the cooperatives in each group. Panel B reports similar comparisons at the section level, comparing those allocated to either one of the  $C_{\theta}^s$ ,  $T_{\theta}^s$ , or  $C_{\theta}^s$  groups.

<sup>9</sup>We did not randomly vary the share of the season covered by the DTS, as we expected full-season coverage. As described below, we uncover limited and highly heterogeneous use of the traceability application across and within treated cooperatives, which is unlikely to be explained by exogenous (e.g. technical) reasons and is more likely related to strategic resistance to DTS (see Section 6).

Among the 134 sections tested for differences in key characteristics, we find no evidence of meaningful imbalances in their pre-intervention characteristics.

At farmer- and middleman-level, we further ensure that participants in the  $C_0^s$ ,  $T_\theta^s$ , and  $C_\theta^s$  groups are on average comparable, based on characteristics not related to their sales or purchases (as we draw our information from the farmer and middleman phone surveys which were implemented during the season when the DTS was promoted). Results are reported in Tables A1 and A2 and only point to limited differences across groups.<sup>10</sup> Last, the implementation of both farmer and middleman surveys did not yield any attrition with respect to our sampling plan, which we ensured through repeated phone calls every time the person was not reachable or available.

**Implementation.** The implementation of our design broadly followed what had been planned, further ensuring the internal validity of the estimations discussed below.

**Training and equipment provision for treated middlemen.** To ensure the successful implementation of the project, the cooperatives selected for treatment were informed about the new DTS and the associated randomized control trial several months before the start of the experiment. Within each treated cooperative, employees of the cocoa exporting company were responsible for explaining the functionality and utility of the mobile application, while our research team provided cooperatives with the list of treated sections (and thus middlemen). Cooperatives were then tasked with informing the selected middlemen about the mobile application implementation and to plan training sessions where middlemen were taught how to use the application and address any questions they might have. 94% of them attended at least one and possibly more of the training sessions planned between July and early September 2022 (Table 7).<sup>11</sup>

At the time of the first middleman survey in October 2022, 90% of the selected middlemen had functional tablets ready for use with the mobile traceability application, and 95.5% declared having

---

<sup>10</sup>More farmers in the spillover sections (i.e., control sections from treated cooperatives) (72%) than in the pure control (63%) and treated (66%) sections have an active mobile money account. Significantly more farmers in treated sections (75%) than in spillover sections (59%) reported during the survey that they often sell to pisteurs rather than to the middleman working for their cooperative. At the middleman level, the only difference is that treated middlemen are better educated than middlemen in control sections (79% vs 70%).

<sup>11</sup>Among those middlemen who were not able to attend any training session, 62.5% felt they had received sufficient information on how to use the application directly from the cooperative they work with, and only 3 middlemen reported receiving no information about the application when initially surveyed.

Table 6: Balancing tests

	(1) Control		(2) Treated 1/3		(3) Treated 2/3		(4) Total		T-test P-value		
	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(2)-(3)	
<b>Panel A: Cooperatives</b>											
Average number of farmers per cooperative	7	1093.857 (138.434)	7	1479.571 (179.659)	8	1359.375 (228.980)	22	1313.136 (110.531)	0.113	0.339	0.686
Average number of farmers per section	7	290.000 (84.062)	7	289.714 (41.500)	8	317.875 (25.359)	22	300.045 (18.672)	0.996	0.521	0.571
Average number of trans. per farmer	7	5.668 (1.208)	7	7.555 (1.527)	8	5.224 (0.595)	22	6.107 (0.660)	0.350	0.745	0.175
Average volume of cocoa bought per farmer	7	1694.064 (298.903)	7	2059.131 (265.868)	8	1811.526 (113.396)	22	1852.935 (131.008)	0.378	0.718	0.404
Average volume of cocoa per transaction	7	404.668 (63.116)	7	351.799 (57.306)	8	435.313 (59.154)	22	398.990 (33.792)	0.546	0.728	0.328
Mean Average Deviation from Benford's Law (MAD)	7	4.538 (0.823)	7	4.883 (0.902)	8	5.612 (1.604)	22	5.038 (0.676)	0.782	0.562	0.699
<b>Panel B: Sections</b>											
	(1) Control		(2) Spillover		(3) Treated		(4) Total		T-test P-value		
	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(2)-(3)	
Average number of farmers per section	46	166.457 (22.703)	73	154.781 (18.112)	65	152.800 (14.704)	184	157.000 (12.074)	0.682	0.606	0.905
Average volume of cocoa bought per farmer	46	1839.325 (307.652)	73	1876.987 (181.877)	65	1867.138 (156.201)	184	1864.092 (141.415)	0.914	0.934	0.918
Average volume of cocoa per transaction	46	400.881 (106.131)	73	361.997 (42.045)	65	373.374 (48.151)	184	375.737 (40.826)	0.726	0.808	0.691
Average number of trans. per farmer	46	5.836 (1.521)	73	5.930 (0.759)	65	5.727 (0.573)	184	5.835 (0.598)	0.954	0.945	0.648
Mean Average Deviation from Benford's Law (MAD)	46	5.262 (1.151)	73	5.077 (0.656)	65	5.559 (0.801)	184	5.293 (0.570)	0.886	0.828	0.386

Source: Authors' calculation.

Note: The values displayed for t-tests are p-values. Standard errors are clustered at the cooperative level. <sup>\*\*\*</sup>, <sup>\*\*</sup>, <sup>\*</sup> and <sup>\*</sup> indicate significance at the 1, 5, and 10 percent critical level. The figures reported come from the 2020-2021 cocoa season, prior to any treatment.



received their username and password to access the application.<sup>12</sup> By the time of the second survey in November 2022, all treated middlemen had a tablet and login credentials. However, nearly a quarter of the middlemen (24%) reported needing an assistant to use the application.<sup>13</sup>

**Table 7:** Implementation, equipment distribution, and compliance with the DTS use according to treatment assignment

	Assigned to	
	Control	Treatment
<b>Equipment and implementation</b>		
Middlemen with training	1.00%	93.75%
Middlemen with a functioning tablet		
<i>In October 2022</i>	1.00%	90.25%
<i>In November 2022</i>	1.00%	100%
Middlemen with login and password		
<i>In October 2022</i>	1.00%	95.60%
<i>In November 2022</i>	1.00%	100.00%
Farmers with personal ID cards	0.00%	95.79%
<b>Compliance</b>		
Cooperatives using the app	0.00%	100%
Sections using the app	1.40%	82.35%
Middlemen using the app	1.00%	72.6%

*Source:* Authors' calculation.

**Note:** 93.75% of middlemen assigned to the treatment group had been trained to use the traceability application at the start of the survey.

**Farmer-level information campaign and QR-coded ID cards.** Farmers in treated sections were provided with personal identification cards equipped with a QR code linked to their individual profiles. At the time of the farmer survey, 96% of farmers from treated sections did possess their ID cards, which are necessary for the middlemen to record their sales in the mobile traceability application.<sup>14</sup>

Farmers were informed about the DTS by local lead farmers, the middlemen themselves, and/or the exporting company's sustainable sourcing manager based in each cooperative. During the farmer survey, 97.5% of farmers in the treated sections stated that they had indeed received information and

<sup>12</sup>7% had non-functional tablets, and 3% reported not having received one. The reasons cited by middlemen who reported not having a tablet at the time were as follows: the middleman did not know how to use a tablet, no tablet was available at the cooperative, and the middleman was absent during the distribution of devices.

<sup>13</sup>63% reported having an assistant available to help them, but 15% indicated that their assistant was not always available. The remaining 22% reported not having an assistant to support them, either due to unavailability (66.7%) or because they could not afford to hire one (33.3%).

<sup>14</sup>Of the remaining 4%, half did receive their card but reported to have lost it.

explanations about the traceability application. Among them, 76% could explain that they needed to provide their ID card to their middlemen for each cocoa sale. However, less than half indicated that they were required to keep their ID card with them (47%) and to sign the middleman's tablet after each sale (48.5%).

Overall, the effective allocation of treatment largely followed the randomized design. Moreover, since treated farmers and middlemen were supposed to be equipped with the ID card and the application, respectively, there were no cases of spillovers between the control and treated groups that could bias the internal validity of our design.

### 5.3 Estimation framework

The effect of the DTS can be assessed from the transaction level information available from the farmer and middleman surveys, the mid-season and end-season versions of cooperatives' traceability registers, and the DTS data.<sup>15</sup> Accordingly, the DTS-induced changes in side-sourcing and related misreporting can be inferred from discrepancies across these data sources, in response to treatment status at the section level, with large discrepancies indicative of large  $\epsilon$ : more side-sourced conventional cocoa.

We detail our main estimation approach in the following set of equations, where the omitted category is the pure control group of sections in the control cooperatives:  $C_0^s$ . We rely on estimates at both transaction-level outcome variables  $y_t$  and at section-level ones  $y_s$ , alternatively measuring whether the transaction was digitized (hereafter "take-up" estimates in Section 6.2) and whether the reported information differed across reporting sources and treatment status (hereafter Intent to Treat (ITT) estimates in Section 7):

$$y_{st} = \alpha_c + \beta_1 T_\theta^s + \beta_2 C_\theta^s + \gamma X^t + \nu^{st} \quad (3.1)$$

where  $\alpha_c$  is a cooperative-specific constant term, and  $X^t$  is a vector of transaction-level characteristics (date and volume).  $\nu^{st}$  is a composite error term, with a section-level component that we account for by clustering our estimated standard errors at the section level. By design, we expect  $\hat{\beta}_1 > 0$  and  $\hat{\beta}_2 = 0$  for the "take-up" estimates. As for ITT estimates, if the DTS is effective in reducing discrepancies between our

<sup>15</sup>As described in Section 3, farmer-level and trader-level survey information are close to identical.

various data sources, one expects  $\hat{\beta}_1 < 0$ , while  $\hat{\beta}_2$  should be positive if the discrepancies are reported onto control sections in the treated cooperatives.

Next, we further leverage the variation in treatment intensities at the cooperative level ( $\theta$ ), to further assess the extent of reported side-sourcing onto control sections of treated cooperatives:

$$y_{st} = \alpha_c + \beta_1 T_{\frac{1}{3}}^s + \beta_2 C_{\frac{1}{3}}^s + \beta_3 T_{\frac{2}{3}}^s + \beta_4 C_{\frac{2}{3}}^s + \gamma X_t + \nu^{st} \quad (3.2)$$

Accordingly, evidence of cooperative-level strategic misreporting should imply that  $\hat{\beta}_2 \geq \hat{\beta}_4$ , while we do not expect *a priori* differences between  $\hat{\beta}_1$  and  $\hat{\beta}_3$ .

Last, we investigate in Equation 3.3 whether misreporting strategies vary with transaction-level characteristics through the added interaction terms and the corresponding estimated parameters  $\hat{\delta}_1 - \hat{\delta}_4$ :

$$y_{st} = \alpha_c + \beta_1 T_{\frac{1}{3}}^s + \beta_2 C_{\frac{1}{3}}^s + \beta_3 T_{\frac{2}{3}}^s + \beta_4 C_{\frac{2}{3}}^s + \gamma X_t \quad (3.3)$$

$$+ \delta_1 X_t' \times T_{\frac{1}{3}}^s + \delta_2 X_t' \times C_{\frac{1}{3}}^s + \delta_3 X_t' \times T_{\frac{2}{3}}^s + \delta_4 X_t' \times C_{\frac{2}{3}}^s + \nu^{st}$$

## 6. Support vs resistance to traceability

On the technical side, and apart from isolated cases, the DTS was fully functional by October 1st, the first day of the cocoa season, with the vast majority of middlemen being equipped and trained with the use of DTS. Yet, its use remained far from complete and heterogeneous across and within treated cooperatives. We describe these issues in the following paragraphs and further assess whether and how the promotion of DTS effectively led to the digitization of cocoa transactions in treated sections.

### 6.1 Initial resistance to DTS

Combining our own field observations during the study, with the various datasets described in Section 3, we first describe how the DTS was perceived and supported/opposed by the different actors of the value chain. At farmer-level, the DTS was broadly supported since the beginning of the season. Amongst the

farmers surveyed in October and November, 71.5% of them expected that the DTS would ensure that they receive their certification premium. In some cooperatives, some farmers refused to participate in the project for several weeks but not so much in relation to the DTS itself than as a means to protest against their cooperative, which had not paid them the traceability premium in recent years.

Middlemen's support for the DTS was comparatively lower. At survey time, middlemen reported two main concerns. First, the use of the DTS would generate additional work for them. Not only did they have to continue to manually track their cocoa purchases using paper receipts, but they now had to enter this same information in the mobile application, without any compensation for the extra workload. Second, 24% of the middlemen argued that they would need assistance to use the DTS mobile application, given their limited digital skills. The supply of these assistants and their remuneration sparked tensions with the cooperatives which were reluctant to bear the associated costs. The issue was swiftly solved, however, and only 6 middlemen (22% of middlemen requiring an assistant) did not have such assistance by mid-October. In other sections, where farmers were typically paid below the official cocoa price, middlemen refused to digitize their purchase receipts in the traceability application, on which the purchase price would then appear.

At cooperative-level, one-third of the treated cooperatives openly expressed reservations during initial project meetings, regarding the use of the DTS. Other cooperatives, which did not explicitly express their opposition, only began using the application late in the season, arguing that a succession of implementation problems constrained its use (e.g., internet connection issues, repeated issues with middlemen), although agents from the cocoa exporting company working within the cooperatives were most of the time unable to confirm the existence of these problems.

## **6.2 Use of the DTS**

Initial resistances by cooperatives and middlemen were mostly addressed, such that DTS was eventually used in every treated cooperative and 82% of the treated sections in the following six months (see Table 7). Overall, 9,634 transactions of certified cocoa were recorded on the DTS by middlemen, for a total of 2,827 tonnes of beans throughout the season.

However, although all 15 treated cooperatives used the application, they did so with very different intensity: one cooperative reported only one digitized transaction for the entire season, while the

**Table 8:** Use of the DTS across cooperatives and sections during the 2022-2023 main cocoa season

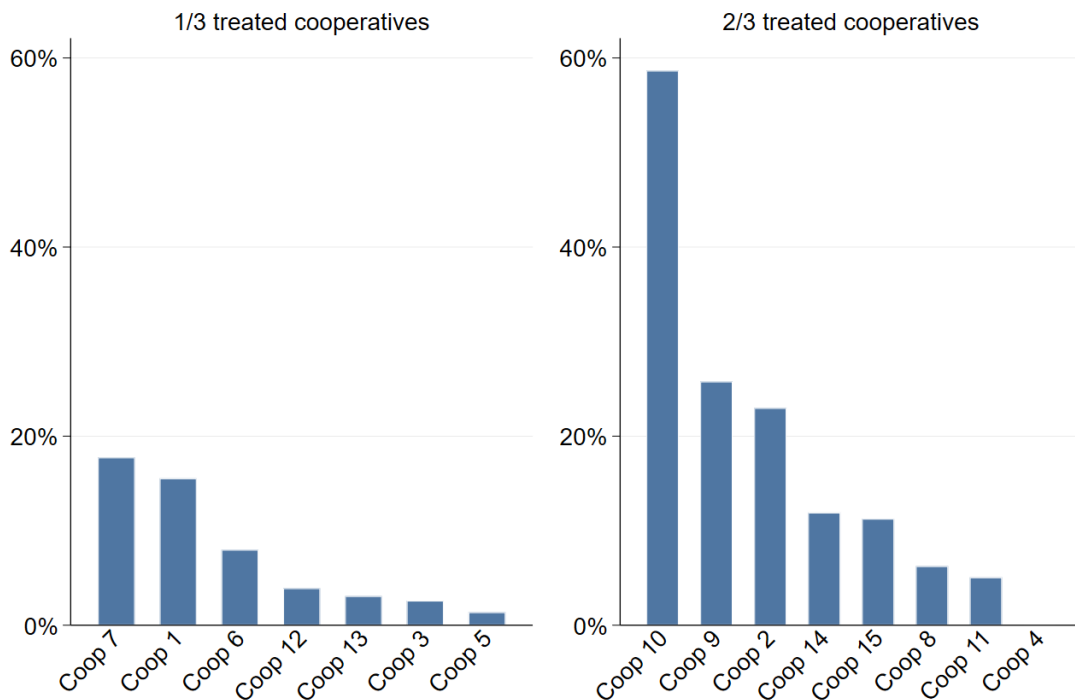
	Number of sections (1)	% of treated sections (2)	% treated sections which used DTS (3)	First use of DTS (4)	Latest start use of DTS (5)	% of cocoa reported in DTS (6)
Coop 1	11	1/3	75	Oct 02	Nov 03	15.53
Coop 2	5	2/3	100	Oct 24	Dec 07	22.95
Coop 3	17	1/3	100	Nov 04	Feb 07	2.59
Coop 4	4	2/3	33	Dec 19	19 Dec	0.04
Coop 5	9	1/3	100	Oct 03	10 Jan	1.37
Coop 6	7	1/3	100	Oct 05	Nov 05	7.98
Coop 7	18	1/3	100	Nov 18	Dec 15	17.74
Coop 8	6	2/3	100	Oct 20	Nov 09	6.24
Coop 9	8	2/3	100	Oct 27	Oct 27	25.78
Coop 10	9	2/3	100	Oct 02	Oct 05	58.32
Coop 11	27	2/3	79	Nov 18	Jan 22	5.06
Coop 12	11	1/3	75	Oct 25	Dec 28	3.91
Coop 13	12	1/3	50	Dec 07	Dec 26	3.08
Coop 14	8	2/3	100	Oct 02	Oct 25	11.91
Coop 15	5	2/3	100	Nov 09	Nov 09	11.26

Source: Authors' calculation.

**Note:** The proportion of cocoa reported in DTS corresponds to the ratio between the total volume digitized within treated sections and the total volume of cocoa supplied within those same sections, as reported in cooperative traceability registers.

most compliant one reported 3,076 transactions. In Table 8 we summarize the heterogeneity of DTS usage between treated cooperatives, and across the sections within each. Accordingly, five cooperatives (cooperatives #1, #5, #6, #10, and #14) started using the DTS in the first days of the season in at least one of their sections, while others (#4 and #13) started at least two months later (Column 4). This heterogeneity is further apparent in the overall volume of their transactions that were effectively recorded in the DTS. One-third of treated cooperatives digitized less than 5% of their cocoa purchases in the treated sections, 47% digitized between 5% and 20% of their purchases, and 20% digitized more than 20%. One cooperative (#10) digitized up to 60% of the volume originating from its treated sections (Column 6). Such heterogeneity is also found at the intra-cooperative level, between the different treated sections within a cooperative (Figure A2). These descriptive statistics further suggest that cooperatives where two-thirds of the sections were selected for the use of the DTS digitized more of their transactions than those where only one-third was selected (Figure 7).

Table 8 further shows significant intra-cooperative heterogeneity. Accordingly, the share of treated sections that effectively reported at least one transaction in the DTS varies from 33% to 100%. Further, while there are only a few days between the first and last section to start using the DTS in some

**Figure 7:** Proportion of volume from treated sections reported in the DTS, at cooperative-level

Source: Authors.

**Note:** The figure shows the proportion of the volume of cocoa reported in the DTS in treated sections, at cooperative-level.

cooperatives (e.g. #6, #8, or #10), one finds that some treated sections within the same cooperative started using the DTS up to 3 months apart (#3, #5, or #11).

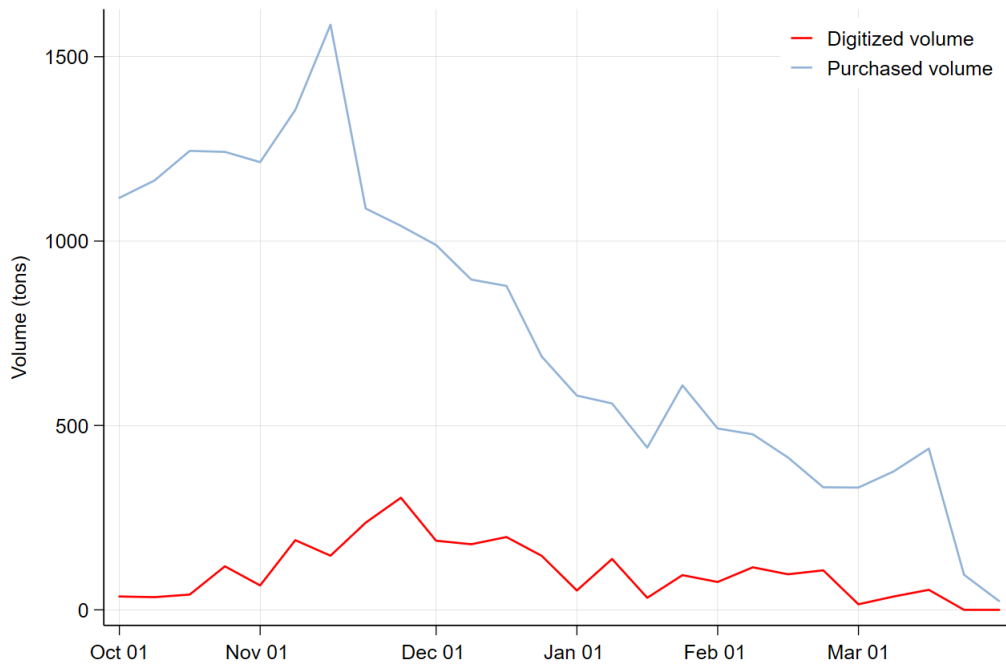
Figures reported in Table 8 are somewhat insufficient to characterize the incomplete use of the DTS at the section level. For instance, two sections may similarly digitize 50% of their transactions, although for very different reasons. For instance, one may have faced technical difficulties that delayed the use of DTS after which all transactions were digitized, while the other started using the DTS early on but only used it for part of its transactions. The strategically limited use of the DTS is more likely in the second case. In Figure 8, we report for each section the share of transactions digitized among those that were made since the first use of the DTS in the section. Accordingly, while sections may have started using the DTS at different times, one should expect that most transactions be digitized after that point such that most points be at 100% on the vertical axis.

Results in Figure 8 show otherwise, with the vast majority of sections having less than half of their transactions digitized after their first use of the DTS. In fact, one only finds three sections with more than 80% of their transactions digitized in the DTS after its first use in the section: two are from the same



in Figure 9 the number of digitized transactions had reached its peak by the end of November as more middlemen effectively started to use the application, after which it decreased. Observing less control by the exporter's agents and/or that many middlemen were only partly using the application without any related sanctions, most middlemen and/or cooperatives may have reduced their efforts to use the DTS.

**Figure 9:** Daily volume of cocoa digitized in the traceability application by treated sections



Source: Authors' calculation.

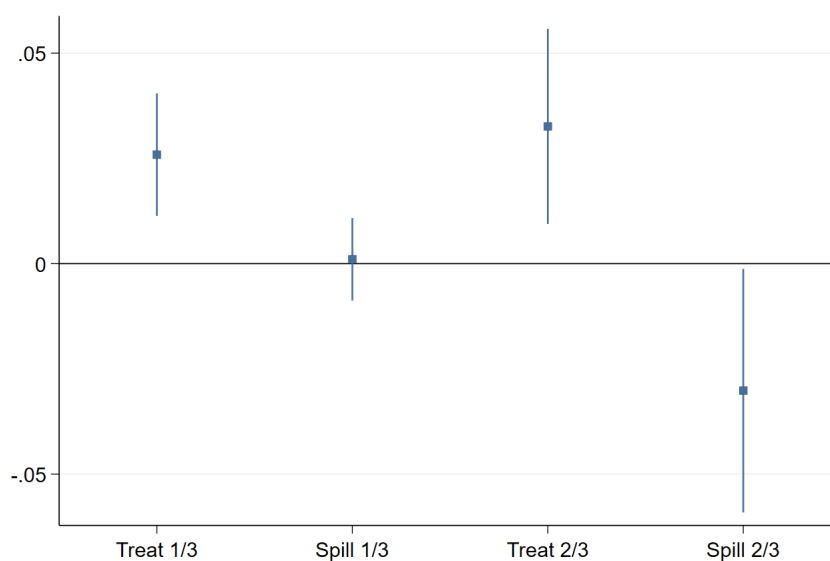
**Note:** The graph shows middlemen's use of the traceability application during the main cocoa season (red line), and the volume of cocoa purchased by cooperatives as reported in their final traceability registers (blue line), for treated sections.

A second possible explanation is one of selective entry into the use of the DTS, or selective use of the DTS for specific transactions. Accordingly, middlemen and/or cooperatives with more to lose from the use of the DTS would try and postpone its use until later in the season, and when they do start to use it, to limit its use to a subset of their transactions only.

In Figure 10 we report the parameter estimates for  $\hat{\beta}_1$ ,  $\hat{\beta}_2$ ,  $\hat{\beta}_3$  and  $\hat{\beta}_4$  from the estimation of Equation 3.2 at transaction-level. The outcome variable,  $y_{st}$  is a dummy variable indicating whether the transaction was digitized. The full set of results for this estimation is given in Table A3 in Appendix.

Figure 10 points to a significant effect of a section being allocated to treatment (i.e. its middlemen being trained and equipped with the DTS application). The effect is however small. On average, less than 5% of the transactions from treated sections were digitized, regardless of the treatment intensity ( $\theta = \frac{1}{3}$  or  $\theta = \frac{2}{3}$ ) at cooperative-level.



**Figure 10:** Treatment effects on transaction digitization

Source: Authors' calculation.

**Note:** This figure plots the coefficients of treatment variables from Equation 3.2 (Table A3, column 2). The sample includes all transactions reported in the cooperatives' final traceability registers. The dependent variable takes the value of 1 if the transaction has been digitized, and 0 otherwise.

**DTS use by value chain actors: key lessons.** Overall, results from this section point to unequal support of the DTS across actors of the cocoa value chain. While farmers appear supportive, some middlemen and some cooperatives were opposed to it in the early days of the season and only used it partially later on. While important heterogeneity exists across treated cooperatives, the same is true across treated sections of the same cooperative, further confirming that side-sourcing (which the DTS constrains) is found at both middleman- and cooperative-levels. Overall, the use of the DTS remained limited despite the exporters' team efforts early in the season, and late-season transactions were even less likely digitized than earlier ones.

## 7. Impact of traceability system

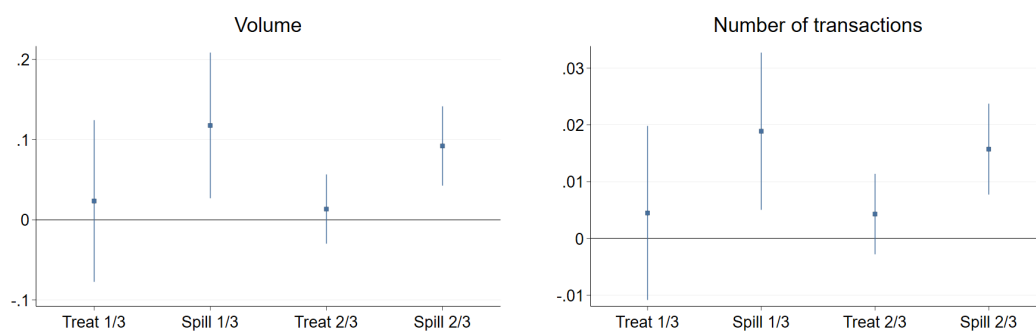
In this section, we investigate whether and how the DTS affected purchases and reporting amongst middlemen and cooperatives who were selected for its use in the 2022-2023 cocoa season.

## 7.1 Impact on purchases

Using the empirical framework described in Section 4 and the farmer surveys, we first estimate the Intent to Treat effects of the DTS on cocoa sales reported by farmers. Parameter estimates from Equations 3.2 are reported in Figure 11, where the dependent variable  $y_f$  is the volume (left panel) and the number of transactions (right panel) reported per surveyed farmers per week, respectively. In each figure, we report the estimated  $\hat{\beta}$  coefficients associated with the treatment of the section where the farmer is located ( $T_{\frac{1}{3}}^s, C_{\frac{1}{3}}^s, T_{\frac{2}{3}}^s, C_{\frac{2}{3}}^s$ ), the excluded category being the sections of the control cooperatives ( $C_0^s$ ).<sup>17</sup>

The results presented in Figure 11 suggest no significant effect of the DTS on farmers' sales in treated sections compared to pure control ones. We however find evidence of significantly higher volumes and number of transactions reported by farmers in "spillover" sections ( $C_{\frac{1}{3}}^s$  and  $C_{\frac{2}{3}}^s$ ). As the decision on which farmer to visit at what time is essentially in the hands of middlemen (possibly instructed by cooperatives), these results point to a strategic change of procurement decisions towards an increased reliance on non-treated sections of treated cooperatives. This is further supported by results from the middlemen surveys, in which they report purchases from surveyed farmers (Table A5 and Figure A3 in Appendix).

**Figure 11:** Treatment effects on volume and number of transactions reported per surveyed farmers per week



Source: Authors' calculation.

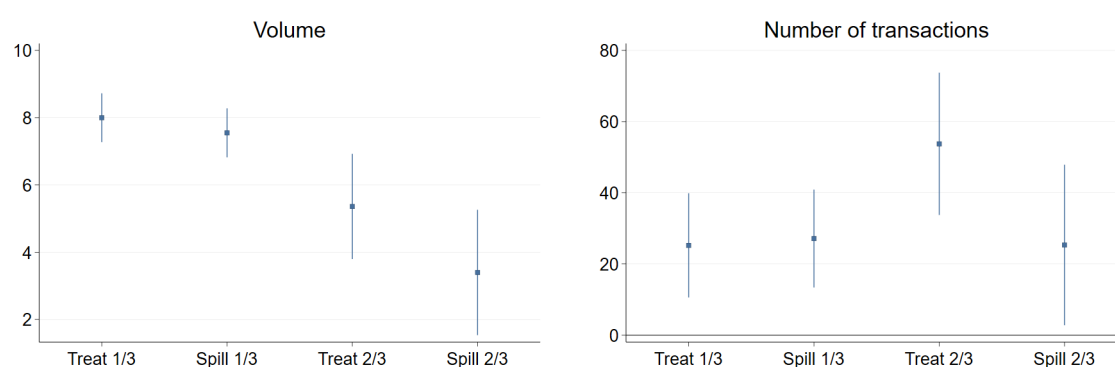
**Note:** The figure plots the coefficients for treatment variables from Equation 3.2 (Table A4, columns 2 and 5). The sample includes one observation per surveyed farmer per week. Dependent variables are volume in kilos using the inverse hyperbolic sine-transformation (IHS), and the number of transactions declared by farmers during the farmer survey.

<sup>17</sup>The full set of estimates are reported in Table A4 in Appendix.

## 7.2 Impact on reported volumes and transactions at cooperative-level

Next, we assess how the DTS intervention affected the reported cocoa purchases in the first (mid-season one) cooperatives' traceability registers. Cooperatives do manipulate the registers sent to the exporter throughout the season, in a way that end-of-season audits are less likely to spot suspiciously high volumes in some of the reported transactions (see Section 4). Thus, we assume that the first register is closer to reality in the sense that, while it may include conventional cocoa reported as certified one, cooperative agents have yet fully been able to hide it. Thus, our first set of estimates is based on data from the first traceability registers, with which we estimate the effect of treatment allocation on volume and the number of transactions reported per week at section-level, for the part half of the season.

**Figure 12:** Treatment effects on volume and number of transactions reported in first traceability registers



*Source:* Authors' calculation.

**Note:** The figure plots the coefficients for treatment variables from Equation 3.2 (Table A6, columns 2 and 5). The sample includes one observation per section per week. Dependent variables are volume in kilos using the inverse hyperbolic sine-transformation (IHS), and the number of transactions reported in the first traceability registers.

Figure 12 reports parameter estimates from Equation 3.2 estimated using traceability registers' data, where each observation is the per-week and per-section value of the outcome variable. We focus on the early months of the main cocoa season (October to mid-December), thereby covering the same period as the reports obtained from farmers. Therefore, meaningful differences in patterns between results in Figure 11 and those of Figure 12 are indicative of misreporting by middlemen and/or cooperatives.

We find substantial differences between treatment effects estimated on data from farmer surveys and those obtained from cooperatives' first traceability registers. Accordingly, registers from treated cooperatives show significant increases in reported volumes in all their sections (treated and non-

treated) as compared to those of control cooperatives. These effects appear larger when one-third of the sections were targeted by the DTS, as compared to those where two-thirds were targeted. The smallest - although positive - effects are found in control sections of two-thirds treated cooperatives. Results regarding the number of transactions reported in a week for a given section are also at odds with those obtained from the farmers' survey reports. All sections of treated cooperatives, irrespective of their own treatment status, show a larger number of transactions as compared to sections in control cooperatives. The effect is somewhat homogeneous across treated and spillover sections, except for treated sections in two-thirds treated cooperatives where the effect is twice as large.

At this stage, we cannot fully uncover strategic reasons for the substantial increase in section-level volumes and number of transactions that one finds in this first version of traceability registers for treated cooperatives, compared with control ones. In light of the results presented in Section 6, however, one can propose that cooperatives that were able to postpone the start of the DTS use (most of them), used the corresponding time to increase their sourcing of conventional cocoa, while reporting it as certified in the traceability registers. As the DTS was gradually being put into use, however, they increased the sourcing of conventional cocoa from the control sections of treated cooperatives.

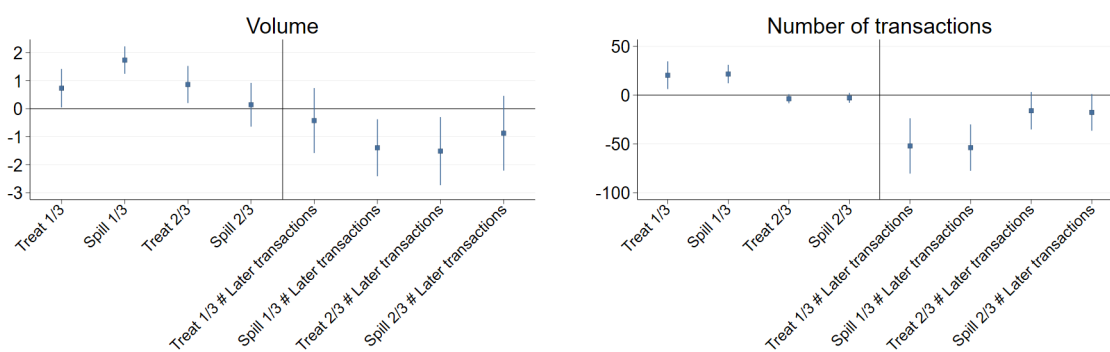
### 7.3 Impact on data manipulation

We next investigate whether DTS treatment at cooperative- and section-level led to changes in the manipulation of transaction data in cooperatives' registers, which we described in Section 4. To this end, we explore differences between mid-season cooperative registers that were shared with the exporter, with the end-of-season ones, focusing on the period common to both registers. To further facilitate comparisons with the period explored in the above analyses, we also interact all treatment variables with a dummy indicator indicating whether the considered transactions referred to the period also covered by the farmer surveys (October-November) or later (December-January).

**Section-level volumes and transactions.** We first estimate the impact of the DTS on variations in the key indicators of cooperative traceability registers between the first and final versions, at the section level. In Figure 13 we plot the parameter estimates from Equation 3.3, where the outcome variable is the differences in weekly section-level volume (left panel) and number of transactions (right panel), between the first and final traceability registers. In each panel, the left part plots the  $\hat{\beta}$  coefficients

associated with the treatment status of the section, while the right one plots the  $\hat{\delta}$  ones, corresponding to the additional effect of the transaction occurring later in the season.<sup>18</sup>

**Figure 13:** Treatment effects on the difference in volume and number of transactions between first and final traceability registers



Source: Authors' calculation.

**Note:** The figure plots the coefficients for treatment variables and interactions between treatment and transaction date from Equation 3.3 (Table A8, columns 3 and 6). The sample includes one observation per section per week, and transactions from October to mid-January. Dependent variables are the difference in transaction volume in kilos using the inverse hyperbolic sine-transformation (IHS), and the difference in the number of transactions reported in traceability registers. The "Later transactions" variable is a dummy taking the value of 1 if transactions took place in December and January, and 0 if they took place in October and November.

We find significant differences in reporting between the first and final traceability registers in terms of overall volume and number of transactions for one-third treated cooperatives, compared to control ones (Table A8). Sections in one-third treated cooperatives increase the overall volume initially reported in the first traceability register during the season, and mainly inflate transactions that took place at the start of the cocoa season (October/November) (Figure 13, left panel). We find a similar pattern regarding the number of transactions, as the sections in one-third treated cooperatives increase the number of transactions made during the period covered by the first traceability registers, especially at the start of the season (Figure 13, right panel). Findings are less conclusive for the two-thirds treated cooperatives, for which there is no significant variation in terms of the number of transactions between the two traceability registers compared to the sections of control cooperatives. The variation in terms of volume between the first and final traceability register is also only significant for the treated sections.

Overall, treated cooperatives report significantly more volume and transactions than control cooperatives in their first (and final) traceability registers (Tables A6 and A7). Additionally, treated

<sup>18</sup>Thus, the left part of each panel corresponds to the treatment effects associated with the part of the season also covered by the farmer survey.

cooperatives make more changes to their registers during the season than control ones. This is especially true for one-third treated cooperatives, which inflate their reported volume as well as the number of transactions, while the effect is much smaller for two-thirds treated cooperatives, which adjust their initially reported volume without increasing the number of transactions.

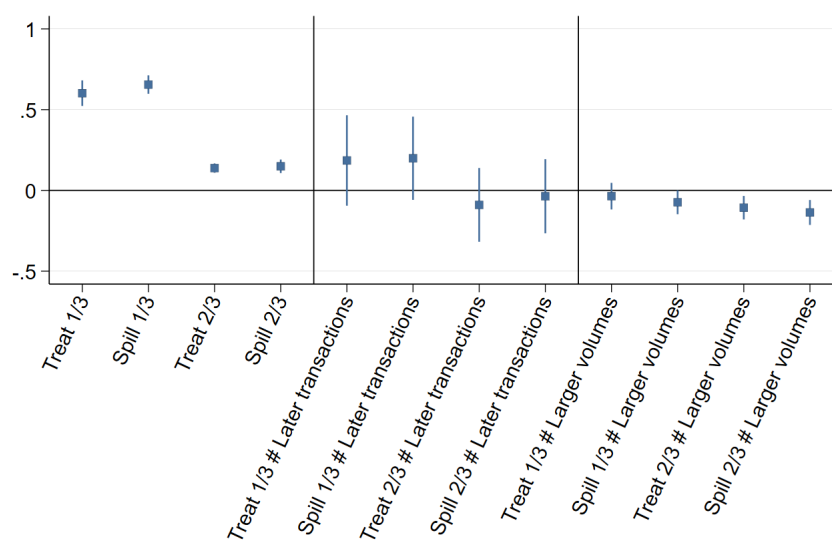
**Tracking transactions.** Next, we investigate whether the promotion of the DTS in randomly selected cooperatives and sections affected one's capacity to track transactions across the first and final traceability registers shared by the cooperatives to the exporter. Specifically, we use all transactions reported in the final traceability registers that relate to the period covered by the first ones, and assess whether they can effectively be found in the latter.

We report treatment parameters from Equation 3.3 where the outcome variable is a dummy indicating whether transactions reported in the final traceability register can be found perfectly or imperfectly in the first traceability register.<sup>19</sup> We further perform similar analyses using as a dependent variable whether a transaction in the last traceability register perfectly matches a transaction in the first register, vs. imperfectly matches (see Table A9 columns 4,5, and 6, and Figure A5).

We find that there is a greater probability that information on a transaction reported in treated cooperatives matches between the first and final register, than in control cooperatives (Table A9, columns 1 and 2). The effect is particularly large in sections of one-third treated cooperatives, where the probability of a match is 66% to 71% more likely than in the pure control sections. It remains positive but much more limited for sections of two third-treated cooperatives. That both control and treated sections in treatment cooperatives exert similar treatment effects is a clear indication of cooperative-level actions in strategic misreporting.

That one-third treated and two-thirds treated cooperatives show significantly different patterns further suggests that treatment intensity yielded different data manipulation behaviors across the cooperatives. Figure 14 is also illustrative of these aspects, showing that larger transactions reported in sections of two-thirds treated cooperatives are less likely to be matched across the two registers, while the effect is much lower and insignificant for sections in one-third treated cooperatives. Although the results are not significant, findings plotted in Figure 14 further suggest that late transactions in one-third

<sup>19</sup>We define a perfect match as transactions for which the same date, the same farmer, and the same volume can be found in both registers, and an imperfect match as transactions for which we find the same farmer and the same date, but different volumes, or the same farmer and the same volume, but a different date, between the first and final traceability registers.

**Figure 14:** Treatment effect on transaction matching between first and final traceability registers

Source: Authors' calculation.

**Note:** The figure plots the coefficients for treatment variables, interactions between treatment and transaction date, and interactions between treatment and transaction volume, from equation (3) (Table A9). The sample includes all transactions reported in final traceability registers matching the timeframe of first traceability registers. The "Later transactions" variable is a dummy taking the value of 1 if the transaction took place at a later time than the average transaction date, i.e. after November 3, computed on the sample of transactions in the final traceability registers that fall into the period covered by the first traceability register. The "Larger volumes" variable is a dummy taking the value of 1 if the transaction volume is greater than the average volume of transactions recorded in the final traceability registers, i.e. greater than 188 kilos.

treated cooperatives are more likely to be matched between registers, whereas no such effect is found for two-thirds treated cooperatives.

Overall, findings show that transactions reported by sections (treated and non-treated) of treated cooperatives are more likely to be matched between the first and final traceability registers than transactions in sections of control cooperatives. This effect is much greater for sections in one-third treated cooperatives, where transactions made at the end of the season are also more likely to be matched between the two traceability registers. Taken together, findings from section 7 suggest different misreporting patterns at cooperative-level according to treatment status. Sections in one-third treated cooperatives modify less the transactions initially reported in their first traceability registers (leading to a greater proportion of transactions matching between first and final traceability registers), as their data modification strategy relies mainly on adding new (potentially fraudulent) transactions (Figure 13 and Table A8, columns 4, 5, and 6). In contrast, sections in two-thirds treated cooperatives add a moderate number of new transactions between their first and final traceability registers. Thus, their only option for adjusting (potentially side-sourced) volumes between mid- and end-season traceability

registers is to modify existing transactions reported in the first one (leading to a lower proportion of transactions matching between first and final registers).

## 8. Conclusion

This paper investigates the extent to which ESS-certified cocoa in Côte d'Ivoire genuinely originates from ESS-certified farmers. Specifically, we examine whether and to what extent upstream supply chain actors, such as farmers, middlemen, and cooperatives, engage in side-sourcing, i.e. sourcing conventional cocoa but labeling it as certified. In partnership with a top 5 global cocoa exporting company, we also take advantage of the implementation of a digital traceability system (DTS) aimed at reducing opacity in transactions between farmers and middlemen, and thus, cooperatives, to observe its potential effects on the levels and patterns of side-sourcing practices by value chain actors.

We rely on a large, original, and unique set of data. First, we use farmer and middleman survey data about cocoa transactions, as well as cooperative traceability registers, where cooperatives report each certified cocoa transaction and its characteristics. We triangulate information on alleged certified cocoa transactions from these three sources of information and investigate the existence of reporting differences between supply chain actors. We interpret reporting differences as a potential indication of side-sourcing. Second, we assess the intent-to-treat effect of implementing a DTS on a range of outcomes: cocoa purchases by middlemen from certified farmers, the volume and number of transactions reported by cooperatives in their traceability registers, and the data manipulation of traceability registers during the cocoa season by cooperatives. We rely on a randomized controlled trial, where a random subset of sections was selected (treated) from a random subset of (treated) cooperatives for their middlemen to use the DTS, during the 2022-2023 main cocoa season. Within these treated cooperatives, we further varied the share of sections selected for treatment: either one-third or two-thirds.

We find large reporting differences in terms of volumes declared as sold/purchased by certified farmers and middlemen on the one hand, and the certified volumes purchased from these same farmers over the same period by cooperatives on the other. More specifically, we find that cooperatives record a volume of certified cocoa three times greater than that declared by farmers and middlemen. These reporting



differences are found within all cooperatives and sections, indicating a widespread phenomenon of misreporting at the cooperative level, likely to hide side-sourcing practices.

The implementation of a DTS by the exporter in the upstream part of its value chain raised reluctance among some middlemen and cooperatives, leading to a heterogeneous, albeit relatively low, intensity of use of the DTS. As a result, some cooperatives have digitized up to 58% of their certified cocoa purchases in the traceability tool, while other treated cooperatives have barely digitized any.

Yet, we still find evidence of several DTS implementation effects on treated cooperatives. First, we show that farmers in treated sections have sold relatively less cocoa to middlemen working with cooperatives implementing the DTS, compared to control sections in treated cooperatives. Similarly, middlemen in treated sections have bought less cocoa from farmers in treated sections, compared to middlemen and farmers in control sections of treated cooperatives, indicating a strategic shift in cocoa sourcing by middlemen (potentially driven by cooperatives). Second, our results show that sections of treated cooperatives report significantly higher volumes and more transactions in their first traceability register, compared with sections of control cooperatives. One possible explanation for this finding is that cooperatives adopted the DTS rather later than planned, providing them with the opportunity to increase their supply of conventional cocoa at the beginning of the cocoa season while declaring it certified in the traceability registers. Finally, the DTS implementation also increases data manipulation of traceability registers at cooperative-level, but cooperatives engage in different misreporting strategies depending on treatment intensity. Sections in one-third treated cooperatives modify less the transactions initially reported in their first traceability registers, resulting in more transactions being matched between the first and final traceability registers, as their data modification strategy relies mainly on adding new (potentially fraudulent) transactions. Sections in two-thirds treated cooperatives add a moderate number of new transactions between their first and final traceability registers, but rather modify existing transactions reported in the first traceability register to adjust their volumes between mid- and end-season traceability registers. This leads to a lower proportion of transactions matching between their first and final registers.

This paper comes with limitations. First, our analyses are based on a relatively small sample of certified cooperatives, although representative of our partner's supply chain. Second, due to constraints imposed by our partner exporting company, we were not able to implement the DTS in 100% of the treated cooperatives' sections, but only in one- or two-thirds of the sections. This may represent an important

limitation, as misreporting mechanisms and sourcing practices from cooperatives are likely to differ in such cases. Finally, the DTS take-up was rather low, considering that not all transactions made by treated middlemen were digitized. The effects measured in this paper are therefore more the result of the "announcement effect" (and the potential associated fear of increased monitoring of transaction information and traceability registers in treated cooperatives), than a direct effect of the DTS.

This work suggests that further research is needed to better identify, measure, and mitigate side-sourcing in certified agricultural value chains. In particular, future studies should aim to properly identify the actors responsible for side-sourcing, and to better measure the extent to which middlemen and cooperatives are individually accountable for such behavior. This paper also highlights the complexity of designing and implementing a digital traceability tool, at a time when stakeholders in the cocoa sector (exporting companies, certification bodies, as well as the Ivorian government) consider it a promising option for achieving traceability at farmer-level.

## References

- Akoyi, K. T. and Maertens, M. (2018). Walk the talk: private sustainability standards in the Ugandan coffee sector. *The Journal of Development Studies*, 54(10):1792–1818. [108]
- Almunia, M., Hjort, J., Knebelmann, J., and Tian, L. (2024). Strategic or confused firms? evidence from “missing” transactions in Uganda. *Review of Economics and Statistics*, 106(1):256–265. [109]
- Astill, J., Dara, R. A., Campbell, M., Farber, J. M., Fraser, E. D., Sharif, S., and Yada, R. Y. (2019). Transparency in food supply chains: A review of enabling technology solutions. *Trends in Food Science & Technology*, 91:240–247. [109]
- Bai, C., Quayson, M., and Sarkis, J. (2022). Analysis of blockchain’s enablers for improving sustainable supply chain transparency in africa cocoa industry. *Journal of Cleaner Production*, 358:131896. [109]
- Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American philosophical society*, pages 551–572. [121]
- Bergquist, L. F. and Dinerstein, M. (2020). Competition and entry in agricultural markets: Experimental evidence from kenya. *American Economic Review*, 110(12):3705–3747. [108]
- Bergquist, L. F., McIntosh, C., and Startz, M. (2021). *Search cost, intermediation, and trade: Experimental evidence from Ugandan agricultural markets*. eScholarship, University of California. [108]
- Boutin, D., Jouvin, M., and Olié, L. (2024). Assessing dishonesty in cocoa value chains: Lab-in-the-field evidence from middlemen in côte d’ivoire. [106], [109]
- Bymolt, R., Laven, A., and Tyzler, M. (2018). Demystifying the cocoa sector in ghana and côte d’ivoire. *The Royal Tropical Institute (KIT): Amsterdam, The Netherlands*. [110], [111]
- Casaburi, L. and Reed, T. (2022). Using individual-level randomized treatment to learn about market structure. *American Economic Journal: Applied Economics*, 14(4):58–90. [108]
- DeFries, R. S., Fanzo, J., Mondal, P., Remans, R., and Wood, S. A. (2017). Is voluntary certification of tropical agricultural commodities achieving sustainability goals for small-scale producers? a review of the evidence. *Environmental research letters*, 12(3):033001. [108]
- Demir, B. and Javorcik, B. (2020). Trade policy changes, tax evasion and Benford’s law. *Journal of Development Economics*, 144:102456. [121]
- Gardner, T. A., Benzie, M., Börner, J., Dawkins, E., Fick, S., Garrett, R., Godar, J., Grimard, A., Lake, S., Larsen, R. K., et al. (2019). Transparency and sustainability in global commodity supply chains. *World Development*, 121:163–177. [109]
- Gualandris, J., Longoni, A., Luzzini, D., and Pagell, M. (2021). The association between supply chain structure and transparency: A large-scale empirical study. *Journal of Operations Management*, 67(7):803–827. [109]
- Hao, J., Bijman, J., Gardebroek, C., Heerink, N., Heijman, W., and Huo, X. (2018). Cooperative membership and farmers’ choice of marketing channels—Evidence from apple farmers in Shaanxi and Shandong Provinces, China. *Food Policy*, 74:53–64. [108]
- International Cocoa Organization (2024). Quarterly Bulletin of Cocoa Statistics. Technical Report 1, International Cocoa Organization. [110], [111]

- Kaiser, M. (2019). Benford's law as an indicator of survey reliability—can we trust our data? *Journal of Economic Surveys*, 33(5):1602–1618. [121]
- Kalischek, N., Lang, N., Renier, C., Daudt, R. C., Addoah, T., Thompson, W., Blaser-Hart, W. J., Garrett, R., Schindler, K., and Wegner, J. D. (2023). Cocoa plantations are associated with deforestation in Côte d'Ivoire and Ghana. *Nature Food*, 4(5):384–393. [111]
- Kozicka, M., Tacconi, F., Horna, D., and Gotor, E. (2018). Forecasting cocoa yields for 2050. [119]
- Lambin, E. F. and Thorlakson, T. (2018). Sustainability standards: Interactions between private actors, civil society, and governments. *Annual Review of Environment and Resources*, 43(1):369–393. [108]
- Meemken, E.-M. (2020). Do smallholder farmers benefit from sustainability standards? a systematic review and meta-analysis. *Global Food Security*, 26:100373. [108]
- Meemken, E.-M., Barrett, C. B., Michelson, H. C., Qaim, M., Reardon, T., and Sellare, J. (2021). Sustainability standards in global agrifood supply chains. *Nature Food*, 2(10):758–765. [106]
- Mengistie, B. T., Mol, A. P., and Oosterveer, P. (2017). Governance of agro-pesticide through private environmental and social standards in the global cut flower chain from ethiopia. *Ambio*, 46(7):797–811. [106]
- Nitidae and EFI (2021). Traceability and transparency of cocoa supply chains in Côte d'Ivoire and Ghana. Technical report, Nitidae and EFI. [109]
- Oya, C., Schaefer, F., and Skalidou, D. (2018). The effectiveness of agricultural certification in developing countries: A systematic review. *World Development*, 112:282–312. [108]
- Ponte, S. and Gibbon, P. (2005). Quality standards, conventions and the governance of global value chains. *Economy and society*, 34(1):1–31. [108]
- Renier, C., Vandromme, M., Meyfroidt, P., Ribeiro, V., Kalischek, N., and Zu Ermgassen, E. K. (2023). Transparency, traceability and deforestation in the ivorian cocoa supply chain. *Environmental Research Letters*, 18(2):024030. [110]
- Rustagi, D. and Kroell, M. (2022). Measuring honesty and explaining adulteration in naturally occurring markets. *Journal of Development Economics*, 156:102819. [106], [109]
- Sadhu, S., Kysia, K., Onyango, L., Zinnes, C., Lord, S., Monnard, A., and Arellano, I. R. (2020). Sadhu et al., 2020. Technical report, National Opinion Research Center (NORC), University of Chicago. [111]
- Sambridge, M. and Jackson, A. (2020). National COVID numbers—Benford's law looks for errors. *Nature*, 581(7809):384–385. [121]
- Sellare, J., Meemken, E.-M., Kouamé, C., and Qaim, M. (2020). Do sustainability standards benefit smallholder farmers also when accounting for cooperative effects? evidence from côte d'ivoire. *American Journal of Agricultural Economics*, 102(2):681–695. [108]
- Shumeta, Z., D'Haese, M., and Verbeke, W. (2018). A two-step econometric estimation of covariates of side selling: The case of coffee cooperatives in Southwest Ethiopia. *The Journal of Development Studies*, 54(10):1775–1791. [108]
- Stoop, P., Ramanan, N., Geens, H., Lambrecht, A., and Dekeister, S. (2021). Technical brief on cocoa traceability in West and Central Africa. [106], [108], [109]
- Taylor, P.L. (2005). In the market but not of it: Fair trade coffee and forest stewardship council certification as market-based social change. *World development*, 33(1):129–147. [108]

Wollni, M. and Fischer, E. (2015). Member deliveries in collective marketing relationships: Evidence from coffee cooperatives in Costa Rica. *European Review of Agricultural Economics*, 42(2):287–314. [108]

World Bank (2019). Au pays du cacao comment transformer la Côte d'Ivoire. page 64. [106], [109], [111]

## **Appendix**

### **Appendix A: Balance tests**

#### **Farmer survey**

Table A1: Basic characteristics of surveyed farmers across groups

	(1) Control		(2) Spillover		(3) Treated		T-test P-value		
	N	Mean (std. dev.)	N	Mean (std. dev.)	N	Mean (std. dev.)	(1)-(2)	(1)-(3)	(2)-(3)
Male	405	0.941 (0.015)	696	0.934 (0.011)	551	0.918 (0.013)	0.703	0.252	0.340
Farmer age	405	43.277 (0.372)	696	43.336 (0.603)	551	43.405 (0.432)	0.932	0.819	0.913
Married	405	0.983 (0.005)	696	0.984 (0.008)	551	0.966 (0.009)	0.874	0.093*	0.173
Ivorian	405	0.674 (0.082)	696	0.644 (0.041)	551	0.704 (0.033)	0.733	0.725	0.172
Has been to school	405	0.509 (0.045)	696	0.473 (0.024)	551	0.470 (0.028)	0.468	0.455	0.939
Has another source of income	405	0.551 (0.021)	696	0.579 (0.043)	551	0.554 (0.032)	0.557	0.939	0.379
Household size	405	7.560 (0.144)	696	7.454 (0.125)	551	7.690 (0.138)	0.569	0.512	0.228
Has an active mobile money account	405	0.630 (0.034)	696	0.717 (0.024)	551	0.657 (0.023)	0.042**	0.498	0.020***
Number of farms	405	1.052 (0.010)	696	1.082 (0.015)	551	1.082 (0.015)	0.097*	0.112	0.991
Overall farm size	405	3.367 (0.101)	696	3.386 (0.077)	551	3.279 (0.064)	0.880	0.450	0.334
Report usually newer engaging in side-selling	202	0.619 (0.087)	347	0.594 (0.066)	275	0.753 (0.071)	0.814	0.230	0.002***
The farmer is patient	405	0.546 (0.027)	696	0.568 (0.030)	551	0.568 (0.033)	0.584	0.594	0.992
The farmer is risk-averse	405	0.272 (0.009)	696	0.299 (0.014)	551	0.279 (0.022)	0.114	0.743	0.461
Proportion of farmers having been surveyed during round 1	405	0.501 (0.001)	696	0.501 (0.001)	551	0.501 (0.001)	0.889	0.822	0.522

Source: Authors' calculation.

Note: The values displayed for t-tests are p-values. Standard errors are reported in parentheses and clustered at the cooperative level. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent critical level.

## Trader survey



Table A2: Basic characteristics of surveyed traders across groups

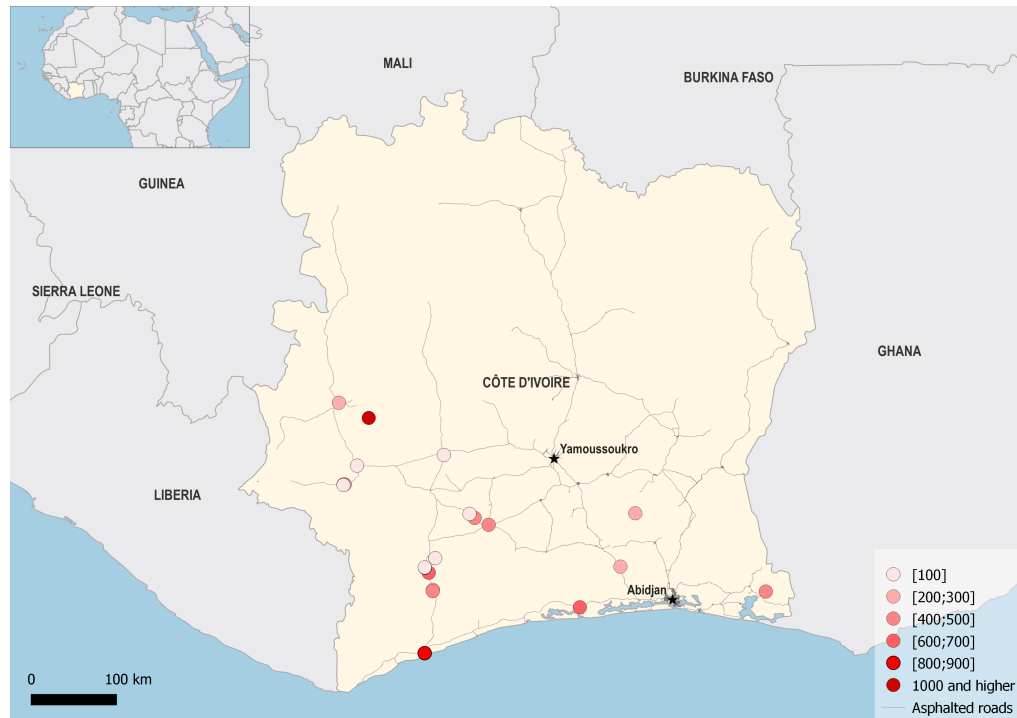
	(1) Control		(2) Treated		(3) Total		T-test P-value (1)-(2)
	N	Mean/SE	N	Mean/SE	N	Mean/SE	
Main trader	194	0.794 (0.071)	113	0.761 (0.064)	307	0.782 (0.053)	0.711
Age	194	42.794 (0.580)	112	43.920 (0.847)	306	43.206 (0.478)	0.301
Ivorian	194	0.649 (0.046)	113	0.681 (0.056)	307	0.661 (0.039)	0.627
Has some education	194	0.701 (0.042)	112	0.795 (0.028)	306	0.735 (0.033)	0.051*
Trader is also a cocoa farmer	194	0.809 (0.037)	113	0.814 (0.043)	307	0.811 (0.031)	0.921
Years of experience for current cooperative	194	6.067 (0.574)	112	5.589 (0.400)	306	5.892 (0.419)	0.311
Has a mobile money account	194	0.938 (0.020)	112	0.955 (0.022)	306	0.944 (0.016)	0.532
Sometimes pay farmers through mobile money	168	0.167 (0.032)	101	0.178 (0.038)	269	0.171 (0.028)	0.782
Own a motorcycle	194	0.861 (0.038)	113	0.788 (0.047)	307	0.834 (0.030)	0.222
Get help from family members	194	0.572 (0.057)	113	0.496 (0.048)	307	0.544 (0.042)	0.299
Hire extra workers	194	0.753 (0.039)	113	0.681 (0.058)	307	0.726 (0.030)	0.366
Risk lover	194	0.278 (0.028)	113	0.345 (0.046)	307	0.303 (0.023)	0.254
Commission per kilos of cocoa (FCFA)	194	39.459 (2.075)	113	37.920 (1.529)	307	38.893 (1.620)	0.480
The cooperative set a volume target	194	0.284 (0.027)	113	0.336 (0.065)	307	0.303 (0.028)	0.467

Source: Authors' calculation.

Note: The values displayed for t-tests are p-values. Standard errors are reported in parentheses and clustered at the cooperative level. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent critical level.

## Appendix B: Cooperatives misreporting

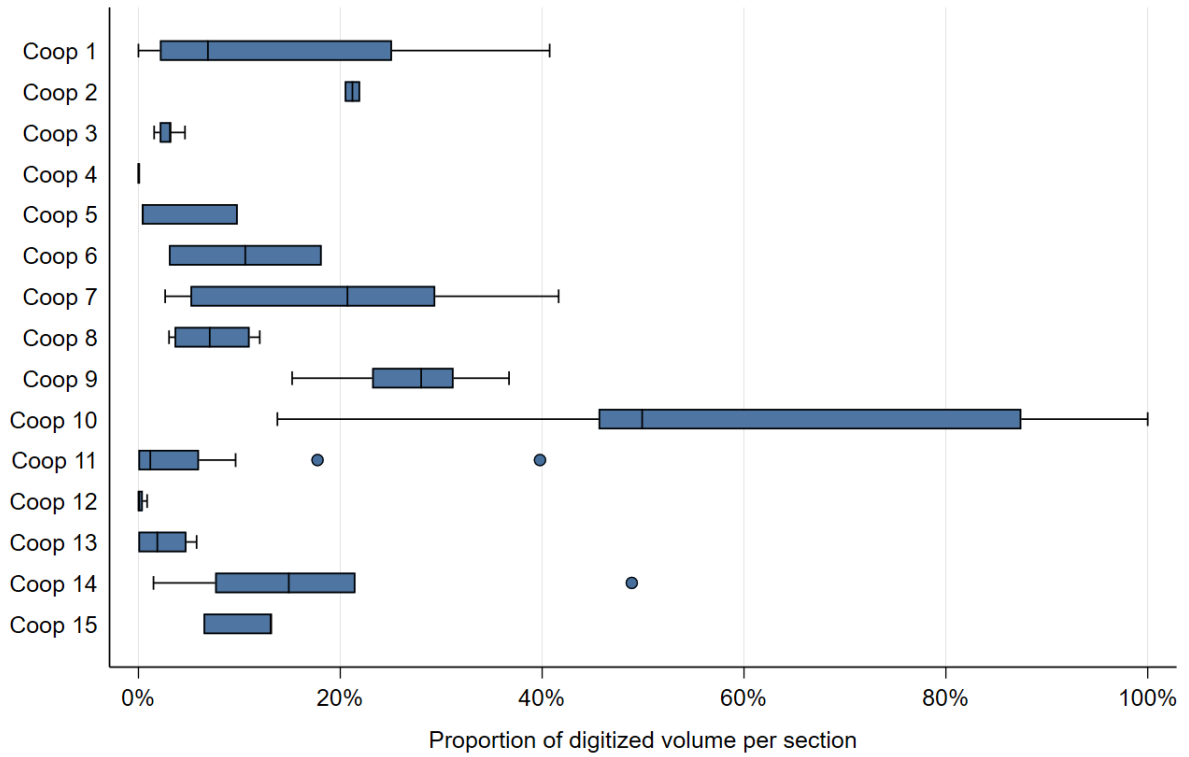
**Figure A1:** Geographical distribution of cooperatives and volume reporting surplus (as a % of volume reported by farmers)



Source: Authors.

## Appendix C: Compliance

**Figure A2:** Proportion of digitized volume across sections, by cooperative



Source: Authors' calculation.

## Appendix D: DTS usage

**Table A3:** Treatment effect on transaction digitization

	(1)	(2)	(3)
Treat	0.038*** (0.008)		
Spill	-0.007 (0.007)		
Late transaction	-0.022*** (0.006)	-0.022*** (0.006)	-0.000 (0.000)
Large volume	0.006* (0.003)	0.007* (0.003)	0.000 (0.000)
Treat 1/3		0.026*** (0.007)	0.031** (0.012)
Spill 1/3		0.001 (0.005)	-0.002 (0.004)
Treat 2/3		0.033** (0.012)	0.052** (0.017)
Spill 2/3		-0.030* (0.015)	-0.042** (0.014)
Treat 1/3 × Late transaction			-0.032** (0.012)
Spill 1/3 × Late transaction			-0.008* (0.004)
Treat 2/3 × Late transaction			-0.064** (0.019)
Spill 2/3 × Late transaction			0.008 (0.007)
Treat 1/3 × Large volume			-0.006 (0.008)
Spill 1/3 × Large volume			0.001 (0.002)
Treat 2/3 × Large volume			0.029** (0.010)
Spill 2/3 × Large volume			-0.009 (0.007)
Cooperative FE	Yes	Yes	Yes
N	276,938	276,938	276,938
R2	0.100	0.121	0.119

Source: Authors' calculation.

**Note:** Standard errors in parenthesis, clustered at section level. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . The sample includes all transactions reported in the cooperatives' final traceability registers. The dependent variable is a dummy taking the value of 1 if the transaction has been reported in the digital traceability tool, 0 otherwise. The "Late transaction" variable is a dummy taking the value of 1 if the transaction took place at a later time than the average transaction Late transaction, i.e. after December 4, computed on the sample of transactions in the final traceability registers. The "Large volume" variable is a dummy taking the value of 1 if the transaction volume is greater than the average volume of transactions recorded in the final traceability registers, i.e. greater than 188 kilos.

## Appendix E: Impact of digital traceability system

### At farmer-level

**Table A4:** Treatment effect on volume and number of transactions reported per surveyed farmers per week

	Volume			Number of trans.		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	0.028 (0.048)			0.005 (0.007)		
Spill	0.115* (0.045)			0.018** (0.007)		
Late transaction	0.044* (0.021)	0.044* (0.021)	2.460 (3.078)	0.006 (0.003)	0.006 (0.003)	0.001 (0.007)
Treat 1/3		0.023 (0.051)	2.238 (3.946)		0.004 (0.008)	0.005 (0.009)
Spill 1/3		0.118* (0.046)	3.879 (3.080)		0.019** (0.007)	0.015* (0.007)
Treat 2/3		0.013 (0.022)	-2.540 (1.908)		0.004 (0.004)	0.004 (0.004)
Spill 2/3		0.092*** (0.025)	5.898 (3.579)		0.016*** (0.004)	0.017* (0.007)
Treat 1/3 × Late transaction			-1.643 (4.657)			0.004 (0.011)
Spill 1/3 × Late transaction			4.237 (4.390)			0.011 (0.010)
Treat 2/3 × Late transaction			1.709 (3.694)			0.006 (0.009)
Spill 2/3 × Late transaction			-1.703 (5.822)			0.003 (0.012)
Cooperative FE	Yes	Yes	Yes	Yes	Yes	Yes
N	16,520	16,520	16,520	16,520	16,520	16,520
R2	0.035	0.035	0.034	0.025	0.035	0.035

Source: Authors' calculation.

**Note:** Standard errors in parenthesis, clustered at section level. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . The sample corresponds to one observation per surveyed farmer per week. The sample includes transactions from October to mid-December. The dependent variable "Volume" corresponds to the inverse hyperbolic sine-transformation (IHS) of volume (in kilos). The "Late transaction" variable is a dummy taking the value of 1 if transactions reported by farmers took place in November and December, and 0 if they took place in October.

## At middleman-level

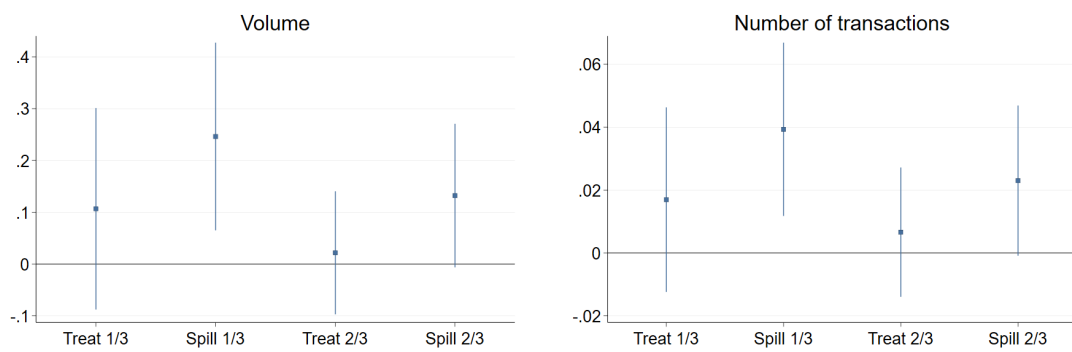
**Table A5:** Treatment effect on volume and number of transactions reported per farmers per week by middlemen

	Volume			Number of trans.		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	0.116 (0.093)			0.019 (0.014)		
Spill	0.241** (0.090)			0.038** (0.014)		
Late transaction	0.131*** (0.032)	0.131*** (0.032)	12.972** (4.630)	0.020*** (0.005)	0.020*** (0.005)	0.026** (0.010)
Treat 1/3		0.107 (0.099)	7.582 (8.500)		0.017 (0.015)	0.010 (0.016)
Spill 1/3		0.246** (0.092)	14.427* (6.614)		0.039** (0.014)	0.040** (0.014)
Treat 2/3		0.022 (0.060)	-0.949 (2.812)		0.007 (0.010)	0.015 (0.011)
Spill 2/3		0.132 (0.070)	6.246 (5.023)		0.023 (0.012)	0.022 (0.015)
Treat 1/3 × Late transaction			-5.176 (6.838)			0.006 (0.015)
Spill 1/3 × Late transaction			-5.682 (6.709)			-0.008 (0.014)
Treat 2/3 × Late transaction			-7.599 (5.785)			-0.020 (0.013)
Spill 2/3 × Late transaction			-8.505 (7.809)			-0.004 (0.018)
Cooperative FE	Yes	Yes	Yes	Yes	Yes	Yes
N	16,520	16,520	16,520	16,520	16,520	16,520
R2	0.062	0.062	0.062	0.047	0.062	0.062

Source: Authors' calculation.

**Note:** Standard errors in parenthesis, clustered at section level. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . The sample corresponds to one observation per surveyed farmer per week. The sample includes transactions from October to mid-December. The dependent variable "Volume" corresponds to the inverse hyperbolic sine-transformation (IHS) of volume (in kilos). The "Late transaction" variable is a dummy taking the value of 1 if transactions reported by farmers took place in November and December, and 0 if they took place in October.

**Figure A3:** Treatment effects on volume and number of transactions reported per surveyed farmers per week by middlemen



Source: Authors' calculation.

**Note:** The figure plots the coefficients for treatment variables from Equation 3.2 (Table A5, columns 2 and 5). The sample includes one observation per surveyed farmer per week, and transactions from October to mid-December. Dependent variables are volume in kilos using the inverse hyperbolic sine-transformation (IHS), and the number of transactions declared by middlemen during the middleman survey.



## At cooperative-level

**Table A6:** Treatment effect on volume and number of transactions reported per section per week in first traceability registers

	Volume			Number of trans.		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	7.894*** (0.312)			34.803*** (4.863)		
Spill	6.769*** (0.316)			22.716*** (4.618)		
Late transaction	-3.063*** (0.292)	-3.063*** (0.292)	-0.359 (0.644)	-18.837*** (3.818)	-18.837*** (3.819)	-10.159 (8.559)
Treat 1/3		7.407*** (0.264)	9.662*** (0.322)		26.133*** (4.845)	43.257*** (9.174)
Spill 1/3		7.047*** (0.289)	8.714*** (0.402)		27.670*** (4.457)	37.857*** (6.982)
Treat 2/3		5.434*** (0.635)	4.473*** (0.641)		50.344*** (7.579)	37.897*** (7.410)
Spill 2/3		3.518*** (0.770)	2.620** (0.836)		24.139** (8.778)	16.444 (8.930)
Treat 1/3 × Late transaction			-6.462*** (0.781)			-37.218** (14.287)
Spill 1/3 × Late transaction			-5.481*** (0.812)			-25.656* (12.000)
Treat 2/3 × Late transaction			-1.100 (0.708)			12.067 (9.720)
Spill 2/3 × Late transaction			-1.206 (0.741)			4.147 (9.481)
Cooperative FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2,140	2,140	2,140	2,140	2,140	2,140
R2	0.680	0.682	0.709	0.347	0.351	0.360

Source: Authors' calculation.

**Note:** Standard errors in parenthesis, clustered at section level. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . The sample includes one observation per section per week, based on transactions in first traceability registers. The sample includes transactions from October to mid-December, covering the same period as the farmer survey. The dependent variable "Volume" corresponds to the inverse hyperbolic sine-transformation (IHS) of volume (in kilos). The "Late transaction" variable is a dummy taking the value of 1 if transactions took place in November and December, and 0 if they took place in October.

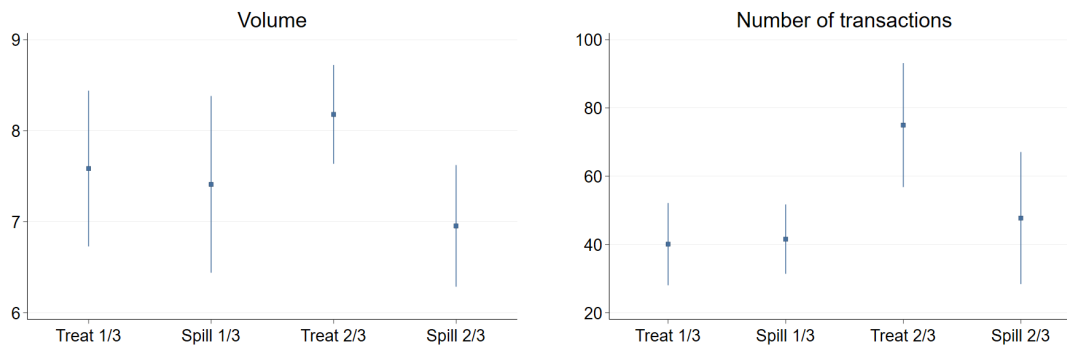
**Table A7:** Treatment effect on volume and number of transactions reported per section per week in final traceability registers

	Volume			Number of trans.		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	7.913*** (0.450)			49.082*** (5.585)		
Spill	7.224*** (0.481)			36.455*** (5.082)		
Late transaction	-3.614*** (0.207)	-3.614*** (0.207)	-2.536*** (0.189)	- 44.509*** (4.608)	- 44.509*** (4.609)	-36.451*** (4.557)
Treat 1/3		7.585*** (0.434)	8.498*** (0.517)		40.115*** (6.114)	52.803*** (12.418)
Spill 1/3		7.411*** (0.492)	8.274*** (0.531)		41.580*** (5.156)	57.065*** (8.845)
Treat 2/3		8.179*** (0.275)	7.871*** (0.311)		74.989*** (9.207)	64.385*** (9.882)
Spill 2/3		6.955*** (0.339)	6.051*** (0.519)		47.759*** (9.807)	30.946** (10.345)
Treat 1/3 × Late transaction			-2.775*** (0.625)			-31.621 (17.427)
Spill 1/3 × Late transaction			-2.682*** (0.452)			-36.817** (12.832)
Treat 2/3 × Late transaction			-0.507 (0.387)			11.636 (7.561)
Spill 2/3 × Late transaction			0.600 (0.647)			23.165** (7.709)
Cooperative FE	Yes	Yes	Yes	Yes	Yes	Yes
N	5,564	5,564	5,564	5,564	5,564	5,564
R2	0.643	0.644	0.653	0.303	0.305	0.317

Source: Authors' calculation.

**Note:** Standard errors in parenthesis, clustered at section level. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . The sample includes one observation per section per week, based on transactions in final traceability registers. The sample includes transactions from October to March. The dependent variable "Volume" corresponds to the inverse hyperbolic sine-transformation (IHS) of volume (in kilos). The "Late transaction" variable is a dummy taking the value of 1 if transactions took place in January, February, or March, and 0 if they took place in October, November, or December.

**Figure A4:** Treatment effect on volume and number of transactions reported per section per week in final traceability registers



Source: Authors' calculation.

**Note:** The figure plots the coefficients for treatment variables and interactions between treatment and transaction date from Equation 3.3 (Table A7, columns 2 and 5). The sample includes one observation per section per week, based on transactions in final traceability registers. The sample includes transactions from October to March. Dependent variables are volume in kilos using the inverse hyperbolic sine-transformation (IHS), and the number of transactions reported in final traceability registers.

**Table A8:** Treatment effect on variation in volume and number of transactions reported per section per week between first and final traceability registers

	Volume			Number of trans.		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	1.242*** (0.200)			9.836*** (2.544)		
Spill	1.320*** (0.190)			10.136*** (2.226)		
Late transaction	0.156 (0.191)	0.156 (0.191)	1.018* (0.425)	-11.321** (3.917)	-11.321** (3.918)	15.691 (8.768)
Treat 1/3		0.919*** (0.229)	0.731* (0.348)		9.726* (3.756)	20.468** (7.201)
Spill 1/3		1.505*** (0.205)	1.734*** (0.248)		10.199*** (2.768)	21.678*** (4.797)
Treat 2/3		0.582* (0.269)	0.863* (0.336)		1.089 (2.433)	-3.641 (2.357)
Spill 2/3		0.133 (0.300)	0.140 (0.396)		1.210 (2.707)	-2.811 (2.550)
Treat 1/3 × Late transaction			-0.425 (0.588)			-52.076*** (14.383)
Spill 1/3 × Late transaction			-1.396** (0.516)			-53.797*** (12.067)
Treat 2/3 × Late transaction			-1.516* (0.617)			-15.976 (9.730)
Spill 2/3 × Late transaction			-0.878 (0.677)			-17.630 (9.536)
Cooperative FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2,996	2,996	2,996	2,996	2,996	2,996
R2	0.047	0.049	0.053	0.151	0.151	0.177

Source: Authors' calculation.

**Note:** Standard errors in parenthesis, clustered at section level. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . The sample includes one observation per section per week, based on transactions in first and final traceability registers from the period covered by the first traceability register. The sample includes transactions from October to mid-January. The dependent variable "Volume" corresponds to the inverse hyperbolic sine-transformation (IHS) of volume (in kilos). The "Late transaction" variable is a dummy taking the value of 1 if transactions took place in December and January, and 0 if they took place in October and November.

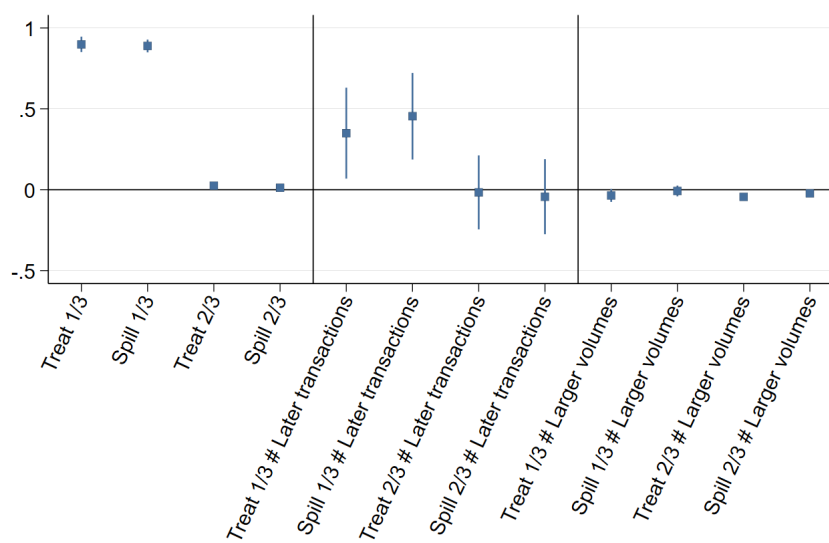
**Table A9:** Treatment effect on matched transactions between first and final traceability registers

	Perfect or imperfect match vs no match			Perfect match vs imperfect match		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	0.668*** (0.024)			0.948*** (0.014)		
Spill	0.708*** (0.022)			0.942*** (0.012)		
Late transaction	-0.066* (0.032)	-0.066* (0.032)	-0.097 (0.111)	0.055 (0.030)	0.055 (0.030)	-0.000 (0.116)
Large volume	0.030** (0.010)	0.030** (0.010)	0.098** (0.034)	0.014*** (0.004)	0.015*** (0.004)	0.038*** (0.009)
Treat 1/3		0.664*** (0.030)	0.602*** (0.040)		0.929*** (0.020)	0.899*** (0.024)
Spill 1/3		0.711*** (0.024)	0.655*** (0.029)		0.955*** (0.017)	0.889*** (0.020)
Treat 2/3		0.081*** (0.014)	0.138*** (0.015)		-0.048 (0.029)	0.025** (0.009)
Spill 2/3		0.112*** (0.024)	0.149*** (0.021)		-0.067* (0.027)	0.012 (0.009)
Treat 1/3 × Late transaction			0.185 (0.142)			0.350* (0.143)
Spill 1/3 × Late transaction			0.199 (0.131)			0.454*** (0.136)
Treat 2/3 × Late transaction			-0.090 (0.116)			-0.017 (0.116)
Spill 2/3 × Late transaction			-0.036 (0.116)			-0.043 (0.118)
Treat 1/3 × Large volume			-0.036 (0.042)			-0.036 (0.020)
Spill 1/3 × Large volume			-0.073 (0.038)			-0.008 (0.017)
Treat 2/3 × Large volume			-0.107** (0.037)			-0.044*** (0.010)
Spill 2/3 × Large volume			-0.137*** (0.039)			-0.022 (0.012)
Cooperative FE	Yes	Yes	Yes	Yes	Yes	Yes
N	128,629	128,629	128,629	68,408	68,408	68,408
R2	0.792	0.792	0.798	0.957	0.957	0.963

Source: Authors' calculation.

**Note:** Standard errors in parenthesis, clustered at section level. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . The sample includes all transactions reported in final traceability registers matching the timeframe of first traceability registers (columns 1, 2, and 3), or transactions that matched perfectly or imperfectly between the first and final traceability registers (columns 4, 5, and 6). The sample includes transactions from October to mid-January. The "Late transaction" variable is a dummy taking the value of 1 if the transaction took place at a later time than the average transaction date, i.e. after November 3, computed on the sample of transactions in the final traceability registers that fall into the period covered by the first ones. The "Large volume" variable is a dummy taking the value of 1 if the transaction volume is greater than the average volume of transactions recorded in the final traceability registers, i.e. greater than 188 kilos.

**Figure A5:** Treatment effect on transactions matching perfectly vs imperfectly between first and final traceability registers



Source: Authors' calculation.

**Note:** The figure plots the coefficients for treatment variables, interactions between treatment and transaction date, and interactions between treatment and transaction volume, from Equation 3.3 (Table A9, column 6). The sample includes transactions that matched perfectly or imperfectly between the first and final traceability registers. The sample includes transactions from October to mid-January. The "Later transactions" variable is a dummy taking the value of 1 if the transaction took place at a later time than the average transaction date, i.e. after November 3, computed on the sample of transactions in the final traceability registers that fall into the period covered by the first traceability register. The "Larger volumes" variable is a dummy taking the value of 1 if the transaction volume is greater than the average volume of transactions recorded in the final traceability registers, i.e. greater than 188 kilos.



## General Conclusion

This thesis investigates non-compliance with certification standards among various players in the cocoa value chain in Côte d'Ivoire. Using a rich and original set of first-hand data, the three chapters of the thesis aim to detect and measure opportunistic behavior among certified farmers, cocoa middlemen, and cooperatives. This work contributes to filling an important gap in the literature on certification schemes in agricultural value chains, which has so far focused on the effect of such standards on the economic benefits of participating farmers.

The first chapter investigates the reliance on child labor by certified cocoa farmers, using an indirect questioning method, the list experiment, designed to reduce the social desirability bias associated with questioning individuals about sensitive issues. I find that between 21% and 25% of surveyed farmers have used child labor to perform dangerous tasks on their cocoa farms in the previous year, a practice that is well known to be prohibited by the government and certification standards. I also highlight the relevance of using indirect questioning methods to measure child labor use in this context, as twice as many farmers report having used child labor when questioned indirectly rather than directly.

Chapter 2 aims to measure dishonesty among cocoa middlemen when faced with information asymmetries. We implement a lab-in-the-field experiment, where players take part in a "die-under-cup-task" cheating game in which they have the opportunity to cheat to maximize their winnings. Divided into 3 sessions, our experiment involves a risk of observability and penalty designed to simulate the effect of a traceability tool on the opportunistic behavior of cocoa middlemen. We find that 78% of cocoa middlemen cheated at least once during the game, taking advantage of information asymmetries to engage in opportunistic behavior that maximizes their payoffs. 59% of middlemen seized this opportunity each time it benefited them, but 22% of them never cheated. We further show that the introduction of monitoring mechanisms and sanctions in case of cheating are two effective mechanisms for reducing dishonest behavior among part of the middlemen.



Chapter 3 investigates side-sourcing prevalence upstream in the certified cocoa value chain. We rely on a rich and original data set, and compare information reported by farmers, middlemen, and cooperatives on their sales/purchases of cocoa allegedly produced on certified cocoa farms. We find large discrepancies between volumes declared by farmers and cooperatives across all cooperatives in our sample, indicating widespread reliance on side-sourcing. We implement a randomized controlled trial to assess the impact of introducing a digital traceability system (DTS) on the behavior of value chain actors. We observe strong resistance to the DTS by middlemen and cooperatives, translating into a limited use by treated cooperatives and middlemen. We further show that treated cooperatives changed their sourcing and misreporting strategies compared to control cooperatives, but overall no impact of DTS on side-sourcing.

Overall, the findings presented in these three chapters contribute to the literature on certification schemes in agricultural value chains, highlighting actors' opportunistic behavior and non-compliance with certification standards. The results provide evidence of the existence of market failures at each stage of the upstream segments of the supply chain, and suggest that mislabeling related to certification standards is an important issue.

Further research is needed to investigate traceability and mislabeling issues within certified agricultural value chains. For instance, although Chapter 1 provides evidence that certified cocoa farmers rely on child labor, it cannot be ruled out that certification schemes do have a positive (decreasing) effect on child labor. Therefore, it is essential to compare the proportion of child labor used by certified and conventional farmers, using indirect questioning methods. This would allow us to determine whether adherence to certification standards is associated with lower use of child labor when compared to conventional farmers. Additional research is also needed to determine the role of middlemen in side-sourcing practices. Chapter 2 only provides insights into their behavior in situations of information asymmetry, from which they can take advantage to increase their profits, and Chapter 3 leaves room for further investigation of their participation in side-sourcing.





# Detailed Table of Contents

<b>General Introduction</b>	1
References . . . . .	11
<b>Chapter 1.</b> _____	
Addressing Social Desirability Bias When Measuring Child Labor Use: An Application to Cocoa Farms in Côte d'Ivoire	13
1. Introduction . . . . .	14
2. Child labor in the Ivorian cocoa sector . . . . .	18
2.1 Cocoa sustainability standards and child labor . . . . .	18
2.2 Social desirability bias among certified cocoa farmers . . . . .	20
2.3 The list experiment . . . . .	22
3. Research design . . . . .	23
3.1 The data . . . . .	23
3.2 Design and implementation of the list experiment . . . . .	24
3.3 Estimation strategy . . . . .	28
4. Results . . . . .	31
4.1 Differences between direct questioning and LE responses . . . . .	31
4.2 Robustness checks . . . . .	33
4.2.1 Potential source of measurement errors . . . . .	33
4.2.2 Sensitivity of LE estimates to design changes . . . . .	35
4.2.3 List experiment and cognitive load . . . . .	39
5. Conclusions . . . . .	39
References . . . . .	42
Appendix . . . . .	46
<b>Chapter 2.</b> _____	
Information Asymmetries, Risk and Dishonesty: A Behavioral Experiment Among Middlemen in Certified Cocoa Value Chain	53
1. Introduction . . . . .	54
2. Context . . . . .	58
2.1 Traceability in the certified cocoa supply chain . . . . .	58
2.2 Detecting dishonest behavior . . . . .	59
3. Experimental design . . . . .	60
3.1 Protocol . . . . .	60
3.2 Field implementation . . . . .	63

4.	Pre-lab quantitative survey . . . . .	64
5.	Cheating behavior of cocoa middlemen . . . . .	67
5.1	Cheating rates . . . . .	67
5.2	Cheater profiles . . . . .	69
6.	Effect of observability and sanction among occasional cheaters . . . . .	72
6.1	Empirical specification . . . . .	73
6.2	Findings . . . . .	74
7.	Conclusion . . . . .	77
	References . . . . .	80
	Appendix . . . . .	83

**Chapter 3.**

---

	Traceability Along the Value Chain: the Case of Certified Cocoa in Côte d'Ivoire	105
1.	Introduction . . . . .	106
2.	Context . . . . .	109
2.1	Cocoa value chain in Côte d'Ivoire . . . . .	110
2.2	Conventional vs ESS-certified cocoa . . . . .	111
2.3	Current traceability system . . . . .	112
2.4	A new digital traceability application . . . . .	113
3.	Data . . . . .	114
4.	Side-sourcing in ESS-certified cocoa value chains . . . . .	117
4.1	Reporting discrepancies across farmers, middlemen, and cooperatives . . . . .	117
4.2	Who engages in side-sourcing? . . . . .	118
4.3	Strategic misreporting . . . . .	122
5.	Assessing the impact of a new digital traceability system: a randomized controlled trial	126
5.1	Design . . . . .	127
5.2	Experimental integrity . . . . .	129
5.3	Estimation framework . . . . .	133
6.	Support vs resistance to traceability . . . . .	134
6.1	Initial resistance to DTS . . . . .	134
6.2	Use of the DTS . . . . .	135
7.	Impact of traceability system . . . . .	140
7.1	Impact on purchases . . . . .	141
7.2	Impact on reported volumes and transactions at cooperative-level . . . . .	142
7.3	Impact on data manipulation . . . . .	143
8.	Conclusion . . . . .	147
	References . . . . .	150
	Appendix . . . . .	153

<b>General Conclusion</b>	171
Detailed Table of Contents	177
List of Tables and Figures	179



## List of Tables and Figures

<b>Chapter 1</b>		<b>13</b>
Figure 1.	Sequencing of the list experiments and direct questions . . . . .	25
Table 1.	List experiments and direct questions . . . . .	26
Table 2.	Basic characteristics of respondents in groups A and B . . . . .	29
Table 3.	No-design effect test results . . . . .	30
Table 4.	Distribution of farmers according to their answer to the list experiments . . . . .	31
Table 5.	Proportion of farmers using child labor estimated by list experiment and direct questioning method . . . . .	32
Table 6.	Survey characteristics across waves . . . . .	36
Table 7.	Regression analysis of list experiments (LEs) across survey waves . . . . .	37
Table 8.	Proportion of farmers who passed the list experiment comprehension test . . . . .	40
Table A1.	Descriptive statistics of basic farmers characteristics . . . . .	46
Table A2.	Basic characteristics of respondents across groups - 2022-V1 . . . . .	47
Table A3.	Basic characteristics of respondents across groups - 2022-V2 . . . . .	47
Table A4.	Farmers' use of labor (using direct questions) . . . . .	47
Table A5.	List experiments in 2022-V1 . . . . .	48
Table A6.	List experiments in 2022-V2 . . . . .	49
Table A7.	Distribution of farmers according to their answer to the list experiment - "harvesting and pod-breaking" . . . . .	50
Table A8.	Distribution of farmers according to their answer to the list experiment - "farm preparation" . . . . .	50



Table A9.	Distribution of farmers according to their answer to the list experiment - “employed child labor” . . . . .	50
Table A10.	Wording of the list experiments tests and proposed items . . . . .	51
<b>Chapter 2</b>		<b>53</b>
Figure 1.	Structure of the experimental game . . . . .	61
Table 1.	The five paired lottery-choice decisions . . . . .	62
Figure 2.	Distribution of true die roll outcomes and outcomes declared by players for Session 1 . . . . .	68
Figure 3.	Distribution of individual cheating rates . . . . .	70
Table 2.	Individual determinants of cheating . . . . .	72
Table 3.	Effect of history loss, previous cheating, and game features on cheating . . . . .	75
Figure A1.	Lab-in-the-field experiment setting and tools . . . . .	83
Figure A2.	Survey areas . . . . .	84
Table B1.	Balancing test comparing individuals who accepted vs. those who declined participation in the game . . . . .	92
Table C1.	Socio-economic characteristics of traders . . . . .	93
Table C2.	Characteristics of cocoa traders’ activities . . . . .	94
Table C3.	Traders’ relationship with farmers . . . . .	95
Figure D1.	Distribution of true die roll outcomes and outcomes declared by players (Session 2) . . . . .	96
Figure D2.	Distribution of true die roll outcomes and outcomes declared by players (Session 3) . . . . .	97
Table D1.	Effect of the probability of losing on unconditioned cheating, using linear probability model (LPM) . . . . .	97
Table D2.	Payoffs distribution according to players’ cheating profile . . . . .	98

Table D3.	Differences in payment modes with farmers by players' cheating profiles . . . .	99
Table E1.	Linear Probability Model (LPM) on the subsample of occasional cheaters . . . .	101
Table E2.	Linear Probability Model (LPM) on overall sample . . . . .	102
<b>Chapter 3</b>		<b>105</b>
Table 1.	Descriptive statistics of cocoa value chain actors . . . . .	110
Figure 1.	Products, services, and information flows in cocoa value chains . . . . .	114
Table 2.	Cocoa transactions: sources of information . . . . .	115
Figure 2.	Weekly reporting differences between farmers, middlemen, and cooperatives	118
Table 3.	Farmers' average theoretical and observed yields during the first part of the main cocoa season (October-December) . . . . .	119
Figure 3.	Volume reported by cooperatives for surveyed farmers during the recall period (as a % of volume reported by farmers) . . . . .	121
Figure 4.	Correlation between differences in volume declared by cooperatives compared with farmers and 2020-2021 MAD . . . . .	122
Table 4.	Descriptive characteristics of sections across years, for the main cocoa season .	124
Figure 5.	Proportion of sections for which a key indicator is higher in cooperatives' traceability registers than in farmers' reporting . . . . .	125
Table 5.	Differences in key indicators between first and final traceability registers of cooperatives during the 2022-2023 cocoa season . . . . .	126
Figure 6.	Control and treatment groups at the cooperative and section level . . . . .	129
Table 6.	Balancing tests . . . . .	131
Table 7.	Implementation, equipment distribution, and compliance with the DTS use according to treatment assignment . . . . .	132
Table 8.	Use of the DTS across cooperatives and sections during the 2022-2023 main cocoa season . . . . .	136

Figure 7.	Proportion of volume from treated sections reported in the DTS, at cooperative-level . . . . .	137
Figure 8.	Correlation between starting date of the traceability application and level of digitization, at section-level . . . . .	138
Figure 9.	Daily volume of cocoa digitized in the traceability application by treated sections	139
Figure 10.	Treatment effects on transaction digitization . . . . .	140
Figure 11.	Treatment effects on volume and number of transactions reported per surveyed farmers per week . . . . .	141
Figure 12.	Treatment effects on volume and number of transactions reported in first traceability registers . . . . .	142
Figure 13.	Treatment effects on the difference in volume and number of transactions between first and final traceability registers . . . . .	144
Figure 14.	Treatment effect on transaction matching between first and final traceability registers . . . . .	146
Table A1.	Basic characteristics of surveyed farmers across groups . . . . .	154
Table A2.	Basic characteristics of surveyed traders across groups . . . . .	156
Figure A1.	Geographical distribution of cooperatives and volume reporting surplus (as a % of volume reported by farmers) . . . . .	157
Figure A2.	Proportion of digitized volume across sections, by cooperative . . . . .	158
Table A3.	Treatment effect on transaction digitization . . . . .	160
Table A4.	Treatment effect on volume and number of transactions reported per surveyed farmers per week . . . . .	161
Table A5.	Treatment effect on volume and number of transactions reported per farmers per week by middlemen . . . . .	162
Figure A3.	Treatment effects on volume and number of transactions reported per surveyed farmers per week by middlemen . . . . .	163

Table A6.	Treatment effect on volume and number of transactions reported per section per week in first traceability registers . . . . .	164
Table A7.	Treatment effect on volume and number of transactions reported per section per week in final traceability registers . . . . .	165
Figure A4.	Treatment effect on volume and number of transactions reported per section per week in final traceability registers . . . . .	166
Table A8.	Treatment effect on variation in volume and number of transactions reported per section per week between first and final traceability registers . . . . .	167
Table A9.	Treatment effect on matched transactions between first and final traceability registers . . . . .	168
Figure A5.	Treatment effect on transactions matching perfectly vs imperfectly between first and final traceability registers . . . . .	169



