THÈSE de DOCTORAT

de l'UNIVERSITÉ TOULOUSE CAPITOLE

Présentée et soutenue par

Madame Johanna OBST

Le 19 février 2024

Trois essais sur l'économie de la certification volontaire

École doctorale : **TSE Toulouse Sciences Economiques**

Spécialité : **Sciences Economiques - Toulouse**

Unité de recherche : **TSE-R - Toulouse School of Economics - Recherche**

Thèse dirigée par *Madame Emmanuelle AURIOL et Monsieur Stéphane STRAUB*

Composition du jury

Rapporteure : Mme Julie SUBERVIE
Rapporteur : M. Eduardo SOUZA-RODRIGUES
Directrice de thèse : Mme Emmanuelle AURIOL
Président : M. Mathias REYNAERT

Three essays on the economics of voluntary certification

Johanna Isman

Thesis submitted to

Toulouse School of Economics, University Toulouse I Capitole

in partial fulfilment for the award of the degree of

DOCTOR OF PHILOSOPHY

in Economics

December 2023

Acknowledgements

This doctoral thesis would not have been possible without the support of faculty members and staff at the Toulouse School of Economics and my friends and family. I owe a significant amount of gratitude to my advisors, Emmanuelle Auriol and Stéphane Straub. Their encouragement and help throughout the last five years have been tremendously influential in completing this work. I am also profoundly thankful to Mathias Reynaert. Regular exchanges with him helped me overcome numerous challenges, particularly in the work on the first chapter of this thesis. I further thank Isis Durrmeyer and Eduardo Souza-Rodrigues for many conversations that helped me progress at the beginning and end of the work on this thesis, respectively. Moreover, I am very grateful to Sara Shahanaghi, who mentored me as part of the Women in Economics Association. For their helpful comments, I also thank Jens Abildtrup, Victor Aguirregabiria, Stefan Ambec, Facundo Argañaraz, Wesley Blundell, Nicolas Bonneton, Zohra Bouamra-Mechemache, Sylvain Chabe-Ferret, Manon Desjardins, Pierre Dubois, Raymond Fisman, Ana Gazmuri, Olivier de Groote, Christian Hellwig, Koen Jochmans, Yao Luo, Josepa Miquel-Florensa, Anouch Missiran, Katrina Mullan, Benjamin Ouvrard, Harry Paarsch, Jérome Renault, François Salanié, Anna Sanktjohanser, and many seminar and conference participants.

Throughout the work on this thesis, my family and friends have been there to accompany me through the ups and downs. Most of all, I am deeply grateful to my husband and parents for believing in me and supporting me in many ways. Thanks to the love and example of my parents, I developed the perseverance and confidence needed to complete this thesis. Thanks to my husband, I have not given up even when my perseverance and confidence were seriously challenged. Among the other friends and family members who regularly shared an open ear, uplifting words, and advice are Luise Eisfeld, Jonas Gathen, Moritz Hengel, Claire Ivins, Oscar Jara, Alina Ozhegova, Hanna Zinner, as well as my siblings and their partners.

Finally, I thank numerous stakeholders of certification schemes for providing valuable insight through interviews and an online survey. I am grateful to the IT team at TSE and my friends who helped with the technical implementation and translation of the survey conducted for the first chapter of the thesis. This work has benefitted from a State grant managed by the National Research Agency under the "Investissements d'Avenir" program (ANR-18-EURE-0021).

Abstract

English version

This thesis studies the economics of voluntary certification of quality that consumers and investors do not observe, so-called credence attributes. The first two chapters take an empirical approach, focussing on the certification of sustainable forest management in wood production, which aims at reducing negative externalities such as carbon emissions and biodiversity loss.

The first chapter investigates how stricter enforcement of certification rules affects forest plots' participation and quality, here sustainability, when forest managers choose not only the standard and label but also one of competing certifiers. Those companies audit quality for the same label with potentially varying levels of rigor. Label owners enforce their standards by excluding excessively lenient certifiers. I build a structural model of such voluntary certification and estimate it with novel web-scraped and survey data on the Forest Stewardship Council's (FSC) standard for sustainable wood production. I find considerable differences in certifiers' levels of rigor and show that forest managers are willing to pay substantially more for relatively lenient certifiers. Counterfactuals reveal that increasing certifiers' minimum level of rigor raises quality in certified forests but reduces participation. That leads to a hump-shaped relationship between minimum rigor and aggregate quality in certified and uncertified forests. The results highlight a general limitation of voluntary certification. However, they also show FSC's scope to incentivize further reduction of negative externalities.

In the second chapter, my co-authors and I investigate if labels such as FSC's can motivate companies to reduce negative externalities even by restricting production volume. We study the impact of modifications made to the FSC standard in 2014 that attempted that. Specifically, we analyze the effects of the requirement to preserve a minimum of 80% of forested areas in their natural state (called intact forest landscape or IFL) within certified forests. We study the effects on certification decisions and the conservation of IFL. To examine those, we link geographic information on forest concessions with remote sensing of the forest area, details about FSC certification, and audits in countries characterized by substantial areas of IFL. We use a difference-in-differences framework. Our findings reveal that in Russia, after the modification in the FSC standard, concessions with IFL were less inclined to get or remain FSC-certified compared to those lacking IFL. Additionally, we find indications that

the change contributed to the conservation of IFL in certified forests. However, the drivers behind these improvements remain elusive, as our study does not uncover reductions in tree cover loss within the IFL.

The third chapter of the thesis uses game theory to study accreditation, i.e., the auditing and licensing of certifiers. Many countries and label owners establish accreditation bodies for that purpose. I analyze how the establishment and regulation of accreditation bodies affect certifiers' incentives for fraud and welfare. I provide motivating evidence based on increased certification after the international recognition of the Uruguayan accreditation body. I then analyze public-perfect equilibria in an infinitely repeated game of a mass of buyers, a monopolist supplier, a monopolist auditing company, and one or more accreditation bodies. I show that a necessary condition for a welfare-improving effect of accreditation bodies' existence is that buyers are sufficiently sophisticated or that accreditation is compulsory. The model highlights that accreditation bodies should not be profit-maximizing companies.

French version

Cette thèse étudie l'économie de la certification volontaire de la qualité non-observés par les consommateurs et les investisseurs. Les deux premiers chapitres se concentrent sur la gestion durable des forêts, qui peut réduire les externalités négatives telles que la perte du stock de carbone et de la biodiversité.

Le premier chapitre étudie comment une application plus stricte des règles de certification affecte la participation des parcelles forestières et leur qualité lorsque les gestionnaires forestiers choisissent non seulement le label, mais aussi l'un des certificateurs concurrents. Ces entreprises contrôlent la qualité pour le même label avec des niveaux de rigueur potentiellement variables. Les propriétaires de labels font respecter leurs normes en excluant les certificateurs trop indulgents. Je construis et estime un modèle structurel de ce type de certification à l'aide de nouvelles données de panel sur la norme du Forest Stewardship Council (FSC). Je quantifie des différences considérables dans les niveaux de rigueur des certificateurs et montre que les gestionnaires forestiers sont prêts à payer beaucoup plus pour des certificateurs relativement indulgents. Des analyses contrefactuels révèlent que l'augmentation du niveau minimum de rigueur des certificateurs améliore la qualité des forêts certifiées mais réduit la participation. Il en résulte une relation en forme de bosse entre la rigueur minimale et la qualité globale des forêts. Les résultats montrent une limitation générale de la certification volontaire, mais aussi que FSC peut inciter à réduire davantage les externalités négatives.

Dans le deuxième chapitre, mes co-auteurs et moi cherchons à savoir si des labels peuvent inciter des entreprises à réduire les externalités négatives même en limitant le volume de production. Plus précisément, nous étudions l'impact des modifications de la norme FSC en 2014, qui exigent la préservation d'un minimum de 80% de paysages forestiers intacts et non perturbés (IFL) dans les forêts certifiées. Nous analysons les effets sur leur taux de certification et la conservation des IFLs. Pour ce faire, nous relions les informations géographiques sur les concessions forestières à la télédétection de la zone forestière, aux détails de la certification FSC dans des pays avec beaucoup de IFLs. A base de la méthode des doubles différences, nos résultats révèlent qu'en Russie, après la modification de la norme FSC, les concessions avec IFL étaient moins enclines à obtenir ou à rester certifiées FSC que celles qui n'avaient pas d'IFL. En outre, nous trouvons des indications que le changement a contribué à la conservation de l'IFL dans les forêts certifiés. Cependant, les facteurs à l'origine de ces améliorations restent insaisissables, car notre étude ne révèle pas de réduction de la perte de couverture arborée au sein de l'IFL.

Le troisième chapitre de la thèse utilise la théorie des jeux pour étudier l'accréditation, c'est-à-dire l'audit et l'octroi de licences aux certificateurs. De nombreux pays et propriétaires de labels mettent en place des organismes d'accréditation (OA) à cette fin. J'analyse comment la création et la réglementation des OAs affectent les incitations des certificateurs à la fraude et le bien-être économique. Je quantifie des faits motivantes basées sur l'augmentation de la certification après la reconnaissance internationale de l'OA uruguayen. J'analyse ensuite les équilibres publics parfaits dans un jeu répété à l'infini entre une masse d'acheteurs, un fournisseur monopoliste, une société d'audit monopoliste et un ou plusieurs OAs. Je montre qu'une condition nécessaire pour que l'existence des OAs ait un effet d'amélioration du bien-être est que les acheteurs soient suffisamment sophistiqués ou que l'accréditation soit obligatoire. Le modèle prouve que les OAs ne doivent pas être des entreprises maximisant leur profit.

Contents

Chapter 1

Auditor Leniency and Participation in Voluntary Forest Certification

Johanna Isman[1,2]

[1]Toulouse School of Economics

[2]This work has benefitted from support by Emmanuelle Auriol's research grant at the Institut Universitaire de France and data collection assistance by Jia Wang, research assistant at Assurance Services International.

1.1 Introduction

Millions of firms seek third-party certification of voluntary standards to signal unobserved quality of their products or services (Dranove and Jin, 2010).[3] Examples range from credit risk and seller quality to organic agriculture and sustainable forestry. Certification mitigates information asymmetries, particularly in the case of *credence attributes*, which consumers and investors cannot observe at a reasonable cost, not even after purchase (Darby and Karni, 1973). Furthermore, certification of sustainability standards can help reduce negative production externalities (Bizzotto and Harstad, 2020), such as the loss of biodiversity and carbon stocks. Yet, well-known non-profit institutions criticize lax enforcement, and several empirical studies find small or no significant differences in the sustainability of certified and uncertified production.[4]

How would stricter enforcement in voluntary certification affect quality in the industry and welfare? Understanding and predicting those effects is not only crucial for private and public institutions that own certification labels.[5] It also helps policy-makers understand the potential and limitations of voluntary certification in regulating externalities.[6] From a theoretical perspective, the impact of stricter enforcement is ambiguous. While it could improve sustainability or other qualities of certified production, it may also reduce participation in certification and, thus, the quality of dropouts (Hui et al., 2023).

This paper studies that empirical trade-off by building and estimating a structural model of voluntary certification in which firms choose between competing certifiers. The model captures the way voluntary certification of most credence qualities is organized. Various certifiers audit forests' quality according to the same standard but potentially with varying levels of rigor. These companies provide the same label in the eyes of consumers and investors. Strategically acting, profit-maximizing firms may, thus, shop for more lenient certifiers (Bar and Zheng, 2019). To counterbalance, label owners appoint an accreditation body that

[3]Quality management system certification according to the ISO 9001 standard alone covers more than a million companies (ISO, 2023).

[4]Examples of such empirical studies are Blackman et al. (2018), Villalobos et al. (2018), Rico-Straffon et al. (2023). See Blackman and Rivera (2011) and DeFries et al. (2017) for reviews of various sectors. Greenpeace International (2021) and the International Consortium of Investigative Journalists (Alecci, 2023) denounce a lack of rigor in certifying land-based commodities.

[5]Voluntary standards set by public authorities include the EU rules for organic certification and the German "Green Button" for sustainable clothing production (European Commission, 2018; Federal Ministry for Economic Cooperation and Development, 2023).

[6]For example, EU institutions discussed whether to recognize certification as proof of compliance when drafting the EU deforestation regulation (European Union, 2023).

inspects and licenses certifiers. To increase the minimum level of rigor, the accreditation body suspends the market access of excessively lenient certifiers, which has side effects on firms' participation and welfare.

I quantify certifiers' rigor differences, firms' preferences for leniency, and effects of counterfactual certifier suspensions with novel data on Forest Stewardship Council (FSC) certification. FSC owns one of the most widespread sustainability standards, covering roughly one-tenth of the global wood production (FAO, 2018). The FSC standard is the most demanding and transparent standard for sustainable forest management.[7] Nevertheless, some FSC certifiers have been accused of leniency, even regarding cases of illegal logging.[8]

This paper contributes to the literature on certification and enforcement in three ways. First, it provides new insights about forest conservation by assessing reasons for the limited effectiveness of certification and the extend to which FSC can improve in view of certifiers' rigor and firms' preferences. Second, the paper extends the broader literature on certifiers' incentives and oversight. While, for example, Duflo et al. (2013a) highlight auditors' incentives problems and the positive effects of stricter control for mandatory standards, I quantify the more ambiguous effects in voluntary certification. Finally, the present paper complements Hui et al. (2023), who study tightened thresholds in the certification of experience quality, which consumers observe and review after purchase. I focus on credence quality, whose certification requires field audits, and I account for the widespread separation of label owners and certifiers in that setting.

To address the research question, I collected novel panel data on FSC-certified forests. I scraped data on all forests that were FSC-certified for at least one year between 2015 and 2019, henceforth referred to as *once-certified* forests. Most importantly, the panel includes yearly observations of participation, the chosen certifier, and the number of violations reported by the certifier during the annual audit. These reported violations of the FSC standard matter since forest firms must correct them to become or remain certified. I extracted these variables from the audit reports published on FSC's website. For the analysis, I focus on various world regions that jointly represent one third of all FSC certificates and provide over 6,000 observations. These regions include the Americas and large parts of Asia. I also conducted

[7]Jurists, forest scientists, and NGOs share this view (Clark and Kozar, 2011; Greenpeace International, 2021; Gutierrez Garzon et al., 2020; Ludwig et al., 2014).

[8]See Conniff (2018); Earthsight (2020). FSC consequently asked their accreditation body to "perform a check of the performance of the certification body" (FSC, 2021a), which suggests that they recognize the role certifiers' ("certification bodies"') rigor plays.

an online survey in 24 languages asking forest managers about the certification fees they pay. The responses provide over 380 certification fee quotes from the regions the paper focuses on, joined with information about the corresponding certificates. I use those data to predict certification fees for all certifiers and forests.

Next, I model voluntary certification as a 3-stage game, played each year. Each world region and year jointly define a market. In the first stage, certifiers set markups as market-level percentages of their marginal cost. In the second stage, forest firms choose whether to become FSC-certified and, if yes, their most preferred certifier. In the third stage, the chosen certifiers audit the forests and detect and report a fraction of forest firms' actual violations of the FSC standard. Correcting those violations is a cost to forest firms' surplus. The reported fraction of violations is a measure of certifiers' audit rigor. Forest firms anticipate certifiers' rigor when choosing a certifier. Conditional on a rich set of audit and forest characteristics, I model both certifiers' level of rigor and the number of firms' actual violations as exogenous variables rather than strategic choices. Certifiers' market-level rigor *differences* depend on their constant rigor types, the distances to their headquarters and inspections by the accreditation body. While the types account for adverse selection, the variation with distances and inspections account for potential moral hazard or behavioral effects in a parametric way.[9]

I estimate the model by backward induction. The abovementioned modeling choices allow me to identify certifiers' rigor differences, i.e., their *relative rigor* compared to a baseline. That is necessary since I observe neither certifiers' level of rigor nor forest firms' actual violations of the FSC standard directly. To obtain reliable estimates of relative rigor from violation reports, I need to account for the potential selection of more or less compliant forest plots into more or less rigorous certifiers. A rich set of audit and forest-level predictors, selected by LASSO regularization, does that to a substantial extent. In addition, I build a control function based on a distributional assumption about the unobserved factors affecting selection and compliance. I combine the approach of Lee (1983) for selection among multiple alternatives in a linear model with the approach of Terza (1998) for binary selection in nonlinear models. To my knowledge, I am the first to do so. The model's estimates suggest that certifiers differ significantly in terms of their relative rigor levels. All else being equal, the least rigorous certifier reports only around one third of the violations that the most rigorous certifier reports. An inspection by the accreditation body is associated with doubling the number of violation

[9]In Appendix 1.10.1.5, I discuss to which extent the model nests the possibility of strategic long-term choices of rigor.

reports by increasing the inspected certifier's relative rigor. Variation in the likelihood of inspections creates within-certifier variation in the relative rigor levels which forest expect for each certifier and market.

In the second stage, I exploit that cross-market rigor variation to identify firms' preference for less rigorous certifiers. To do so, I use demand estimation techniques from the empirical industrial organization literature (Berry et al., 1995; McFadden, 1977). I account for choice variation along various certifier characteristics and unobservable popularity captured by certifier fixed effects. Nevertheless, both prices and expected relative rigor may correlate with unobserved shocks to firms' demand for certifiers. For rigor, the reason is that the accreditation body reacts to changes in market shares by adapting the frequency of accreditation inspections which affects certifiers' rigor. To account for those sources of endogeneity, I use differentiation instruments that capture the degree of market isolation of certifiers in terms of their characteristics (Gandhi and Houde, 2019). On average, the estimates suggest that forest firms are willing to pay more than 4,000 USD more for a certifier that reports one standard deviation of violations less than their competitor. That is substantial, constituting approximately 40% of the average certification fee and half of firms' average estimated net benefit from FSC certification. Considerable variation in the timing of forest firms' decisions to join FSC or to drop out, coupled with variation in market structures and certifiers' characteristics, helps identifying those net benefits. Firms' estimated willingness to pay to stay with their previous certifier is of similar magnitude as the willingness to pay for leniency, capturing large switching costs that limit shopping for leniency.

Using the estimated model, I conduct two sets of counterfactual exercises. First, I simulate 10 to 500% increases in the minimum level of expected relative rigor worldwide, henceforth referred to as minimum rigor. That is, I shift the lower rigor levels of any certifier and market up to that new level. Such a change mechanically increases the quality of certification. I measure quality as the expected number of violation reports, since more violation reports lead to the correction of more violations and, thus, more compliance. I solve for forest firms' new choice probabilities and certifiers' new prices, keeping everything else constant. The results suggest that the effect on quality among all once-certified forest plots is positive up to increases of approximately 250%. Overall, the relationship between minimum rigor and quality exhibits a hump-shape.

In the second set of counterfactuals, I start by considering the enforcement of increases in minimum rigor. Since many voluntary standards, including FSC, work with multiple certifiers, label owners and their accreditation bodies cannot directly shift minimum rigor. Instead, they do so by suspending excessively lenient certifiers' accreditation and thus their market access. I simulate the accreditation suspensions of the most lenient, the two most lenient and the three most lenient certifiers, by removing them from the choice sets. The trade-off between the quality of certification and participation remains. The overall effects of all three suspension examples on aggregate quality across certified and uncertified forests and welfare are positive. For example, I predict that the suspension of the most lenient certifier leads to a 2.5% increase in aggregate quality, despite almost 6% reduction in participation. I attempt to approximate the potential social benefits and welfare based on survey results on consumers' stated valuation of FSC certification. Despite the reduction in the availability of certified wood and a cost of almost one million USD for forest firms and certifiers, I predict that the suspension of the most lenient certifier could raise welfare by approximately 23 million USD.

Finally, I compare the effects of suspensions to the effects of equivalent increases in minimum rigor for each market. For each set of suspended certifiers, I implement an alternative scenario where I do not remove that certifier but rather shift their rigor to the next most rigorous certifier's rigor in the same market. The set of available rigor levels in each market is then the same as that in the case of a suspension. I find that such equivalent increases in minimum rigor explain less than half of the changes due to suspending the most lenient certifier. Suspensions lead to larger drops in participation but also larger increases in the quality of certification than equivalent increases in minimum rigor. Switching costs and preferences for other certifier characteristics can explain that. Suspension not only changes the available set of rigor levels but also forces forest firms to transfer to another certifier. Transfers are costly, so that more forest firms drop out instead of transfer, when faced with such a situation. At the same time, in the case of a suspension, more forest firms transfer to a certifier that is more rigorous than to the next most lenient certifier. Overall, whether suspensions have better or worse quality and welfare effects than equivalent rigor shifts depends on the characteristics and market shares of the targeted certifier.

Contribution to the literature

The present paper extends a rich literature that investigates reasons for forest degradation and deforestation (Assunção et al., 2023; Balboni et al., 2021; Burgess et al., 2012) and their reduction through private and public policies (Alix-Garcia and Wolff, 2014; Assunção et al., 2023; Simonet et al., 2019; Sims and Alix-Garcia, 2017; Souza-Rodrigues, 2019). In particular, this work complements studies of the effectiveness of forest management certification in preventing forest degradation or deforestation, which certification can do in two ways. First, it can compensate firms' conservation efforts by involving only those forest plots that were already more sustainably managed than average ones before certification. While Goodman et al. (2019) and Kalonga et al. (2016) document such contexts, other studies find no ex-ante difference or even more tree cover loss in later-certified forests (Blackman et al., 2018; Rico-Straffon et al., 2023). Second, certifiers can make forest management more sustainable, for example by reporting violations of the standard and requiring firms to correct them to maintain their certification. Several papers study such causal effects by comparing changes in ex-ante similar forests. Most of those papers find no significant effects (Blackman et al., 2018; Panlasigui et al., 2018; Rico-Straffon et al., 2023; Villalobos et al., 2018), with some exceptions (Miteva et al., 2015; Tritsch et al., 2020). The present paper investigates two potential reasons for the limited effectiveness of FSC and similar certification schemes in some contexts[10] by analyzing the incentives of certifiers and the impact of the voluntary nature of those certification schemes. This work goes beyond existing studies by predicting to what extent stricter accreditation may improve the current level of forest conservation.

More generally, the present paper addresses the economic incentives of inspectors and certifiers. Much of this stream of literature is theoretical (Auriol and Schilizzi, 2015; Bizzotto and Harstad, 2020; Bolton et al., 2012; Mathis et al., 2009; Stahl and Strausz, 2017). Empirically, firms' shopping for less rigorous auditors has been documented for the certification of mandatory environmental or product safety regulations (Chu et al., 2021; Duflo et al., 2013a,b, 2018; Hubbard, 2002) and credit ratings (Becker and Milbourn, 2011; Jiang et al., 2012). To my knowledge, the only quantitative analysis of firms' preference for leniency in voluntary certification is Bar and Zheng (2019), focusing on food safety certification. I identify differences in leniency and firms' willingness to pay for leniency in *voluntary sustainability certification*. There are many reasons to expect important differences in the degree to which firms in this context shop for leniency compared to the settings studied in other papers.[11] More important than such differences are the normative implications. Where

[10]One caveat is that most of those rigorous, quantitative studies look at tree cover loss, whereas FSC may be more effective in other dimensions that are more difficult to measure at a large scale.

[11]Ex-ante, it is not clear if shopping for leniency is exacerbated or mitigated in the context of voluntary certification. On the one hand, punishment for leniency can be more severe in mandatory certification. Credit rating agencies also face more considerable risks of losing their reputation as (i) their brand is more observable to clients than are the brands of certifiers working for the same label as FSC, and (ii) clients are likely to more appropriately judge differences in their overall accuracy ex-post. On the other hand, intrinsic motivation may be more present in voluntary certification than in mandatory certification and credit ratings. Food safety differs from environmental protection in that the former is more of an experience than a credence attribute,

participation is mandatory, counteracting shopping for leniency by punishing lenient auditors or assigning them randomly typically has positive effects (Duflo et al., 2013a). In voluntary certification, the outcome is less clear, as firms may stop participating when auditors become too rigorous. I complement Bar and Zheng (2019) by investigating this trade-off. I estimate a structural model that allows me to simulate the quality and welfare effects of suspending lenient certifiers.

Finally, the present paper joins Hui et al. (2023) in empirically investigating the trade-off between stringency and participation in voluntary certification. Hui et al. (2023) exploit rich data and a change in eBay's seller certification to analyze the effects of increased stringency on the distribution of quality and other market outcomes. My data and context do not allow me to estimate such a rich model incorporating moral hazard. However, I complement that paper by focusing on a certification standard that is very different from eBay's seller certification but representative of many voluntary schemes, particularly those addressing ethical and environmental issues. Certification on internet platforms such as eBay typically reveals experience attributes, observed after purchase. Platforms can certify easily by summarizing the reviews of previous customers. I investigate the trade-off between stringency and participation in the voluntary certification of credence attributes, which is not even observable after purchase. The certification of such quality attributes often requires extensive field audits, which has led label owners to outsource certification to multiple certifiers. As my paper shows, this structure has significant consequences for the enforcement of increased stringency and consequent effects on quality. Moreover, I analyze the effects with a view towards externalities, which matters for interpreting welfare effects.

Overview

The remainder of this paper is structured as follows. Section 1.2 describes the institutional setting. Section 1.3 presents the data. Section 1.4 outlines the model of violation reporting, demand for certifiers, and pricing. Section 1.5 details the estimation and presents and discusses the results. Section 1.6 covers the counterfactual analysis. Section 1.7 concludes the paper.

---

similar to credit ratings.

1.2 Institutional setting

This section describes the key features of the FSC certification system. Many of these features are typical for most voluntary certification schemes of credence qualities and motivate the model's focus. Other features are more specific and will help interpret the quantitative results of the estimated model.

1.2.1 Forest management units

Approximately 10% of the global forest area is certified according to sustainable management standards (UNECE/FAO, 2019). This paper analyzes certification decisions over time at the forest management unit (FMU) level. FMUs are equivalent to establishments in other industries. They are defined forest areas managed by the same firm according to a joint plan (FSC, 2017).[12] A certificate can cover an individual FMU or a group of FMUs, jointly responsible for compliance. I refer to the whole entity covered by one certificate as an FMU but account for jointly certified groups of FMUs in the analysis.

FMUs seeking certification can choose between two main, globally recognized sets of standards: the Forest Stewardship Council (FSC) standard and a wide range of other standards recognized by the Programme for the Endorsement of Forest Certification (PEFC). This paper focuses on FSC for two reasons. First, only FSC publishes audit results, which are indispensable for empirical analysis. Second, the FSC standard is the most demanding standard, also compared to national regulations.[13] The analysis thus considers all other standards as less demanding outside options. FSC certification covers roughly one tenth of the global wood production and approximately 4% of the world's forest area, equivalent to the size of Mexico.[14]

1.2.2 The FSC certification standard

Figure 1.1 illustrates the FSC certification scheme, which is organized similarly to most other voluntary certification schemes.[15] After choosing the FSC standard, the FMU chooses among certifiers approved by FSC's accreditation body. The chosen certifier decides on the FMU's

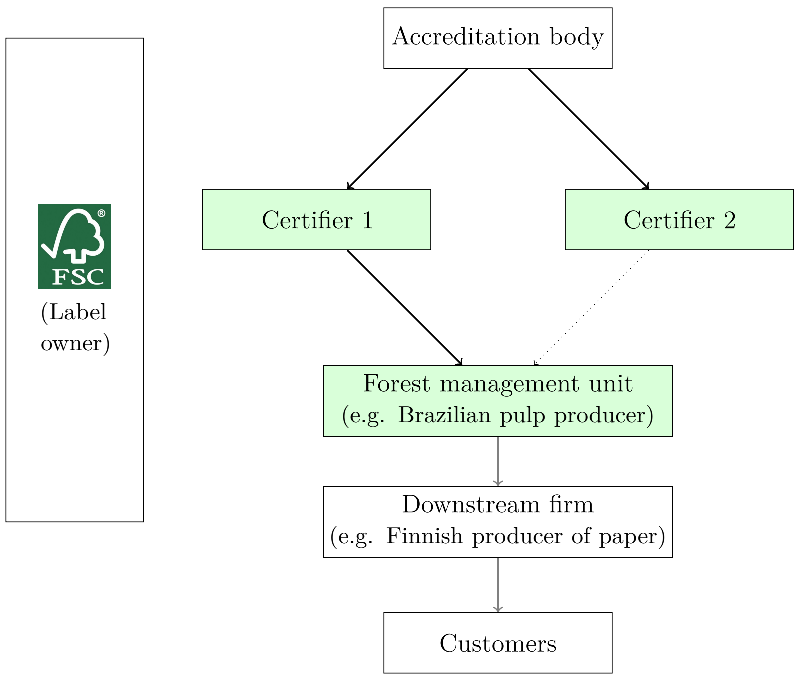
[12]FMUs are typically predefined administratively for publicly owned forests, while they can be more flexibly adapted by private owners.

[13]Jurists, forest scientists, and NGOs share this view (Clark and Kozar, 2011; Greenpeace International, 2021; Gutierrez Garzon et al., 2020; Ludwig et al., 2014).

[14]See Appendix Figure 1.5, FAO (2018) and UNECE/FAO (2019).

[15]In particular, major voluntary certification schemes let firms choose their certifier, including those covering EU organic production and ISO standards. Fair Trade is a rare exception.

Figure 1.1: Illustration of the FSC certification scheme



Notes: Black arrows symbolize monitoring through audits and inspections, the dotted arrow counterfactual monitoring if Certifier 2 would have been chosen. Gray arrows symbolize the supply chain. The blue color highlights the firms which are the focus of this paper. The position of FSC on the left symbolizes the relevance of its standards and label for all levels of the monitoring and supply chain.

certification based on audits, as described below. Downstream firms can use the FSC label on wood, paper, and other forest products if they source sufficiently from FSC-certified forests.[16] This paper restricts attention to FMUs' certification according to FSC's forest management standard, henceforth referred to as *the FSC standard*.[17]

The FSC standard regulates the environmental and social externalities of forest management rather than traditional quality characteristics of forest products. The quality on which the FSC standard focuses, thus, involves credence rather than experience attributes, i.e., not even observed after purchase (Darby and Karni, 1973). The rules include reducing the amount of logging in areas of high conservation value, sustainable logging volumes, and worker safety equipment. All rules stem from the following ten principles (FSC, 2015):

1. Compliance with Laws
2. Workers' Rights and Employment Conditions
3. Indigenous Peoples' Rights
4. Community Relations
5. Benefits from the Forest
6. Environmental Values and Impacts
7. Management Planning
8. Monitoring and Assessment
9. High Conservation Values
10. Implementation of Management Activities

The standard has national and regional versions, but all versions must align with a common set of core rules.[18] The violations considered in this paper are mostly violations of those core rules, but I also control for regional differences.

Firms' costs and benefits from FSC certification are heterogeneous. The costs include opportunity costs from compliance with the standard and certification fees for the certifiers.[19]

[16]That certification follows FSC's chain-of-custody standard.

[17]Most certified FMUs are also certified according to the chain-of-custody standard since they sell their forest products. They typically choose their certifier and pay their fees jointly for both standards. Since the audit results for chain-of-custody certification are not public, I can only account for the impact of the forest management standard on a firm's certifier choice. I exclude a small number of FMUs that are not simultaneously certified according to the chain-of-custody standard.

[18]These core rules include the "Principles and Criteria" and additional policies and procedures specifying their interpretation. These rules are written and voted on by an assembly of FSC members from environmental NGOs, the wood industry, and representatives of communities from forest areas.

[19]These costs also include anegligible FSC membership fees, which amount to only 55-160 USD for a median-sized certified forest, depending on the forest type (FSC, 2016a).

There is little public information on such costs. FSC does not regulate certification fees, except requiring that these fees follow a general pricing schedule.[20] Small-scale surveys suggest a total cost of approximately 6 to 40 USD per hectare and year for FMUs of less than 4,000 hectares and 0.07 to 0.5 USD for large FMUs of more than 400,000 hectares in the Americas, with similar costs in Japan (Chen et al., 2010; Cubbage et al., 2009; Sugiura and Oki, 2018). However, the design of these surveys suggests that the responses do not include all opportunity costs from compliance with the FSC standard, such as reduced revenues due to restricted logging rates.[21]

The economic benefits of FSC certification vary across markets and firms. FSC certification often secures continued business with downstream buyers.[22] Some markets provide a direct premium of up to 56% for FSC-certified wood.[23] According to a meta-analysis covering mostly advanced economies, the median consumer is willing to pay a price premium of approximately 10% of the retail price (Cai and Aguilar, 2013). A survey commissioned by FSC suggests that about every second consumer in 33 countries recognizes the FSC label and approximately every fourth consumer is willing to pay more for an FSC-certified product than for an uncertified product (FSC and IPSOS, 2023). Overall, small-scale surveys and case studies suggest that additional revenues exceed the total cost of certification for only roughly every second certified FMU.[24] For other FMUs, less tangible benefits drive participation. For example, certification makes it easier to obtain licenses to import wood products into

[20]The rules state that a general schedule of certification fees should be publicly available (FSC, 2015), but FSC does not enforce this requirement entirely. Instead, interested FMUs have to ask for a quote from certifiers. One certifier states that the fees depend on the "operation's size, geographic locations, and the complexity of factors such as forestry activities, high conservation values, stakeholder relations, etc." (NEPCon, 2019). An interviewed association of certified FMUs in Canada suggested that the fees are typically not negotiated.

[21]In Cubbage et al. (2009), only four out of 14 respondents included an estimation of the cost of "changes required to get/maintain certification". Given the cost quoted by those respondents (0.059 USD per hectare at the median) and the formulation of the question, it is unlikely that this cost includes all foregone profits. The situation is similar in Sugiura and Oki (2018). For Sweden, Villalobos et al. (2018) report only average direct payments for FSC and PEFC certifiers, suggesting hectare rates of 1-2 USD for recertification and 10-30 cents for annual surveillance audits.

[22]See Araujo et al. (2009); Overdevest and Rickenbach (2006); Galati et al. (2017); FSC (2018a); Holopainen et al. (2015). Forest management certification is sometimes a condition for participation in public procurement, as is the case in Germany (Ludwig et al., 2014).

[23]Frey et al. (2018) find increased revenues of 10-30% in Vietnam and Kollert and Lagan (2007) report a certification premium of up to 56% for high quality tropical wood in Malaysia. FSC certification may also be a precondition or facilitator for public subsidies (Visseren-Hamakers and Pattberg, 2013). However, most Japanese firms surveyed by Sugiura and Oki (2018) and a Canadian smallholder I interviewed do not see any direct economic benefits. FSC is by far not the only label, for which not all certified producers receive a price premium (Dragusanu et al., 2014; Subervie and Vagneron, 2013).

[24]See Cubbage et al. (2009); Frey et al. (2018); Kitchoukov et al. (2019); Sugiura and Oki (2018); Owari and Sawanobori (2007).

countries with timber legality regulations, such as EU member states.[25] Surveys suggest that many firms seek certification to improve their reputation, potentially hoping for more tangible benefits in the future. Social prestige concerns and intrinsic benefits due to learning about sustainable practices also play a role.[26] For state-owned firms, certification is often a political decision.[27]

1.2.3 FSC certifiers

FSC certifiers are both for-profit and nonprofit companies. They evaluate FMUs' compliance with the FSC standard in annual *surveillance audits* and in more extensive *(re)certification audits* in the first year and every fifth year (FSC, 2015).

Violation reporting: In every audit, certifiers report minor and major violations of compliance with the standard. *Major violations* are significant failures of compliance with FSC's core rules and the focus of this paper.[28] Examples include insufficient monitoring to protect endangered species or a management plan with logging volumes that would not sustain the forest cover. Many violations concern workers' rights, such as insufficient protection clothing, and insurance. Others relate to the local community and the state, such as the omission of consulting indigenous communities before making the harvest plan.[29] Reporting such violations is the main way certifiers enforce compliance with the FSC standard. If certifiers find five or more major violations in one audit, they must suspend the FMU's certificate (FSC, 2009b). If they find under five, they must check their correction in an additional audit after three months and suspend the certificate otherwise.[30] Most FMUs, therefore, correct their reported violations. For a given level of compliance ex ante,

[25] See Holopainen et al. (2015), Gavrilut et al. (2015) and the recognition of an FSC certification body as providing sufficient evidence for timber legality in Cameroon (FSC, 2016b).

[26] See the surveys by Araujo et al. (2009); Overdevest and Rickenbach (2006); Galati et al. (2017) and Paluš et al. (2021).

[27] This has been highlighted in interviews with industry insiders in Germany and Austria.

[28] If not specified otherwise, the violations in this paper are always major violations. Note further that certifiers must upgrade minor violations to major violations if FMUs do not correct these violations within a year.

[29] These examples come from the descriptions of 35 violations publicly reported in 20 audits randomly drawn from all observations with at least one violation in the forest unit panel (described in Section 1.3.1). In a sample of 110 audit summaries from Brazil in 2016, Rafael et al. (2018) find that most of the violations relate to environmental issues, community relations, and workers' rights. Furthermore, Blackman et al. (2017) examine the audit summaries of 35 FMUs in 2000-2013 in Mexico and find that most violations relate to social and economic-legal issues rather than environmental ones.

[30] Certifiers may lift the suspension if the FMU corrects the major violations within a year. Otherwise, the certificate is withdrawn (FSC, 2015). An FMU can apply for a new certificate after the withdrawal of a previous certificate (FSC, 2021c), but doing so requires the correction of all major violations.

more violation reports, thus, translate into greater compliance and, hence, fewer negative externalities and higher quality among the certified.

For FMUs, violation reports generate the cost of an additional audit plus the opportunity cost of correcting the violation or losing the certificate. There is no estimate of those costs in the literature thus far. The additional audit needs only to check that the reported violations have been corrected. Hence, the fee is likely a minor fraction of the usual annual certification fee. Correcting violations requires that the violation be ended and that procedures are adjusted to prevent similar violations in the future. The cost of such corrections varies. Among the violation reports from 20 randomly drawn audits, most corrections require relatively low levels of investment, such as purchasing a missing protection cloth for loggers. About one-third of the corrections seem to require medium-range investments, such as assessing a road's environmental impact before construction. In addition to the direct cost, roughly half of all violation reports seem to bear a small risk of much higher opportunity cost in the future, for example, if an environmental impact assessment concludes that the FMU cannot build a road due to the presence of protected species. A minimal share of violation reports requires the limitation of harvests and may, thus, lead to opportunity costs of over a hundred thousand USD.[31]

Certifiers have some discretion in auditing and reporting violations. For example, they have only to check a certain number of criteria and have some freedom in choosing those criteria. FMUs can transfer to another certifier any year, even though doing so is not encouraged by the FSC.[32] They can only transfer after correcting any major violations reported by the current certifier (FSC, 2010). FMUs obtain guidance and information from other companies when choosing their certifiers (Sugiura and Oki, 2018). Industry insiders suggest that the relevant factors for FMUs' certifier choice are fee differences, existing relationships with certifiers, efficiency, local presence and expertise of the certifier, but also differences in rigor.

Accreditation of FSC certifiers: FSC certifiers must be licensed by FSC's accreditation body, Assurance Services International (ASI).[33] ASI assesses certifiers' competence and

[31]From a case study by the NGO Earthsight (2020), it can also be crudely estimated that an FSC-certified Ukrainian forest enterprise was able to gain around 100,000 USD when illegal logging practices were undetected by their FSC certifier.

[32]If FMUs transfer to another certifier more than once in the five years between (re)certification audits, then they have to perform an additional (re)certification audit, implying increased costs.

[33]ASI is a company that conducts certifier assessments for sustainability standards in various industries (ASI, 2022). FSC is the unique shareholder of ASI but delegated "full business control" to ASI's supervisory board in 2017.

compliance by reviewing documents and inspecting their audits of FMUs and their offices at least once per year per certifier (ASI, 2019). For inspections of audits, ASI either is present as an observer while the certifier audits or inspects the FMU afterward to compare the results with those in the certifier's report. I refer to both cases as *accreditation inspections*. If ASI finds that certifiers do not fulfill their duties, it may suspend their accreditation globally or for some regions. In such cases, the affected FMUs have six months to contract new certifiers to keep their certification valid (FSC, 2015, Art. 1.1.3). Despite certifiers' accreditation, activists and journalists have accused certifiers of issuing certificates despite severe violations of the FSC standard (Alecci, 2023; Earthsight, 2021; FSC Watch, 2020). In addition, FSC member organizations demanded investigating threats to the impartiality of certifiers to maintain the FSC's credibility (FSC, 2019).

1.3 Data and summary statistics

I collect information from various sources to construct three novel datasets covering 2015-2019.[34]

1. Forest management units' (FMUs') characteristics, yearly demand for certifiers, and audit results.
2. Certifier characteristics at the market-level.
3. Certification fees from an anonymous survey of certified FMUs.

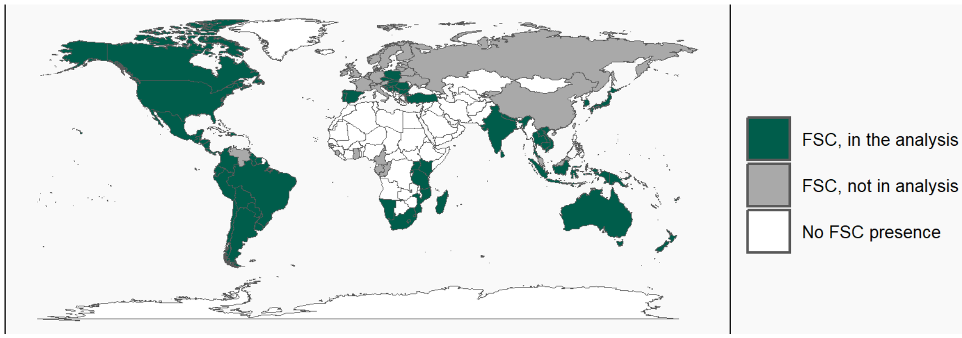
The FMU and the certifier-market data jointly form the main dataset, henceforth referred to as the FMU-certifier panel. The final datasets cover the following regions, as shown in Figure 1.2: the Americas except for Venezuela; Southern and Eastern Africa; Asia except for China and Malaysia; Oceania; and Eastern and Southern Europe except for Russia, Ukraine, and Italy. I exclude observations from other regions due to insufficient representation in the price panel or insufficient variation in market shares to identify the model parameters.[35] The included regions represent one-third of all FSC certificates. Audits in these regions result in

[34]I exclude the years before 2015 since the number of missing audit documents is high in those years. I exclude 2020 to avoid any disruptions caused by the COVID-19 crisis.

[35]In Western Europe, Sweden and Finland, Western and Middle Africa (particularly the Congo Basin), and the countries of the Commonwealth of Independent States (CIS, particularly Russia and Ukraine), there is too little variation in firms' certifier choices to precisely identify market-level alternative constants. In Italy, no firms chose the outside option, i.e., ended their FSC certification or were not certified yet, in 2019. CIS countries, the rest of Northern Europe (particularly UK), and China are not well represented in the survey providing the price panel. In Malaysia and Venezuela, important country-level statistics are missing, but these countries have very few certified FMUs.

more average numbers of violation reports than those in the excluded regions (see Appendix Figures 1.6).

Figure 1.2: FSC presence and the regional focus of this paper



The datasets include all FMUs certified for at least one year in 2015-2019.[36] Throughout the paper, I refer to such FMUs *once-certified* FMUs. There is no comprehensive dataset for those FMUs that have never been certified. The population this paper considers is, thus, a selected sample from the population of all FMUs. The results do not necessarily extend to the entire population of FMUs. I do not observe all relevant variables for all FMUs throughout 2015-2019, but I impute missing observations, as outlined in the following sections.

1.3.1 FMU panel

I construct a panel of yearly observations of once-certified FMUs primarily from information on the FSC website (FSC, 2020). Each FMU-year observation corresponds to one *choice situation*, that is, one choice of whether to participate in FSC certification and, if yes, the choice of one certifier. FSC publishes information about all certificates issued in the past, including the first issue date, termination date, country, size of the certified area, group certificate members, products and tree species, and audit documents (see Appendix Figures 1.7 and 1.8). I web-scraped that information and downloaded all audit documents in May and June 2020. The information does not specify the owner of the FMU.[37]

[36] I include those FMUs whose certificates became valid only in 2020 since they may still have received their first audit and made their certifier choice in 2019.

[37] This missing information implies that I cannot account for the joint ownership of various FMUs. That is, I have to assume independence of unobserved shocks from the benefits of certification for different FMUs with the same owner. This assumption probably does not affect the analysis greatly since separate certification suggests separate management, as owners have an interest in certifying jointly managed units through one certificate, given that joint certification lowers costs. Only 3% of the certified entities have names which appear on multiple certificates. The maximum number of certificates with the same name as the certified entity is 6. However, entities with different names may still have a joined owner.

From the audit documents, I extract the certifier, year, type of audit and the number of major violation reports. To do so, I write an algorithm using regular expressions. One-third of the numbers of violation reports are also checked manually, to ensure accuracy given the heterogeneity of document formats and languages. The extracted data cover 80% of all choice situations in which FMUs were certified, and include choice situations of 93% of all once-certified FMUs. I impute firms' certifier choices for the missing choice situations following reasonable assumptions. For example, suppose I have no observation for a year preceded and followed by a year with the same certifier. In that case, I assume that the FMU chose that certifier in the missing year since FMUs never switch the certifier two years in a row. Excluding imputed choices dramatically drives up the number of "unstructural zeroes", which can lead to more substantial bias (Briesch et al., 2008). Appendix Sections 1.9.1.1 and 1.9.1.2 detail the reasons for missing data and the data construction process, including the favorable results of accuracy tests.

I match the panel with additional data from various sources. First, I add indicators of corresponding accreditation inspections from the accreditation body's website (ASI, 2022). Second, I add a rich set of annual characteristics of the FMU's country to account for differences in the cost of compliance and demand for certification, including wood product trade values, a corruption index, and the number of certified downstream firms. See Appendix 1.9.1.4 for details. Third, I include an indicator for FMUs which are plantations. For about every fourth FMU, I find such information in audit documents or a recent FSC register. I use these FMUs to train a logit model to classify the remaining FMUs. I select relevant features from the available FMU characteristics with repeated cross-validation and LASSO regularization.[38]

Table 1.1 summarizes the most important characteristics of FMUs, their choices and audits. The data feature time variation in FMUs' participation in FSC certification which helps identify FMUs' willingness to pay for the certification: in 13% of the choice situations, FMUs do not choose FSC certification, yet, and 11% are terminated certificates. While most FMUs stick to their initial certifiers, 3% of all choice situations involve transfers to other certifiers.

Since there is no reliable observed measure of differences in certifiers' audit rigor,[39] this paper

[38]The prediction accuracy in a test sample is about 68%. I also attempt K-Nearest Neighbor matching, but the classification error is much smaller with the logit model.

[39]The only observed signals of rigor differences are past suspensions, complaints by NGOs, and mistakes that FSC's accreditation body reported from inspections in 2017-2019. However, even the accreditation body does not see those incidents as reliable measures of rigor, due to the small number of inspections, among

Table 1.1: Summary statistics of the forest unit panel

Statistic	N	Median	Mean	St. Dev.	Min	Max
Not yet certified	6,250	0	0.13	0.34	0	1
Termination of last year's certificate	6,250	0	0.11	0.31	0	1
Transfer to another FSC-certifier	4,596	0	0.03	0.17	0	1
No. of violation reports (by year)	3,810	0	0.60	1.65	0	33
At least one violation report	3,810	0	0.25	0.43	0	1
Five or more violation reports	3,810	0	0.02	0.14	0	1
Audit inspected by accreditation body	3,810	0	0.03	0.18	0	1
Yrs. with FSC cert.	6,250	5	5.98	5.48	0	26
(Re)certification audit	6,250	0	0.35	0.48	0	1
Classified as plantation	6,250	0	0.24	0.42	0	1
Logging is FMU's primary activity	6,250	1	0.91	0.29	0	1
Certified area in 1000 ha	6,250	12.98	110.09	342.75	0.005	5,986.68
Group certificate (vs. individual)	6,250	0	0.24	0.43	0	1
No. of certificate members	6,250	1	6.64	54.78	1	1,563

exploits variation in major violation reports to identify such differences. The numbers of violation reports per year range from 0 to 33. The median is zero. Appendix-Figure 1.9 shows that the distribution resembles a Poisson, with a longer tail of extreme observations. The average numbersof violation reports varies across certifiers, as Appendix-Figure 1.10 demonstrates. If certifiers audited the same population of FMUs, those averages could be used as a proxy of their rigor. However, this is not the case, as the examples in Appendix-Figures 1.11 and 1.12 illustrate. If less compliant FMUs select into more lenient certifiers, they might end up reporting more violations than more rigorous certifiers. This problem motivates the construction of a model that disentangles these differences. The model will exploit variation in the assignment of inspections by the accreditation body. Such inspections appear in 3% of all audits. They are associated with increased numbers of violation reports, as Appendix-Figure 1.24 (a) suggests.

The model will account for FMUs' characteristics which might affect their choices and violations. Table 1.1 summarizes a few important examples. Most once-certified FMUs are natural forests used for logging. The majority cover large areas of tens of thousands of hectares. A long tail goes up to more than 5 million hectares. Approximately one-fifth are group certificates. The median group has 6 members. The oldest certificate has been valid for 26 years, with the median being four. A more extensive (re)certification audit is needed in about one-third of choice situations.

other things.

1.3.2 Market definition, choice sets and certifier-market panel

I define a market as the combination of the year and a world region. Every year, there is a new market since certifiers tend to charge certification fees annually and FMUs can switch certifiers annually. A market covers a world region defined by two criteria. First, a region corresponds to a United Nations subregion (United Nations, 2023). I exclude certifiers from a region's choice set if they did not certify any FMU in the region in 2015-2019. If they did not certify in the region in the first or last one to four years, I exclude them from the region's choice set for those years, following Briesch et al. (2008). I do so since certifiers are not available in all countries for which they are accredited. In contrast, an office presence in one country can cover countries nearby without public information concerning their availability. Second, if a certifier did not have accreditation for a subset of countries in a given region, I split the region accordingly to have a common choice set within each market. The data cover ten world regions over five years and, hence, 50 markets.

Table 1.2: FSC certifiers active in 2015-2019

Certifier	Global entry	# active markets	Share across markets 2015 (%)	Share across markets 2019 (%)	Head-quarters	# countries with offices	Type	Company revenue 2020 (MM USD)	Past suspen-sions
1	2011	50	38	29	DNK	60	Non-profit	23	0
2	2005	26	6	4	FRA	75	Traded	5228	2
3	2005	29	3	5	NLD	64	Private	17	0
4	2000	24	5	8	DEU	9	Private	159	0
5	1995	40	14	14	GBR	50	Non-profit	30	0
6	1995	40	15	16	USA	10	Private	38	0
7	1995	45	15	16	ZAF	123	Traded	5962	1
Small certifiers	1998 (earliest)	12	3	8	DEU (largest)	137	Mixed		0

I collect information about certifiers from their websites and the accreditation body. Table 1.2 presents their main characteristics. Throughout the paper, I focus on the seven largest certifiers who certified over 90 percent of once-certified FMUs.[40] I merge the remaining certifiers into a competitive fringe, as their certification activities are too small to estimate separate parameters precisely. Most large certifiers have been active since the early years of FSC, but the certifier with the highest market share did not join until 2011. Certifiers differ in terms of their market presence. Market shares vary only slightly over time but greatly across regions, as Appendix Figures 1.13 show. The consequent variation in FMUs'

[40] I further consider two, originally separate certifiers, Certifiers 1(a) and 1(b), as a single certifier since Certifier 1(b)'s certification department was acquired by Certifier 1(a) in 2018. Almost all staff working on FSC certification at Certifier 1(b) moved, and certificates transferred quite easily, as industry insiders suggest.

choice sets is helpful for the model's identification. Most certifiers have headquarters in the Global North but offices in many countries. Two large certifiers are nonprofits specializing in sustainability standards, mainly agriculture and forest management. Most for-profit certifiers do business in a broader range of industries and have higher revenues. The two publicly traded certifiers are the only certifiers that the accreditation body suspended in 2010-2020. Certifier 7's suspension in 2011 led to its permanent exclusion from Brazil. Certifier 2 was suspended worldwide for a few months in 2016-2017 and has been excluded from Russia since 2015 (ASI, 2020).

1.3.3 Price panel

As FSC certifiers do not publish their certification fees, I surveyed FMUs.[41] I e-mailed the survey to all FMUs with valid certificates in June 2020 for whom the audit reports or their websites included an e-mail address. Owing to a response rate of 21%, I obtained 387 fee quotes and corresponding FMU characteristics from the world regions on which this paper focuses. I convert the quotes into real USD prices using the World Bank (2020)'s currency conversion factor and the US paper and wood pulp producer price index (FRED, 2020), with 2015 as the base year. I conducted the survey anonymously and asked about FMUs' characteristics only in broad categories to encourage participation. In particular, I asked about the size of the certified FMU in six categories from "< 1,000 ha" to "> 500,000 ha" and for the 5-year-interval in which the FMU was initially certified. In the analysis, I replace those categorical responses with the within-category averages by market from the FMU panel, i.e., the whole population of once-certified FMUs, as a numeric characteristic. Appendix 1.9.3 describes the design and outcomes in detail.

Table 1.3: Summary statistics of the price panel

Statistic in 1000 USD	N	Median	Mean	St. Dev.	Min	Max
Annual cert. fee in 1000 USD, PPI adj.	387	7.37	9.41	8.24	0.67	58.20
Annual cert. fee in USD/ha	387	0.39	1.48	2.79	0.01	20.05
Certified forest area in 1000ha (market-mean by category)	387	22.84	57.66	155.30	0.35	1,399.48
Classified as plantation	383	1	0.56	0.50	0	1
Group certificate (vs. individual)	387	0	0.35	0.48	0	1
(Re)certification audit	384	0	0.28	0.45	0	1
Yrs. with FSC cert. (market-mean by category)	382	8.88	8.92	6.18	0.00	23.00

Table 1.3 presents summary statistics of the fee quotes and respondents' most important characteristics. There is an enormous amount of variation in certification fees. The median

[41] Certifiers were unwilling to share their fees upon request.

fee among survey respondents is approximately 7000 USD, roughly 0.4 USD per hectare. Fees range from a few cents to more than 25 USD per hectare. Such fees are consistent with case studies (Chen et al., 2010; Cubbage et al., 2009; Sugiura and Oki, 2018). Respondents represent the variety of FMU types and regions quite well, as detailed in Appendix 1.9.3.2. They tend to be slightly larger than in the whole population, are more often plantations, and have been certified for a longer time. The model estimation will account for such sample selection, but the selection is most likely helpful for the response precision, as managers of larger FMUs probably record fees paid for certification even more reliably than managers of smaller FMUs. I check for consistency in the responses in various dimensions and exclude a minimal number of unreliable responses. Many respondents raised additional confidence in their motivation to respond thoroughly through additional, detailed comments they made in writing.

I use the price panel to predict prices for all certifiers and FMUs. Specifically, I predict prices as the product of a certifier-market-level price and an FMU-market-level price factor. I motivate that in the next section, when describing the model. I select regressors of total prices from a large set of variables which vary either by FMU i and market t ($\boldsymbol{f_{it}}$) or by certifier j and market t ($\boldsymbol{x_{jt}}$). The certifier-variant variables include functions of certifiers' market shares to capture variation in markups. I choose relevant regressors using repeated cross-validation with LASSO regularization (Hastie et al., 2001). I regress log prices per log mean certified forest area by category and market, $\log(\overline{area}_{it})$, on the selected regressors. I then use the estimated coefficients and the FMU-certifier-panel to predict certifier-market-level prices p_{jt}, the individual price factors c_{it} and total prices p_{ijt}.

$$\log\left(\frac{p_{ijt}}{\log(\overline{area}_{it})}\right) = \boldsymbol{f}'_{\boldsymbol{it}}\tilde{\boldsymbol{\rho}}^{\boldsymbol{f}} + \boldsymbol{x}'_{\boldsymbol{jt}}\tilde{\boldsymbol{\rho}}^{\boldsymbol{x}} + \tilde{\epsilon}^{p}_{ijt} \tag{1.1}$$

$$\mathbb{E}[p_{jt}|\boldsymbol{x_{jt}}] = \exp(\boldsymbol{x}'_{\boldsymbol{jt}}\tilde{\boldsymbol{\rho}}^{\boldsymbol{x}}) \qquad \mathbb{E}[c_{it}|\boldsymbol{f_{it}}] = \log(\overline{area}_{it})\exp(\boldsymbol{f}'_{\boldsymbol{it}}\tilde{\boldsymbol{\rho}}^{\boldsymbol{f}}) \qquad \mathbb{E}[p_{ijt}|\boldsymbol{f_{it}}, \boldsymbol{x_{jt}}] = c_{it}p_{jt} \tag{1.2}$$

However, the price panel is not a representative but a selected sample of the prices offered to all FMUs by all available certifiers. I only observe prices of (1) FMUs that decided to participate in the survey and (2) for the certifiers chosen by the FMUs. The included regressors already account for many of the characteristics that influence those choices and prices. However, remaining sample selection due to unobserved factors might still cause bias in the prediction exercise. Fortunately, I can control for that selection bias based on assumptions about the joint distribution of the error terms in the different models. First, I assume that the error terms of the utility from the choice to participate in the survey and the choice of the certifier are independent, conditional on the other price regressors. I can then control for those two sources of bias separately. Second, I assume that the error term of the survey participation and the pricing prediction models are jointly normal. The Inverse Mills Ratio of the utility of survey participation predicted from a probit model then accounts for the bias from that selection (Heckman, 1979), as outlined in Appendix 1.9.3.3.1. The control functions derived by Lee (1983) account for bias due to the choice of the certifier.[42]

Table 1.4: Summary statistics of the price predictions

Statistic	N	Median	Mean	St. Dev.	Min	Max
Predicted price p_{ijt}	34,794	7.49	9.44	7.68	0.31	128.29
Predicted certifier-market-level price p_{jt}	34,794	3.21	3.33	1.33	1.27	11.67
Predicted FMU-market-level price factor c_{it}	34,794	2.41	2.90	2.12	0.18	25.80

Table 1.4 shows the price predictions for all FMU-certifier-year observations. While the outlined prediction procedure aims at a good balance between bias and explained variance in out-of-sample prediction, the predictions still explain 72% of the in-sample variation in the price quotes. Differences between FMUs and markets matter most, but certifier-variant predictors still explain 20% of the predicted variation. At all quartiles, the predicted prices are lower than the prices quoted in the survey. This difference is not surprising. FMUs that pay higher prices might see more benefits in FSC certification and contribute more willingly to research. The difference suggests that controlling for selection into the survey is important.

1.4 Empirical model

I build an empirical model of voluntary certification. I apply the model to FSC certification of forest management units (FMUs), but the main ideas are applicable for many certification schemes. The model describes how certifiers' rigor may impact violation reporting; That part of the model allows to identify a proxy of the differences in certifiers' rigor, tackling one of the main challenges of this paper: certifiers' rigor is not observed. Next, the model illustrates the role of certifiers' rigor in FMUs' participation and certifier choices, both directly and indirectly through certifiers' markups. It provides a framework to simulate how an increase

[42]The control functions follow the form in Equation (3.7) in Lee (1983). They are based on the transformation (1.18) of the nested logit error, as described in Section 1.5.1.1 of this paper.

in minimum rigor levels through stricter accreditation can affect violation reporting and participation.

There are T markets, defined by year and world region. In each market t, FMUs and certifiers play the following stage game. In Stage 1, each certifier j sets their certification fee. In Stage 2, each FMU i decides whether to participate in FSC certification and chooses a certifier. In Stage 3, certifiers audit FMUs and detect and report violations of the FSC standard.[43] This section describes the stages in reverse order.

1.4.1 Stage 3: Violation reporting

Violation reports v_{ijt} are an exponential function of certifier, audit and forest management unit (FMU) *types* plus a shock:

$$v_{ijt} = \underbrace{r_j \exp(\boldsymbol{x}^{v\prime}_{ijt}\boldsymbol{\omega})}_{\tilde{r}_{ijt}} \underbrace{\exp(\boldsymbol{f}^{v\prime}_{it}\boldsymbol{\gamma} + \eta_i)}_{v^0_{i1t}} + \epsilon^v_{ijt} \tag{1.3}$$

The empirical distribution motivates the functional form.

r_j denotes a fixed effect of certifier j. It captures certifier differences in violation reporting that vary across neither regions nor years.[44] It is a proxy for certifiers' *rigor type*. Rigor types might reflect long-run decisions and expertise, for example, due to the legal status and intrinsic motivation of owners and managers. The intercept γ_0, contained in $\boldsymbol{f}^{v\prime}_{it}\boldsymbol{\gamma}$, absorbs the rigor type of one of the certifiers. I consider Certifier 1 as that *baseline certifier* since Certifier 1 is present in all markets and has the largest global market share. I normalize r_1 to 1. r_j then measures the fraction of violation reports of Certifier 1 reported by certifier j, everything else equal.

Certifiers' overall rigor may vary across audits. It depends on $\boldsymbol{x}^v_{ijt}$, a vector of two audit characteristics: an indicator for an accreditation inspection and the distance between the market and the certifier's headquarters. These factors reflect hypotheses shared by industry insiders. First, certifiers tend to report more violations when the accreditation body inspects the audit they conducted. Second, the more distant the market is from the headquarters, the lower the degree of quality control of auditors and, hence, the level of rigor.

[43] If an FMU chooses an FSC certifier in the subsequent year, I conclude that it corrected the reported violations, as is necessary under the FSC standard.

[44] The model applies to a few years, a period in which companies are unlikely to change drastically. In the application of this paper, five years are considered.

Violation reports also vary with certifier-invariant characteristics of the audit, the FMU and its market. The vector $\boldsymbol{f^v_{it}}$ includes observable characteristics, while η_i is an unobserved compliance type. $\boldsymbol{f^v_{it}}$ can affect both FMUs' actual violations and certifiers' rigor. For example, FMUs that have been certified for many years might violate the standard less often than newly certified FMUs, but certifiers might also detect more violations among these FMUs, as they have more information from previous audits. I assume only that the rate at which $\boldsymbol{f^v_{it}}$ affects rigor does not differ systematically across certifiers.

Overall, the number of violation reports is the product of relative rigor, $\tilde{r}_{ijt} \equiv r_j \exp(\boldsymbol{x^{v\prime}_{ijt}\omega})$, and baseline violations, $v^0_{i1t} \equiv \exp(\gamma_0 + \boldsymbol{f^{v\prime}_{it}\gamma} + \eta_i)$, plus the shock ϵ^v_{ijt}. *Baseline violations* are the violations reported by the baseline certifier in an audit in the region of its headquarters, without accreditation inspection, apart from the shock ϵ^v_{ijt}. As baseline violations do not vary across certifiers, they capture ordinal differences in FMUs' levels of compliance. *Relative rigor* is the fraction of baseline violations that certifier j would report. As the model focuses on types and exogenous determinants of baseline violations and relative rigor, it describes a setting of adverse selection rather than moral hazard. Shock ϵ^v_{ijt} may create additional, random variation in violation reports. One the one hand, the shock may capture variation in rigor, for example, through auditors' concentration. On the other hand, it may also capture variation in compliance, for example through a shortage of FSC-compliant logging contractors.[45]

FMUs form expectations about their own compliance and certifiers' rigor, when deciding about participation in FSC certification and their certifier. Regarding compliance, I assume that FMUs predict the number of baseline violations conditional on their characteristics, $\mathbb{E}[v^0_{i1t}|\boldsymbol{f^v_{it}}] = \exp(\boldsymbol{f^{v\prime}_{it}\gamma})$. These are their *expected baseline violations.* FMUs do not explicitly account for their unobserved compliance type η_i or the shock ϵ^v_{ijt} in their participation and certifier choice. These assumptions reflect the uncertainty reported by forest managers regarding their own compliance.[46] Nevertheless, η_i may correlate with a shock to the participation and certifier choice, as formalized in the next section. For example, part of forest managers' intrinsic motivation which does not correlate with their observed characteristics may affect not only their compliance, but also their preference for certain certifiers.

[45] A certified forest manager in Canada, for example, reported a shortage of logging firms in her region, which reduces the forest managers' abilities to force the contractors to use FSC-compliant logging methods.

[46] Moreover, I could not consistently estimate η_i from the violations model. Attempts to account for $\exp(\eta_i)$ as random coefficients around $\mathbb{E}[v^0_{i1t}|\boldsymbol{f^v_{it}}]$ were not successful, likely due to the limited degree of variation in the data.

Regarding certifiers' relative rigor, FMUs' expectation is $\mathbb{E}[\tilde{r}_{ijt}|\overline{\boldsymbol{x}}^{\boldsymbol{v}}_{\boldsymbol{jt}}]$, where $\overline{\boldsymbol{x}}^{\boldsymbol{v}}_{\boldsymbol{jt}}$ is the certifier-market-level average of $\boldsymbol{x}^{\boldsymbol{v}}_{\boldsymbol{ijt}}$. That is, FMUs observe certifiers' rigor type, r_j, and the market's distance to the certifiers' headquarters, denoted as x^v_{1jt}. FMUs do not observe which audits will be inspected by the accreditation body, x^v_{2ijt}, but perfectly predict the likelihood of accreditation inspections for a given certifier in the market, i.e. the average $\overline{x}^v_{2jt}$.[47] Given these assumptions, the *expected relative rigor* is

$$ExpectedRigor_{jt} = r_j \exp(\omega_1 x^v_{1jt})\Big(1 + (\exp(\omega_2) - 1)\overline{x}^v_{2jt}\Big) \tag{1.4}$$

These assumptions capture insights shared by interviewed forest managers. On the one hand, they report that they ask managers of other certified FMUs about their experience with certifiers' stringency and expertise when choosing their certifier. On the other hand, they also report uncertainty, here captured by the shock ϵ^v_{ijt} and the indicator for accreditation inspections, x_{2ijt}.

1.4.2 Stage 2: Demand

Each forest management unit (FMU) i chooses whether to participate in FSC certification ($j > 0$) or not ($j = 0$). If FMUs choose FSC certification, they each have to select one available certifier in their market t, $j \in J_t$. Since each FMU makes these choices each year, a unique choice situation is defined by FMU i and market t. The FMU chooses the option that maximizes their surplus u_{ijt}, indicated by y_{ijt}:

$$\begin{aligned} y_{ijt} &= \mathbf{1}(u_{ijt} \geq \max_{k \in J_t} u_{ikt}) \\ u_{ijt} &= \underbrace{\alpha_{it} p_{jt} + \beta^r_{it} ExpectedRigor_{jt} + \boldsymbol{x}^{\boldsymbol{u}'}_{\boldsymbol{jt}} \boldsymbol{\beta}^{\boldsymbol{x}} + \boldsymbol{d}'_{\boldsymbol{ijt}} \boldsymbol{\beta}^{\boldsymbol{d}} + \boldsymbol{f}^{\boldsymbol{u}'}_{\boldsymbol{it}} \boldsymbol{\xi}_{\boldsymbol{j}} + \xi_t + \Delta\xi_{jt}}_{V_{ijt}} + \epsilon^u_{ijt} \end{aligned} \tag{1.5}$$

The characteristics of main interest are p_{jt} and $ExpectedRigor_{jt}$. p_{jt} is the certification fee, $ExpectedRigor_{jt}$ is the expected relative rigor. β^r_{it}, thus, captures FMUs' preference or dislike of certifiers' expected relative rigor. $-\beta^r_{it}/\alpha_{it}$ quantifies the willingness to pay for leniency as $-\alpha_{it}$ captures the marginal utility of income.

The observable certifier characteristics $\boldsymbol{x}^{\boldsymbol{u}}_{\boldsymbol{jt}}$ are the certifiers' distance to their headquarters and an indicator for being available in a given market for the first year. Both characteristics

[47] When FMUs choose certifiers, they do not know whether the accreditation body will inspect their audits by particular certifiers since the accreditation body needs to know the certificates issued by certifiers to assign inspections (ASI, 2019).

may affect the certifiers' familiarity and local expertise in market t. $\boldsymbol{f}_{\boldsymbol{it}}^{\boldsymbol{u}\prime}\boldsymbol{\xi}_{\boldsymbol{j}}$ are certifier fixed effects interacted with a constant and indicators for plantations and group certificates. They account for unobserved popularity of the certifier throughout the population and in those types of FMUs. $\Delta\xi_{jt}$ is an unobserved demand factor that may depend on exogenous shocks and certifier decisions such as advertising.

$\boldsymbol{d}_{\boldsymbol{ijt}}$ is a vector of interactions of certifier and FMU characteristics. First, an indicator for participating in FSC certification for the first time accounts for entry cost. Second, I capture switching cost by including an indicator for having been certified by the same certifier in the previous year. I also include interactions of this indicator with the age of the certificate and an indicator for the end of a five-year cycle at which switching is less costly. Third, $\boldsymbol{d}_{\boldsymbol{ijt}}$ includes an interaction of the number of years certifier j had been accredited in the first year in which FMU i was certified. That captures the importance of certifiers' experience and familiarity in FSC certification at the time of FMUs' entry. Fourth, $\boldsymbol{d}_{\boldsymbol{ijt}}$ includes FMU characteristics interacted with a dummy for FSC certification. These interactions capture variation in certifier-invariant net benefits from participation across FMUs within the same market, while fixed effects ξ_t capture such variation across markets. Specifically, I allow for variation in those net benefits across FMUs with different clades of plants, countries' corruption perception indices and export values of wood chips.

The individual preference parameters α_{it} and β_{it} capture baseline preferences within the whole once-certified FMU population and heterogeneity in those preferences related to predicted FMU characteristics:

$$\begin{aligned} \alpha_{it} &= \bar{\alpha} + \tilde{\alpha} c_{it} \\ \beta_{it}^r &= \bar{\beta}^r + \tilde{\beta}_1^r \mathbb{E}[v_{i1t}^0 | \boldsymbol{f}_{\boldsymbol{it}}^{\boldsymbol{v}}] + \tilde{\beta}_2^r \mathbf{1}\left\{\max\left(\mathbb{E}[v_{i1t}^0 | \boldsymbol{f}_{\boldsymbol{it}}^{\boldsymbol{v}}] ExpectedRigor_{jt}\right)_{it} \geq 5\right\} \end{aligned} \tag{1.6}$$

c_{it} are predicted differences in cost, defined in the next section. $\mathbb{E}[v_{i1t}^0|\boldsymbol{f}_{\boldsymbol{it}}^{\boldsymbol{v}}]$ are FMUs' expected baseline violations, defined in Section 1.4.1. $\mathbf{1}\{\max(\mathbb{E}[v_{i1t}^0|\boldsymbol{f}_{\boldsymbol{it}}^{\boldsymbol{v}}]ExpectedRigor_{jt})_{it} \geq 5\}$ is an indicator for those choice situations in which FMUs expect the most rigorous certifier in their market to report at least five violations and, thus, to suspend their certificate. Using those predictions from the violations and pricing model to capture heterogeneity in preferences, instead of directly observable FMU characteristics allows for reduced dimensionality.

I normalize the mean utility of the outside option of not choosing FSC certification to zero: $u_{i0t} = \epsilon_{i0t}^u$. The taste shock ϵ_{ijt}^u follows an extreme value distribution according to a nested logit with two nests. The first nest is the set of FSC certifiers J_t, and the second is the outside option of no FSC certification, $j = 0$. The nest parameter λ measures the correlation of taste shocks for choosing an FSC certifier relative to choosing no FSC certification. Lower values of λ indicate higher substitutability among certifiers. $\lambda = 1$ implies independent taste shocks, translating into a multinomial logit model. The probability of choosing certifier j is the product of the probability of choosing FSC times that of choosing j conditional on having chosen FSC:

$$s_{ijt} \equiv s_{itFSC} s_{ijt|FSC} \tag{1.7}$$

where

$$s_{ijt|FSC} = \frac{\exp(V_{ijt}/\lambda)}{\sum_{k \in J_t} \exp(V_{ikt}/\lambda)}$$

and

$$s_{itFSC} = \frac{\left(\sum_{k \in J_t} \exp(V_{ikt}/\lambda)\right)^{\lambda}}{1 + \left(\sum_{k \in J_t} \exp(V_{ikt}/\lambda)\right)^{\lambda}}$$

These formulas show that the decision to participate in FSC certification depends on the characteristics of all available certifiers, including their expected relative rigor. Increasing minimum rigor levels or removing lenient certifiers from the choice sets, thus, inevitably affects participation in FSC certification. Estimating the model and simulating those changes allows assessing the size of the effects.

1.4.3 Stage 1: Pricing

I model certifiers pricing strategies to account for price effects in the counterfactual analysis of stricter accreditation. I build on the constant markup model (Train, 2009), which aligns with pricing practices in many industries (Shim and Sudit, 1995). That is, certifiers set markups in each market as a constant factor k_{jt} of the marginal cost c_{ijt}. I specify the markups as follows: the competitive fringe prices at cost ($k_{jt} = 1$), but all larger for-profit certifiers set k_{jt} to maximize profits, in best response to their competitors. Nonprofit certifiers price at a fixed, low margin. The model aligns with three features of many certification markets, including FSC. First, a small number of large certifiers dominate the market. Second, certifiers are differentiated. Third, the cost of certification varies by establishment (here an FMU) and market.

I add two more assumptions. First, I assume that the marginal cost c_{ijt} is the product of a certifier-market-level cost c_{jt} and an individual cost factor c_{it}. This aligns the pricing model with the demand model, where c_{it} is part of the price coefficient. Second, certifiers observe only the average of the individual cost factor per market, $\bar{c}_t = 1/N_t \sum_{i=1}^{N_t} c_{it}$, and market shares, rather than individual choice probabilities, when setting the markup. Each for-profit certifier j thus sets k_{jt} to maximize the expected profit based on that information:

$$\mathbb{E}[\pi_{jt}|c_{jt}, \bar{c}_t, s_{jt}, j \in J_t] = N_t(k_{jt} - 1)c_{jt}\bar{c}_t s_{jt}(\boldsymbol{k_t}) \tag{1.8}$$

where $\boldsymbol{k_t}$ is the vector of all markup factors in market t and $j \in J_t$ means that j is available in market t. Certifiers offer only a single version of certification for a given standard, which, here, is the FSC standard. The first-order condition is thus

$$(k_{jt} - 1)\mathbb{E}\left[\left(\frac{\partial s_{jt}(\boldsymbol{k_t})}{\partial k_{jt}}\right) \Big| c_{jt}, \bar{c}_t, s_{jt}\right] + s_{jt}(\boldsymbol{k_t}) = 0 \tag{1.9}$$

where

$$\mathbb{E}\left[\left(\frac{\partial s_{jt}(\boldsymbol{k_t})}{\partial k_{jt}}\right) \Big| c_{jt}, \bar{c}_t, s_{jt}\right] = \frac{\bar{\alpha} + \tilde{\alpha}\bar{c}_t}{\lambda} c_{jt} s_{jt} \left[1 - \lambda s_{jt} + (\lambda - 1)s_{jt|FSC}\right] \tag{1.10}$$

Overall, the price that an FMU i pays to certifier j is as follows:

$$p_{ijt} = p_{jt}c_{it} = k_{jt}c_{jt}c_{it} \tag{1.11}$$

The model departs from the standard Nash Bertrand pricing model in oligopolistic markets since certification costs vary across individual customers, here the FMUs. The standard model could account for such heterogeneous cost by modelling group-based price discrimination as in D'Haultfoeuille et al. (2019). However, the available data are not rich enough to estimate such a model.[48] The present model provides an alternative that imposes less sophistication on certifiers' pricing policies, but still captures profit-maximization when markups are set.

1.5 Estimation and results

The model is estimated sequentially, in reverse order.

[48]When splitting the FMUs into appropriate groups as in D'Haultfoeuille et al. (2019), the number of FMUs per group becomes so small that there is a large number of zero market shares.

1.5.1 Stage 3: Violation reporting

1.5.1.1 Estimation

To bring the model of violation reporting to the data, I add two assumptions. First, I assume that the unobserved compliance type is normally distributed in the population: $\eta_i | j, \boldsymbol{x^v_{ijt}}, \boldsymbol{f^v_{it}} \sim N(0, \sigma_\eta)$. This implies that $\mathbb{E}[\exp(\eta_i)|j, \boldsymbol{x^v_{ijt}}, \boldsymbol{f^v_{it}}] = \exp\left(\frac{\sigma^2_\eta}{2}\right)$. Second, I assume that the shock ϵ^v_{ijt} is independently and identically distributed for all certifiers and is exogenous to the determinants of relative rigor $\tilde{r}_{ijt}$ and baseline violations v^0_{i1t}.

I estimate the model using **Poisson pseudo-maximum-likelihood (PPML)**. To do so, I follow Santos Silva and Tenreyro (2006) by rewriting model (1.3) as

$$v_{ijt} = \exp(\mu^v_{ijt} + \eta_i + \tilde{\epsilon}^v_{ijt}) \tag{1.12}$$

where

$$\begin{aligned} \mu^v_{ijt} &\equiv log(r_j) + \boldsymbol{x^{v\prime}_{ijt}} \boldsymbol{\omega} + \boldsymbol{f'_{it}} \boldsymbol{\gamma} \\ \tilde{\epsilon}^v_{ijt} &\equiv \ln\left(1 + \frac{\epsilon^v_{ijt}}{\exp(\mu^v_{ijt} + \eta_i)}\right) \end{aligned} \tag{1.13}$$

This reformulation implies that $\mathbb{E}[\exp(\tilde{\epsilon}^v_{ijt})|j, \boldsymbol{x^v_{ijt}}, \boldsymbol{f^v_{it}}, \eta_i] = 1$. A random sample of potential violation reports, v_{ijt}, would thus allow to estimate the model's parameters consistently by minimizing the following log-likelihood (Gourieroux et al., 1984; Wooldridge, 2010):[49]

$$\mathcal{L}^v = \sum_{t=1}^{T} \sum_{i=1}^{N_t} \sum_{j=1}^{J_t} \mathbf{1}\{v_{ijt} \text{ is observed}\} \left[v_{ijt} \log\left(\mathbb{E}[v_{ijt}|j, \boldsymbol{x^v_{ijt}}, \boldsymbol{f^v_{it}}]\right) - \mathbb{E}[v_{ijt}|j, \boldsymbol{x^v_{ijt}}, \boldsymbol{f^v_{it}}]\right] \tag{1.14}$$

where the normal distribution of η_i implies

$$\mathbb{E}[v_{ijt}|j, \boldsymbol{x^v_{ijt}}, \boldsymbol{f^v_{it}}] = \exp\left(\mu^v_{ijt} + \frac{\sigma^2_\eta}{2}\right) \tag{1.15}$$

However, I do not observe a random sample of potential violation reports but the selected sample of violations reported by the certifiers FMUs have chosen. Estimating the model with this sample could lead to **sample selection bias**. In particular, less compliant FMUs might

[49] I estimate standard errors as being robust to heteroskedasticity and serial correlation. These standard errors correct for potential violations of the implicit PPML assumption of proportionality between the conditional mean and variance (Santos Silva and Tenreyro, 2006).

select less rigorous certifiers. These certifiers might, thus, end up reporting more violations than more rigorous certifiers.

To a substantial extent, the **observable controls** $\boldsymbol{f_{it}^{v}}$ account for such selection. I consider a large set of possible predictors derived from the variables in the FMU panel data and their interactions. I select the most relevant predictors by estimating model (1.12) with PPML using LASSO regularization with 5-fold cross-validation to avoid overfitting. I detail the procedure, including a few preliminary steps, in Appendix 1.10.1.1.

Nevertheless, unobservable compliance factors, such as intrinsic motivation, might still affect violations and firms' certifier choices. In the model, the unobserved compliance type η_i captures such factors.[50] In the rest of this section, I formally derive the bias it introduces to account for it with a **control function**.

In the selected sample, the observed mean of violation reports does not correspond to the population mean from Equation (1.15). Instead, it conditions on FMUs' choices $\boldsymbol{y_{it}} \equiv [y_{i1t}, ..., y_{iJ_tt}]$ and their determinants $\boldsymbol{V_{it}} \equiv [V_{i1t}, ..., V_{iJ_tt}]$, defined by the demand model (1.5):

$$\mathbb{E}[v_{ijt}|j, \boldsymbol{x_{ijt}^{v}}, \boldsymbol{f_{it}^{v}}, \boldsymbol{y_{it}}, \boldsymbol{V_{it}}] = \exp(\mu_{ijt}^{v}) \sum_{k\in\{1,...,J_t\}} y_{ikt}\, \mathbb{E}[\exp(\eta_i)|j, \boldsymbol{x_{ijt}^{v}}, \boldsymbol{f_{it}^{v}}, y_{ikt} = 1, \boldsymbol{V_{it}}] \quad (1.16)$$

$\mathbb{E}[\exp(\eta_i)|j, \boldsymbol{x_{ijt}^{v}}, \boldsymbol{f_{it}^{v}}, y_{ikt} = 1, \boldsymbol{V_{it}}]$ might vary across the chosen alternatives k and, thus, introduce sample selection bias. This is the case if η_i is correlated with the differences between the taste shocks $\epsilon_{i1t}^{u} - \epsilon_{ijt}^{u}, ...; \epsilon_{iJt}^{u} - \epsilon_{ijt}^{u}$ from model (1.5), even conditional on $\boldsymbol{f_{it}^{v}}$. Deriving the functional form of $\mathbb{E}[\exp(\eta_i)|j, \boldsymbol{x_{ijt}^{v}}, \boldsymbol{f_{it}^{v}}, y_{ikt} = 1, \boldsymbol{V_{it}}]$ allows to control for the bias.

The problem differs from the standard control function approach in two ways. First, the source of selection is a choice between multiple certifiers. It is a multinomial, not a binary choice. Second, the model is a count data model, so the conditional mean of $\exp(\eta_i)$ is not additive. To address these issues, I combine the approaches of Lee (1983) for multinomial selection problems and of Terza (1998) for count data models with binary selection. To my knowledge, I am the first to do so.

Following Lee (1983), I convert the problem of selection among J_t alternatives to J_t binary selection problems. To do so, I rewrite the selection problem from model (1.5) in terms of

[50] The number of periods in the data is insufficient to consistently estimate η_i as a fixed effect.

maximum order statistics:

$$y_{ijt} = 1 \text{ iff } V_{ijt} \geq e_{ijt} \text{ where } e_{ijt} \equiv max_{k \neq j}(V_{ikt} + \epsilon^u_{ikt} - \epsilon^u_{ijt}) \tag{1.17}$$

The marginal distribution of e_{ijt} is such that $F_j(V_{ijt}) = s_{ijt}$, the conditional choice probability defined in equation (1.7). e_{ijt} is transformed into a standard normal random variable by defining

$$e^*_{ijt} \equiv G_j(e_{ijt}) \equiv \Phi^{-1}(F_j(e_{ijt})) \tag{1.18}$$

e^*_{ijt} and η_i are thus jointly normally distributed with zero means, variances 1 and σ_η and correlation coefficient ρ_j, under the distributional assumption made on η_i. Since $G_j()$ is a strictly increasing function, (1.17) and (1.18) translate into $y_{ijt} = 1$ iff $e^*_{ijt} \leq G_j(V_{ijt})$.

As shown in Appendix 1.10.1.2, I can then follow the steps suggested by Terza (1998) to derive the parametric form of the bias as

$$\mathbb{E}[\exp(\eta_i)|j, \boldsymbol{x^v_{ijt}}, \boldsymbol{f^v_{it}}, y_{ikt} = 1, \boldsymbol{V_{it}}] = exp\left(\frac{\sigma^2_\eta}{2}\right)\frac{\Phi(G_j(V_{ijt}) - \theta_j)}{\Phi(G_j(V_{ijt}))} \tag{1.19}$$

where $\theta_j \equiv \rho_j \sigma_\eta$. If $\theta_j = 0$, η_i and e_{ijt} are independent, then the control function $\frac{\Phi(G_j(V_{ijt})-\theta_j)}{\Phi(G_j(V_{ijt}))}$ equals 1 and is irrelevant. Otherwise, the estimates of $\boldsymbol{\gamma}$ and $\boldsymbol{r}$ would be inconsistent when the control function is not included. A negative θ_j suggests that a higher random effect η_i correlates with a lower e^*_{ijt}, i.e., with unobservable factors that make certifier j more attractive by reducing the value of the best alternative. In other words, a higher degree of non-compliance would correlate with choosing certifier j.

I estimate the model with the control function via PPML, as Egger et al. (2011) do when applying the approach of Terza (1998).[51] I obtain estimates of $G_j(V_{ijt})$ from the standard normal quantiles of the conditional choice probabilities estimated in a version of the demand model (1.5) that replaces expected relative rigor with some predictors of it.[52]

[51]In Egger et al. (2011), selection is binary. These authors can, thus, apply the control function suggested by Terza (1998) without combining it with Lee (1983) as I do.

[52]I estimate a simplified version of the demand model (1.5), but without including expected relative rigor, baseline violations, and prices explicitly. Certifier-market constants capture variation in expected relative rigor across regions, years, and certifiers. Interactions of certifier dummy variables with observed FMU characteristics capture variation in baseline violations.

1.5.1.2 Identification

The previous sections have stated all the required assumptions to identify the rigor parameters r_j and $\boldsymbol{\omega}$. In this section, I provide intuition about the sources of their identification and implied assumptions. Ideally, one would identify r_j by comparing violation reports for pairs of audits, it and $i't'$, conducted by different certifiers, j and 1, but are otherwise equivalent. "Otherwise equivalent" means that they would result in the same number of violation reports if conducted by the same certifier. Similarly, one would identify ω_1 by comparing audits with different distances to the certifiers' headquarters that are otherwise equivalent. The identification of ω_2 would rely on comparing otherwise-equivalent audits with and without accreditation inspection. In practice, one does not need perfectly equivalent pairs for direct comparison but can exploit the model's functional form to use any variation in the degree of equivalence.

The challenge is that I do not observe but need to predict the degree of equivalence. The key assumptions to identify certifiers' relative rigor are, thus, exogeneity of (i) the certifier identity, (ii) the market's distance to certifiers' headquarters, and (iii) accreditation inspections to unobserved determinants of violation reports. That is exogeneity to unobserved factors that render two audits nonequivalent. That exogeneity assumption is needed to interpret expected violation reports as expected changes in violation corrections and, consequently, quality in the counterfactual analysis. Apart from that, the paper's main results rather depend on identifying *expected* relative rigor.[53] Its identification only requires exogeneity to unobserved determinants of violation reports which FMUs *expect*. In the rest of the section, I add references to that type of exogeneity in brackets.

Exogeneity of the certifier identity requires that the rates at which the determinants of violation reports $\boldsymbol{f_{it}}$ affect rigor do not differ systematically across certifiers (in FMUs' expectation). That is, their effect on (expected) violation reports, $\boldsymbol{\gamma}$, must be certifier-invariant. Apart from threats in this regard, the certifier identity would only be endogenous if the observables $\boldsymbol{f^v_{it}}$ did not capture all joint determinants of FMUs' certifier demand and compliance and if the control function was not well identified at the same time. The control function can be identified based on parametric form assumptions or exclusion restrictions, i.e., determinants of demand that do not affect violation reports. The certifier-market-level constants δ_{jt} in the demand model capture variation across markets that cannot stem only

[53] Expected relative rigor drives participation and certifier choices and, thus, the willingness to pay for leniency and changes in choices in the counterfactual simulations of stricter accreditation.

from determinants of violation reports. This variation, thus, satisfies an exclusion restriction. However, within markets, all observed determinants of individual FMUs' demand may also affect violation reports. Therefore, I need to rely on the distributional form assumptions to control for potential within-market selection on unobservables.

Regarding the exogeneity of markets' **distance to certifiers' headquarters**, there is only one negligible concern: Most certifiers have their headquarters in Europe, where (expected) compliance tends to be higher than in other continents. Observable controls account for this across-market variation in compliance.

Accreditation inspections are assigned based on indicators of certifiers' leniency and FMUs' compliance, market shares of certifiers, practical considerations such as the location of assessors, and some random variation.[54] The model controls for many determinants of leniency and compliance and for potential correlation of market shares with expected violation reports, as I discuss in Appendix 1.10.1.3. Conditional exogeneity of the (expected) inspection assignments requires that the accreditation body does not assign inspections based on additional information about leniency or compliance (if FMUs anticipate that). Remaining variation in market shares, practical considerations, and random variation allow then to identify ω_2.

1.5.1.3 Results

Table 1.5 presents the estimates of the main parameters. Columns (1) and (2) show the results of the estimation without and with control functions, respectively. Most of the coefficients θ_j from the control functions are statistically different from zero. As outlined above, this finding suggests that the estimates from Column (1) are inconsistent. In the following, I, thus, interpret the results from Column (2). I include standard errors that are robust to heteroskedasticity and serial correlation.[55]

Rigor types r_j differ significantly, both statistically and economically, across certifiers. Figure 1.3 (a) ranks certifiers by plotting the estimates of r_j with 95% confidence intervals. The lowest ranked certifier, Certifier 7, reports only 56% of the number of violations the baseline certifier reports, everything else equal. Apart from the small certifiers, the best certifier is expected to report 54% more than the baseline certifier. Two facts support the identification

[54] This information is based on ASI (2021) and an interview with the accreditation body's staff.

[55] As a next step, I will bootstrap standard errors to account for the sampling variation in the control functions as well.

Table 1.5: No. of violation reports (by year)

	Without control function (1)	*With control function* (2)
Log(r_j), rigor type:		
Certifier 2	−0.291 (0.246)	−0.442*** (0.135)
Certifier 3	0.350* (0.201)	0.241*** (0.081)
Certifier 4	−0.053 (0.160)	−0.076 (0.112)
Certifier 5	−0.509*** (0.172)	−0.503*** (0.075)
Certifier 6	0.531*** (0.140)	0.429*** (0.054)
Certifier 7	−0.875*** (0.226)	−0.584*** (0.125)
Small certifiers	−0.165 (0.246)	0.024 (0.168)
$\boldsymbol{\omega}$*:*		
Audit inspected by accreditation body	0.776*** (0.156)	0.793*** (0.054)
Average distance to certifier's headquarter by market in 1000 km	−0.034** (0.017)	−0.033*** (0.007)
θ_j from control functions:		
Certifier 1		−0.061 (0.040)
Certifier 2		−0.256*** (0.091)
Certifier 3		−0.222*** (0.057)
Certifier 4		−0.083 (0.064)
Certifier 5		−0.080 (0.054)
Certifier 6		−0.228*** (0.033)
Certifier 7		0.545*** (0.174)
Small certifiers		0.270 (0.219)
Controls $\boldsymbol{f_{it}}$	Yes	Yes
Observations (choice situations)	3,810	3,810

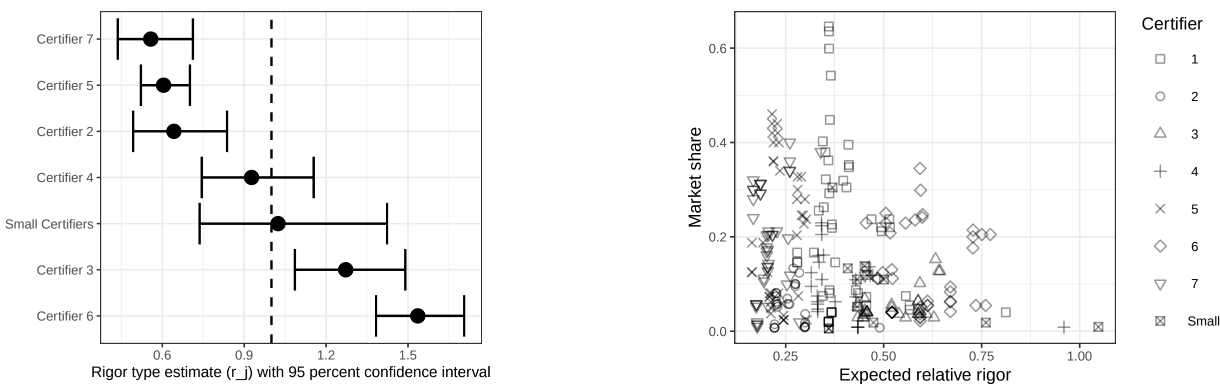
Notes: Standard errors are robust to heteroskedasticity and serial correlation. They do not account for the variance of the generated regressors used as the control function, yet.$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

of the most lenient certifiers: suggested by the estimates: First, the only two certifiers that FSC's accreditation body partially suspended from forest management certification in 2011-2020 appear among the three lowest ranked rigor types. Second, NGOs such as Earthsight (2020) have reported Certifier 7's failure to find serious violations such as illegal logging.

Why should certifiers have different rigor types? Different ownership types and intrinsic motivations of managers or staff might explain such variation. However, different rigor types could appear and survive even if the rigor type was a strategic long-run choice of for-profit certifiers without intrinsic motivation. The costs and benefits from increasing rigor may vary across certifiers. Costs may include increased audit costs and reduced revenues if FMUs prefer leniency. Benefits arise from a reduced likelihood of losing accreditation and market access. Classic reputation theory suggests that larger certifiers with higher profits have reason to be more concerned about losing market access and may, thus, be more rigorous than smaller certifiers (Klein and Leffler, 1981). However, such a pattern is not found here. Certifiers 2 and 7 are very large certifiers both in FSC and other markets. They are the only two publicly traded FSC certifiers. The overall correlation of certification companies' total revenue in

2020 with the relative rigor estimates is negative and statistically significant. Given the low number of certifiers, this finding cannot be interpreted as a pattern in the opposite direction, but, rather, as motivation for further research. An effect of rigor on the likelihood of accreditation suspensions can still rationalize the heterogeneity of rigor if certifiers differ in the degrees to which they discount future profits. Appendix 1.10.1.5 discusses all those potential explanations of rigor heterogeneity in more detail.

Figure 1.3: Relative rigor predictions



(a) Rigor type

(b) Expected relative rigor and market shares by certifier (The darker a shape, the more observations it covers.)

Rigor also varies across accreditation inspections and distance to the certifier's headquarters. The estimates suggest that certifiers report on average 121% more violations in audits with accreditation inspections than in equivalent audits without inspections. Potential explanations for this are moral hazard in uninspected audits or extra rigor in inspected audits due to a behavioral effect, also called Hawthorne effect (Landsberger, 1958). On average, certifiers report 29% fewer violations when they audit 10,000 km farther away from their headquarters, potentially due to reduced quality control by headquarters.

Nevertheless, the certifier-invariant predictors $\boldsymbol{f}_{it}^{v}$ explain much more of the variation in violation reports than differences in rigor. Compared to using the mean number of violation reports, controlling for $\boldsymbol{f}_{it}^{v}$ reduces the cross-validated prediction error by 11%, while the certifier variables r_j and $\boldsymbol{x}_{ijt}^{v}$ only reduce it by 2 additional percentage points (and 2.7% if one excludes the predictors $\boldsymbol{f}_{it}^{v}$). The in-sample R-Squared, adapted for the count data model following Cameron and Windmeijer (1996), is 0.28 for the full model. That suggests that the included variables are predictive of violation reports, but that a substantial share of the variation remains unexplained. Nevertheless, recall that the main purpose of the model is not to predict realized violation reports, but FMUs' expectations of relative rigor and violation reports (next to a rigor ranking that allows to simulate the suspension of the most lenient certifiers).

Table 1.6: Summary statistics of the predictions about violation reporting

Statistic	N	Median	Mean	St. Dev.	Min	Max
Expected violation reports	34,794	0.40	0.61	0.73	0.02	17.23
Expected relative rigor $ExpectedRigor_{jt}$	34,794	0.36	0.39	0.16	0.16	1.05
Expected baseline violations $\mathbb{E}[v^0_{i1t}\|\boldsymbol{f}^{\boldsymbol{v}}_{\boldsymbol{it}}]$	34,794	1.08	1.64	1.81	0.09	25.71

Table 1.6 provides summary statistics of the model's predictions of FMUs' expectations, namely the expected relative rigor $ExpectedRigor_{jt}$, expected baseline violations $\mathbb{E}[v^0_{i1t}|\boldsymbol{f}^{\boldsymbol{v}}_{\boldsymbol{it}}]$ and their product, expected violation reports. Unsurprisingly, expected violation reports have about the same mean, but a slightly smaller range than realized violation reports. The predictions of expected relative rigor suggest that FMUs expect certifiers to report only about 40% of the baseline violations, on average. But that share ranges from 16% to 105%. Figure 1.3 (b) plots those shares against the market share of each certifier in each market. The plot shows substantial variation in expected relative rigor across certifiers and across markets within certifiers. In the next section, I will use the variation across markets to identify FMUs' dislike of rigor.

1.5.2 Stage 2: Demand

1.5.2.1 Estimation and identification

I estimate the demand model in two steps. Following Berry (1994), I split the utility into certifier-market-level utility, δ_{jt}, and within-market variation, $\tilde{V}_{ijt}$:[56]

$$\begin{aligned}
u_{ijt} =& \delta_{jt} + \tilde{V}_{ijt} + \epsilon^u_{ijt} && (1.20)\\
\delta_{jt} =& \bar{\alpha} p_{jt} + \bar{\beta}^r ExpectedRigor_{jt} + \boldsymbol{x}^{\boldsymbol{u}\prime}_{\boldsymbol{jt}} \boldsymbol{\beta}^{\boldsymbol{x}} + \xi_j + \xi_t + \Delta\xi_{jt} && (1.21)\\
\tilde{V}_{ijt} =& \tilde{\alpha} c_{it} p_{jt} + \\
& \left(\tilde{\beta}^r_1 \mathbb{E}[v^0_{i1t}|\boldsymbol{f}^{\boldsymbol{v}}_{\boldsymbol{it}}] + \tilde{\beta}^r_2 \mathbf{1}\left\{ \max\left(\mathbb{E}[v^0_{i1t}|\boldsymbol{f}^{\boldsymbol{v}}_{\boldsymbol{it}}] ExpectedRigor_{jt} \right)_{it} \geq 5 \right\} \right) ExpectedRigor_{jt} + \\
& \boldsymbol{d}'_{\boldsymbol{ijt}} \boldsymbol{\beta}^{\boldsymbol{d}} + \boldsymbol{f}'_{\boldsymbol{it}} \boldsymbol{\xi}_{\boldsymbol{j}} && (1.22)
\end{aligned}$$

[56] This approach seems more reliable than a control function approach, given the difficulties with the compatibility of the latter and a supply model based on profit maximization as discussed in Train (2009).

First, I estimate the **nested logit** model (1.20) with constants δ_{jt} and the determinants of $\tilde{V}_{ijt}$ from (1.22). I maximize the following log-likelihood:

$$\mathcal{L}^u(\mathbf{\Lambda}) = \sum_{t=1}^{T} \sum_{i=1}^{N_t} \sum_{j=1}^{J_T} y_{ijt} \log s_{ijt}(\mathbf{\Lambda}) \tag{1.23}$$

where $\mathbf{\Lambda}$ is the vector of parameters $\tilde{\alpha}, \tilde{\boldsymbol{\beta}}^{\boldsymbol{r}}, \boldsymbol{\beta}^{\boldsymbol{d}}, \boldsymbol{\xi}_{\boldsymbol{j}}, \boldsymbol{\delta}, \lambda$ and $s_{ijt}(\mathbf{\Lambda})$ the nested logit choice probability defined in equation (1.7).

Second, I estimate model (1.21) in a pooled **two-stage least squares (2SLS)** regression. Prices are likely endogenous since the structural error, $\Delta\xi_{jt}$, captures unobserved popularity and may thus affect market shares and hence the markup. Expected relative rigor may also be endogenous since market shares can affect the likelihood of accreditation inspections $\bar{x}_{2jt}$ and since accreditation inspections increase expected relative rigor. Following Gandhi and Houde (2019), I use differentiation instruments to account for the endogeneity of both variables; namely,

1. The number of close-by rivals, $\sum_{k \neq j} 1(|x_{kt} - x_{jt}| < \kappa)^2$, where κ is the standard deviation of x_{kt} across all markets
2. The interaction of that instrument with some mean demographic by market, $\bar{f}_t$

Such instruments are exogeneous for those characteristics x_{jt} for which the utility model controls. Any effect of differentiation in such characteristics on market shares goes through the differences in FMUs' utility levels across certifiers. In the main specification, I construct such instruments for certifiers' experience as x_{jt}, i.e., the number of years they have been FSC-certifiers. The model's combination of certifier and market fixed effects perfectly captures certifiers' experience. I use the average longitude in the market as the mean demographic $\bar{f}_t$.

The instruments are relevant since certifiers with more close-by rivals tend to have lower market shares, everything else equal. The demographic accounts for variation in the effect of close-by rivals on market shares across markets. For prices, lower market shares lead to lower markups. The effect of market shares on expected relative rigor goes through the likelihood of accreditation inspections. The accreditation body generally tends to inspect larger certifiers more often than smaller ones. In some cases, increased competition can also lead to more frequent inspections, especially for relatively new certifiers in a market.

There is one potential concern with the abovementioned phenomenon: The accreditation body may inspect more in markets with more competition because they expect rigor to decrease with increased competition. If this were the case, I might misinterpret the coefficient on expected relative rigor since I do not control for variation in rigor within certifiers across markets. I can test for this by examining the association of the instruments with the variation in violation reports that the predictor variables used above do not explain. To do so, I regress the residual of violation reports $\hat{\epsilon}^{v}_{ijt}$ from model (1.3) on the differentiation instruments. The coefficients on all instruments are far from significant. Hence, variation in unmeasured rigor does not seem to drive the variation induced the instruments by in accreditation inspections and $ExpectedRigor_{jt}$. This is true for the rigor that affects the number of violation reports. Instead, the accreditation body may inspect more in cases of increased competition as they can rely less on historical information about auditors in such contexts. This explanation supports the assumption of instrument exogeneity.

Identification of the price and rigor coefficients, $\bar{\alpha}$ and $\bar{\beta}^r$, relies on variation in prices and rigor across markets within each certifier, following different patterns for different certifiers. Identification requires such variation since fixed effects absorb across-certifier variation that is the same in all markets and across-market variation that is the same for all certifiers. Certifier fixed effects are important since they control for unobserved differences in certifiers' general expertise and efficiency which might correlate with rigor. Market fixed effects matter as FMUs' net benefits of FSC certification are probably sensitive to market-specific changes in demand for certified wood. Replacing market fixed effects with separate region and year fixed effects may omit some of these changes but may allow for the use of additional variation for the identification of $\bar{\alpha}$ and $\bar{\beta}^r$. Appendix-Table 1.26 shows that such replacement would not affect the results much.

The identifying variation for the rigor coefficient $\bar{\beta}^r$, stems mainly from the variation in the likelihood of accreditation inspections since the 2SLS-model (1.21) controls for the distance to certifiers' headquarters. Appendix-Figure 1.24 shows differential variation in the likelihood of accreditation inspections across markets within certifiers. The choice of relevant instruments ensures that the instruments predict sufficient exogenous variation in that likelihood. Appendix 1.10.2.3 presents tests for that.

FMUs' net benefits from participation in FSC certification are identified based on the variation in the timing of FMUs' decisions to join or leave FSC. The identification exploits variation across markets with different choice sets of certifiers and different certifier characteristics. In addition, it uses variation across FMUs with different characteristics within the same market to account for heterogeneity in their benefits.

1.5.2.2 Results

Table 1.7 presents the parameter estimates from the certifier-market-level utility model (1.21). Column (1) presents the OLS results. The 2SLS results from Column (2) are the main specification since I reject the hypothesis of consistent OLS results based on the Wu-Hausman test. Table 1.8 presents the parameter estimates that capture heterogeneity in FMUs' preferences within markets from the nested logit model (1.22). In both tables, I report heteroskedasticity and serial correlation robust standard errors.[57] I present reassuring tests of instruments' relevance in Appendix 1.10.2.3. Tables 1.23 and 1.24 in the Appendix show the coefficient estimates from the first stages.

Table 1.7: Selected preference estimates at the certifier-market-level

	OLS	*2SLS*
	(1)	(2)
$\bar{\alpha}$: Price at certifier-market-level in 1K USD	-0.069^{***} (0.003)	-0.770^{***} (0.178)
$\bar{\beta}^r$: Expected relative rigor	-1.367^{***} (0.046)	-17.366^{***} (1.132)
$\boldsymbol{\beta}^x$:		
First year certifier is available	0.294^{***} (0.019)	1.698^{***} (0.195)
Average distance to certifier's headquarter by market in 1000 km	-0.025^{***} (0.001)	-0.139^{***} (0.016)
Certifier FE	Yes	Yes
Market FE	Yes	Yes
Observations	34,794	34,794
Adjusted R^2	0.649	-1.006
Residual Std. Error (df = 34733)	0.484	1.156
F Statistic	$1{,}072.535^{***}$ (df = 60; 34733)	
Wu-Hausman stat.		723.996^{***} (df=2,34731)
Weak IV stat. (Expected relative rigor)		343.283^{***} (df=2,34733)
Weak IV stat. (Price at certifier-market-level in 1K USD)		77.239^{***} (df=2,34733)

Notes: Heteroscedasticity robust standard errors. Standard errors are not yet corrected for the use of generated regressors and regressands. The dependent variable are the estimates of mean utility by certifier and market, obtained from MLE of the certifier-market constants in the nested logit model of FMUs' certifier choice. The outside option in the nested logit includes both dropping out of FSC certification or not getting certified yet. Dropping out is permanent in the data, such that only the first year after drop-out is included. In the 2SLS regression here, the outside option is excluded since its mean utility is normalized to zero for all markets in the nested logit.$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

The main result is as expected: the estimate of $\bar{\beta}^r$, the coefficient on certifiers' expected relative rigor, is negative. This sign suggests that, all else being equal, FMUs prefer a certifier whom they expect to report fewer violations than other certifiers. That is, they prefer a more lenient certifier. The coefficients $\tilde{\boldsymbol{\beta}}^r$ capture heterogeneity in this preference across FMUs with different levels of compliance. Surprisingly, the point estimates suggests that FMUs with more baseline violations, a proxy for lower compliance, dislike rigor slightly less. This seems to be even more true for FMUs that can expect the most rigorous certifier to find at

[57]The standard errors do not, yet, account for the use of generated regressors. The significance levels are, thus, preliminary and will be updated based on bootstrapping.

Table 1.8: Selected estimates of heterogeneity in preferences across FMU types (Nested Logit)

$\tilde{\alpha}$: Price at certifier-market-level in 1K USD X ...	
Cost factor	0.020*** (0.008)
$\tilde{\boldsymbol{\beta}}^r$: Expected relative rigor X ...	
Expected baseline violations	0.082 (0.097)
≥ 5 violations reports expected from most rigorous certifier	0.519 (1.290)
Forest chose same certifier last year	2.496*** (0.345)
Forest chose same certifier last year x Yrs. with FSC cert.	0.051*** (0.014)
Forest chose same certifier last year x Recertification year x FSC	−0.076 (0.166)
FSC Certifier has office in forest's country	0.549*** (0.141)
Is first FSC certifier (entry cost)	−2.457*** (0.281)
λ:	
Within FSC nest correlation	0.464*** (0.063)
Certifier-market FE	Yes
Group cert. indicator x Certifier FE	Yes
Plantation indicator x Certifier FE	Yes
FMU characteristics x FSC indicator	Yes
Observations (choice situations)	6,250
Log Likelihood	−2,744.536

Notes: MLE of nested logit choice model with R package mlogit. hteteroscedasticity and serial correlation robust standard errors. Standard errors are not yet corrected for the use of generated regressors. The outside option includes both dropping out of FSC certification or not getting certified yet. Dropping out is permanent in the data, such that only the first year after drop-out is included. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

least five violations, i.e., that risk losing their certificates. Yet, these two coefficients are not significant. In particular, the coefficient on the indicator for more than five expected violation reports is very noisy, given the small number of observations in such a situation.

The difference between Columns (1) and (2) in Table 1.7 suggests that the coefficients on prices and expected relative rigor are underestimated in absolute terms if I do not account for their endogeneity. Such underestimation is expected: larger demand shocks at the certifier-market-level should correlate with higher markups, and thus, the OLS results underestimate FMUs' dislike of higher markups. Larger demand shocks should also correlate with more frequent accreditation inspections, which are associated with more violation reporting. Thus, the OLS results underestimate FMUs' dislike of rigor and violation reporting when not accounting for the endogeneity of accreditation inspections.

A few other parameter estimates are worth highlighting. First, the baseline price coefficient $\bar{\alpha}$ is negative, as expected. Second, FMUs with larger cost factors, i.e., larger total prices for their forest area, are less price sensitive. This association may seem counterintuitive. However, this is plausible since FMUs with larger cost factors often have higher revenues, for example, since they tend to have more extensive forests. Third, the net benefit from FSC certification is substantially lower in the first year of participation. This finding reflects a one-time cost of "entering" the certification system. Fourth, FMUs are more likely to stay with their previous certifiers, suggesting the presence of switching costs. Fifth, certifiers are more popular closer to their headquarters and in countries where they have an office, ceteris paribus. Finally, the estimated correlation of taste shocks within the FSC nest $1 - \lambda$ is far from perfect correlation but also substantially different from independence. This finding suggests that certifiers are not perfect substitutes but also not completely different in terms of their yearly unobservable characteristics, captured by the taste shocks.

Table 1.9: Price and rigor elasticity estimates

Statistic	Median	Mean	Min	Max
Price at certifier-market-level	−4	−4	−19	−0
Expected relative rigor	−11	−12	−39	−0

Notes: Percentage point change in conditional choice probabilities per percentage change in certifier-market-level prices and expected relative rigor

Shopping for leniency: in the rest of this section, I analyze the extent of FMUs' preference for leniency and their willingness to pay for it. Table 1.9 presents summary statistics of the elasticities of individual market shares with respect to changes in certifier-market-level prices and expected relative rigor. Recall that expected relative rigor is the portion of baseline violations that FMUs' expect a certifier to report. On average, a 1% increase in this portion is associated with an average decrease in individual market share of 12%. This sizable elasticity is more than three times the average own-price elasticity.

Table 1.10 shows the estimated willingness to pay (WTP) for leniency and other certifier characteristics. My principal measure of the WTP for leniency is the WTP for a certifier that reports a one standard deviation lower fraction of baseline violations than its market competitors, i.e., a certifier that is one standard deviation more lenient than other certifiers in the same market. On average, that WTP is 4123 USD. The average standard deviation in leniency is 0.2, i.e., a 20-percentage-point difference in the reported fraction of baseline violations. The WTP for a certifier that reports one violation less is accordingly higher. However, as I discuss in the next section, the WTP of 4123 USD is already substantial. It

Table 1.10: Estimated willingness to pay

Statistic in 1000 USD for a certifier that ..	Median	Mean	Min	Max
Reports 1 SD lower fraction of violations	4.08	4.12	1.50	10.01
Reports 1 violation less	22.56	34.44	0.99	262.50
Has a 1 SD closer headquarter	0.52	0.52	0.19	1.09
Is last year's certifier	3.46	3.52	3.26	10.06
Has an office in your country	0.76	0.78	0.72	2.21
Skipping the first year with FSC (entry cost)	3.41	3.47	3.21	9.90

Notes: The numbers correspond to the coefficient estimates divided by the price coefficient, multiplied by the within-market standard deviation across certifiers for the numeric variables in rows 1 and 3, and divided by the number of baseline violations in row 2.

is roughly as large as the estimated "entry cost" in the first year of certification and more than half the average predicted certification fee. At the same time, the WTP for other characteristics may limit the degree to which FMUs shop for more lenient certifiers. Most importantly, remaining with the same certifier as in the previous year is of similar value as leniency. This implies that switching to another certifier in years when the previous certifier reports fewer violations due to fewer accreditation inspections would be costly. For most FMUs, their preference for leniency is, thus, relevant primarily for their initial participation and certifier choice or when a suspension of their certifier's accreditation forces them to choose a new one, as the counterfactual analysis will simulate.

1.5.2.3 Discussion

The WTP for leniency captures the net opportunity cost of violation reports for FMUs to the extent that FMUs consider these costs when choosing a certifier. These net costs may include economic private net costs as well as societal benefits if FMUs care about these benefits.

Is the WTP a reasonable estimate of violation reports' economic private net costs? There is no estimate of these costs in the literature so far. However, the discussion in Section 1.2.3 suggests that the economic private cost of violation correction is of the same order of magnitude as the estimated WTP for a certifier that reports one less violation.[58]

How does the perceived cost of violation reports relate to benefits from FSC certification?

[58] For example, if 60% of violations induce a cost of 1000 USD, 30% a cost of 10,000 USD, and 10% a cost of 100,000 USD, then the average cost is 13600 USD. Such a distribution corresponds roughly to the qualified guess of violation correction cost in a small, randomly drawn sample, as described in Section 1.2.3.

Table 1.11: Estimated benefits and costs of FSC certification

Statistic in 1000 USD	Median	Mean	Min	Max
Net benefits, gross of fees and violation correction cost	20.80	20.57	−2.01	41.13
Violation correction cost	8.52	9.29	3.73	25.59
Value of certification fee	3.14	3.24	1.27	11.67
Net benefit, net of fees and violation correction cost	9.03	8.04	−10.73	20.01
Potential additional revenues due to FSC	255.79	1,887.37	0.08	93,122.94

Notes: The numbers are computed according to the model estimates for FMUs' chosen FSC-certifier. The cost of violation correction are the WTP for one less violation report multiplied by the number of violations reported by the chosen certifier. The value of the certification fee is the market-level certification fee since the cost factor is part of the price coefficient. I compute the potential additional revenues by assuming a price premium of 5%, half the median premium from choice experiments, the average price per cubic meter of roundwood imports in the member states of the UNECE in 2017, and the average number of cubic meters produced in FSC-certified forests per hectare in 2017.

Table 1.11 summarizes the model's predictions of the benefits and costs of FSC certification, focusing on the chosen certifier for each FMU. The first row shows the net benefits if there were neither fees nor expected violation reports. The next row presents the opportunity cost of the expected violation reports by the chosen certifier, computed based on the WTP for leniency. At the median, this cost corresponds to almost half of the net benefits in the first row. The fourth row of Table 1.11 shows the net benefits after subtracting the violation correction cost and the less substantial value of the certification fee (in the third row) from the values in the first row. The remaining net benefits are only 9,000 USD at the median, with many negative values in the distribution. This finding suggests that many FMUs are very sensitive to changes in available rigor levels when deciding whether to participate in FSC certification. The counterfactual analysis investigates this sensitivity further.

How do these net benefits and the WTP for leniency compare to the gross benefits from certification, such as additional revenues? The parameter estimates do not allow me to compute the increase in revenue or other gross benefits through FSC certification since most coefficients can capture both cost and benefits. The last row of Table 1.11 summarizes back-of-the-envelope estimates of additional revenue from FSC certification based on the median willingness to pay a premium, as reported in choice experiments in the 2000s (Cai and Aguilar, 2013). At the median, the additional revenue estimates exceed the net benefits by a factor of more than one hundred. What can explain this gap? On the one hand, the additional revenue estimates are likely an upper bound. The willingness to pay a premium stated in choice experiments often exaggerates a consumer's willingness to pay during purchases. Anecdotal evidence further suggests that price premia for certified wood have decreased since the choice experiments covered by Cai and Aguilar (2013). On the other hand, parameters other than the violation correction cost can account for those certification costs not captured by the estimated violation correction cost and certification fees. The estimated violation correction cost is based on the WTP for leniency in terms of the number of violation reports, while other dimensions of leniency can also impact FMUs' cost. For example, some certifiers may report violations that are less expensive to correct than others. In addition, it matters how easily certifiers accept corrections of violations in follow-up audits, information that is not included in the data thus far.

1.5.3 Stage 1: Pricing

I use the estimated demand parameters to compute certifiers' markups and marginal cost computed according to the price model. I do so to account for price effects in the counterfactuals of stricter accreditation. To compute markups, I replace c_{jt} in derivative (1.10) with $\frac{p_{jt}}{k_{jt}}$. The profit-maximizing markup factor is then

$$k_{jt} = \frac{\frac{\bar{\alpha}+\tilde{\alpha}\overline{\tilde{c}_{it}}}{\lambda} p_{jt} \left[1 - \lambda s_{jt} + (\lambda - 1) s_{jt|FSC}\right]}{\frac{\bar{\alpha}+\tilde{\alpha}\overline{\tilde{c}_{it}}}{\lambda} p_{jt} \left[1 - \lambda s_{jt} + (\lambda - 1) s_{jt|FSC}\right] + 1}$$

where I replace the coefficients and variables with the predictions from the model. Table 1.12 presents the markup and marginal cost estimates. Mark-ups constitute 33% of the market price on average, with most observations being below that level.

Table 1.12: Mark-up and marginal cost estimates

Statistic	N	Median	Mean	Min	Max
Prices (1000 USD)	34,794	7.49	9.44	0.31	128.29
Marginal cost (1000 USD)	34,794	6.22	8.08	0.21	110.38
Mark-ups (percentage of cost)	34,794	30.00	33.00	6.79	135.53

1.6 Counterfactual Analysis

I use the estimated model to implement two sets of counterfactual exercises. They investigate whether increasing the minimum level of certifiers' rigor can increase the number of violation reports, hence, corrections of violations and, thus, quality among all once-certified FMUs. In particular, I analyze to what extent effects on participation counteract positive effects on quality among certified FMUs. I also attempt to quantify welfare effects.

In the first set of counterfactuals, I consider direct shifts in the global minimum level of rigor. This exercise explores the effects label owners such as FSC might achieve if they had direct control over minimum rigor. It also allows to illustrate how effects can develop from very small to very large increases of minimum rigor. Finally, this exercise quantifies how costly rigor increases are for certifiers due to FMUs' willingness to pay for leniency.

In the second set of counterfactual scenarios, I draw attention to the enforcement of increased minimum rigor. For many label owners, including FSC, credible threats of suspending market access are arguably the only way to control certifiers' rigor. I, therefore, simulate the suspension of lenient certifiers' accreditation. I then isolate the effects of accreditation suspension due to changes in rigor from the effects due to changed choice sets. To do so, I consider equivalent shifts in those certifiers' rigor. Comparison to these isolated effects allows for discussion of some drawbacks of certification systems with external certifiers.

In all cases, I simulate changes in quality, participation, and agents' surplus. I account for mechanical effects of increased rigor and indirect effects through changes in certification choices and prices. I keep everything else constant and discuss the implications after presenting the results.

1.6.1 Procedure

For each scenario, I solve for the new market equilibrium under the counterfactual assumptions. I allow the FMUs to adjust their certification choices and the certifiers their markups and prices. Specifically, I carry out the following steps:

First, I implement the counterfactual assumptions for 2019, the last year of my data. The observed market structure in 2019 is the baseline scenario.

1. For the first scenario, I consider 10% increases from 0 to 500% in the minimum level of expected relative rigor across all certifiers and markets, $\min\{rigor_{jt} \forall j \in J, t \in T\}$. I call this level the *global minimum*. For each new global minimum, I increase the expected relative rigor to this level wherever it is below at baseline.
2. I consider three cases for the second scenario: suspension of the most lenient, two most lenient, or three most lenient certifiers' accreditation.
    i. I rank certifiers according to their constant rigor type, r_j.[59] I remove the suspended

[59]The ranking according to each certifier's global minimum level of expected relative rigor is the same.

certifiers from FMUs' choice sets J_t.

ii. For each set of suspended certifiers, I also conduct a separate simulation where I do not remove them from the choice sets but shift their expected relative rigor to the level of the next most lenient certifier in each market. The lowest available level of expected relative rigor in each market is then the same as in the case of suspension, but the set of available certifiers is the same as in the baseline scenario.[60]

Second, I solve for the new set of choice probabilities $s_{ijt}(\boldsymbol{k_t'})$ and markups $\boldsymbol{k_t'}$ based on the parameter estimates from Table 1.7, Column (2), and Table 1.8. All other variables and parameters remain constant.[61]

Third, I compute relevant statistics for the baseline and the counterfactuals in 2019.

1. The sum of FMUs' weighted averages of expected violation reports, henceforth referred to as the *unconditional weighted sum*, is a proxy of the aggregate quality changes induced by FSC among all once-certified FMUs. Each FMU's weighted average is the sum of the violation reports the FMU expects from each certifier, weighted by the probability of choosing that certifier. The number of violation reports is zero if the FMU does not participate in certification. The sum of those weighted averages is, thus, $vsum \equiv \sum_{it} \sum_{j \in J_t} \hat{s}_{ijt} \widehat{ExpectedRigor}_{jt} \widehat{E[v_{i1t}^0|\boldsymbol{f_{it}}]}$.
   i. The *conditional* weighted sum $vsum_{FSC}$ is obtained by replacing $\hat{s}_{ijt}$ with $\hat{s}_{ijt|FSC}$, thus, weighting by the probability of choosing each certifier conditional on participating in FSC certification. It is, hence, a proxy of the quality changes induced among certified FMUs, i.e., the quality of certification.
   ii. The weighted sum of expected violations that remain *unreported* due to reductions in participation is the difference between the unconditional and conditional weighted sums.
2. Certifiers' surplus from FSC certification is defined in model (1.8).
3. FMUs' surplus is FMUs' maximum utility in dollar. For the nested logit model (1.5), this is $\frac{1}{\alpha_{it}} \left\{ \ln \left[1 + \left(\sum_{j \in J_t} \exp(V_{ijt}/\lambda) \right)^{\lambda} \right] + C \right\}$ (McFadden, 1977). The unknown constant C

[60] These rigor shifts differ from the shifts in the global minimum analyzed above: the lowest ranked certifiers' leniency varies from market to market, such that their global suspension increases the minimum level of rigor to a different extent in each market. In the case of the suspension of the two and three globally most lenient certifiers, there are even three markets in which those certifiers are predicted to be more rigorous than remaining once, such that the minimum levels of rigor in those markets do not change.

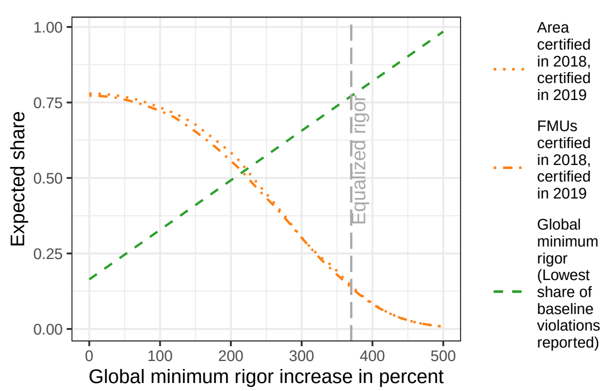
[61] In particular, I do not consider potential effects on rigor above the minimum level, FMUs' true violations and the benefits of certification for FMUs, since the data do not allow me to identify these.

is the utility from the outside option, forestry without FSC certification. The remaining part, thus, constitutes FMUs' surplus from FSC certification.

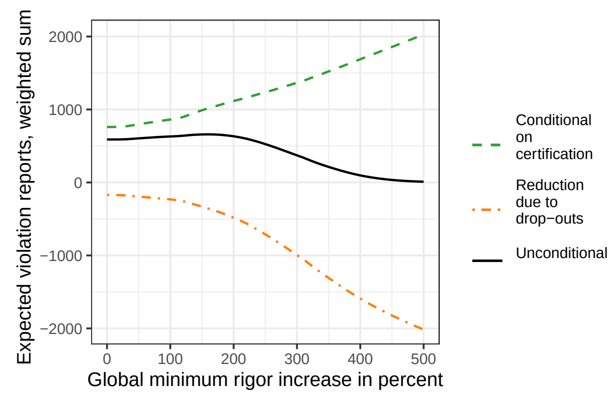
4. A measure of consumer valuation of the counterfactual changes is obtained as follows.
   i. I derive a proxy for consumer valuation of FSC certification per certified hectare at baseline, denoted as CV/ha. To do so, I employ the median stated willingness to pay 10% more for FSC-certified products (Cai and Aguilar, 2013). Assuming that this median value has not decreased until 2017 allows using estimates of wood production per certified hectare, which exist only for 2017. Given those estimates and international wood prices, the estimated consumer valuation per hectare 2017 is approximately 15 USD.[62]
   ii. From that, I deduce $CV/ha/vsum_{FSC}$, the valuation per hectare per unit of certification quality, proxied by the conditional weighted sum of expected violation reports, $vsum_{FSC}$, 0.02 USD in 2017. I assume that this baseline valuation and the production per hectare remain constant until 2019 and across counterfactuals.
   iii. I assume that consumer valuation of FSC certification in the counterfactual is the product of the baseline valuation $CV/ha/vsum_{FSC}$, the changed quality of certification in the counterfactual, $vsum'_{FSC}$, and the changed number of certified hectares, ha': $CV' = (CV/ha/vsum_{FSC}) * vsum'_{FSC} * ha'$. That is, consumer valuation is perfectly elastic to changes in quality.
5. I compute welfare changes as the sum of changes in FMUs' and certifiers' surplus and the changes in consumer valuation. I use consumer valuation as a proxy for social benefits. An underlying assumption is that price premia do not change, such that consumer valuation of the changes translates 1:1 to increases in consumers' surplus, instead of increases in FMUs' surplus.

I discuss the plausibility and implications of all relevant assumptions after presenting the results.

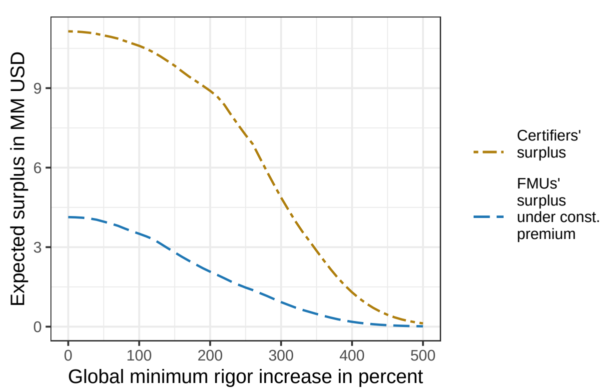
Figure 1.4: Expected changes following direct shifts in global minimum rigor



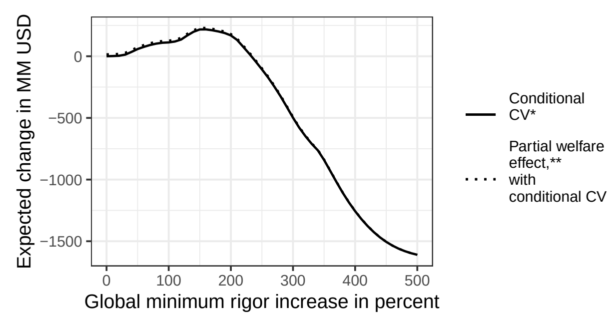
(a) Rigor and participation



(b) Quality, incl. for drop-outs



(c) Surplus from FSC certification



Notes: *Consumer valuation of FSC certification, assuming valuation at 10 percent of wood prices for 2017 baseline quality and constant wood prices. **Sum of CV, FMU and certifier surplus if price premia at baseline CV. Excl. externalities not valued by consumers.

(d) Parametrized change in consumer valuation (CV)

1.6.2 Results and discussion

1.6.2.1 Shifting the minimum level of rigor

Figures 1.4 show how direct shifts of the global minimum rigor and consequent changes in FMUs' certifier choices and prices can affect market outcomes, everything else equal. In each figure, the x-axis represents the enhancement of global minimum rigor in percent of the global minimum rigor at baseline. The global minimum rigor at zero enhancement is the baseline value, the estimate used in the model thus far.

Figure 1.4 (a) shows the trade-off between rigor and participation. The dashed green line represents the new global minimum rigor in each counterfactual. It rises by assumption. The orange lines show two measures of participation. The dot-dashed orange line plots the share of once-certified FMUs participating in 2019. The dotted orange line plots the share in terms of the certified area. Both develop very similarly. At baseline, 23% of once-certified FMUs no longer participated in 2019. Doubling global minimum rigor leads to only 4 percentage points more expected drop-outs. However, participation drops steadily at an increasing rate for moderate to large enhancements in minimum rigor. When minimum rigor more than triples, participation drops very fast at a decreasing rate. Participation approaches zero when the new global minimum exceeds the maximum rigor at baseline. The long-dashed grey vertical line shows this threshold, to the right of which expected relative rigor is equal across all certifiers and markets.

Figure 1.4 (b) shows the predicted effects of the rigor-participation trade-off on quality in the industry. It translates the changes in Figure (a) to those in the weighted sums of expected violation reports. I interpret these numbers as proxies for compliance with the FSC standard and, thus, quality, particularly for reducing negative externalities of wood production. This interpretation is appropriate as long as most reported violations are corrected and wood production does not change significantly apart from these corrections. The dashed green line shows the weighted sum conditional on participation, a proxy for the quality of certification. It rises at an increasing rate with growing minimum rigor. Doubling the global minimum implies an increase of 13%, equalizing rigor across certifiers an increase of 167%. The orange dot-dashed line shows the weighted sum of expected unreported violations due to decreasing

[62] Since the surveys summarized by Cai and Aguilar (2013) ask consumers about their willingness to pay a premium for certified final goods, I also need to assume that the valuation of certification for wood products is proportional to the valuation of certification for rough wood. For wood prices, I use a weighted average of 2017 prices per cubic meter of rough wood from UNECE countries (UNECE and FAO, 2023). I compute the average production of cubic meters per hectare from FSC estimates for 2017 (FSC, 2018b).

participation. This number drops more slowly than the quality of certification increases, until minimum rigor increases by 160%. At this point, the unconditional weighted sum of expected violation reports reaches its maximum, as the solid black line shows. Beyond this, the unconditional effect decreases and becomes negative for extreme changes in global minimum rigor. But at the optimum, the simulation suggests 70 additional expected violation reports, an increase of 12%. This suggests substantial scope for improving quality in the industry by increasing minimum rigor, even if consumers' willingness to pay remains constant.

However, Figure 1.4 (c) shows that improving quality is costly. The dashed blue and the dot-dashed orange lines show FMUs' and certifiers' surplus, respectively. They are similar in shape as participation in Figure (a). My estimated model suggests that certifiers' surplus from FSC forest management certification is approximately 11 million USD at baseline and decreases by 5% if global minimum rigor doubles. If rigor is equalized across certifiers, the profit loss is 81%, according to the model's predictions. Reductions in FMUs' surplus are smaller in absolute numbers but larger in the percentage of their surplus at baseline.

Can the improvements in quality for moderate increases of global minimum rigor justify these costs? The dotted line in Figure 1.4 (d) plots consumer valuation of the counterfactual changes, following the assumptions outlined above, in Section 1.6.1. The solid line shows the sum of this valuation and the changes in FMUs' and certifiers' surplus. It suggests that consumer valuation of quality improvements alone may generate positive welfare effects of increasing global minimum rigor, even for substantial increases. The predicted changes in consumer valuation are so much larger than the effects on FMUs' and certifiers' surplus that the dotted and solid lines almost overlap.

Discussion: The exercise assumes that price premia for FSC certification remain constant. If price premia increased, they would compensate for part of FMUs' loss and reduce drop-outs. That reduction in drop-outs, in turn, would reduce certifiers' losses and increase welfare further. However, zero or little changes in price premia are a realistic scenario, particularly in the short run. Less than 50% of consumers in 33 countries have ever noticed the FSC label (FSC and IPSOS, 2023). A lot fewer consumers will have information about the quality of FSC certification and even fewer will perceive any improvements in this quality. My measure of changes in consumer valuation is not supposed to represent changes in perceived utility. Rather, the idea is as follows: the median consumer valuation at baseline is a proxy of consumers' true utility given the quality of certification at baseline and this true utility changes with the changing quality of certification, even if these changes are not perceived.

The plotted welfare effects must, nevertheless, be interpreted with caution. First, the predictions capture the effects in the regions that this paper focuses on. Second, I assume that consumer valuation is a linear function of the amount of certified wood and the weighted sum of violation reports conditional on certification. In Appendix 1.11.1, I discuss this assumption and consider an alternative measure of consumer valuation, generating similar qualitative results. Third, the social benefit estimates rely on consumers' stated valuation of FSC certification which may not account for all of FSC's positive externalities. Accounting for externalities that consumers do not value may amplify the predicted changes in welfare in absolute terms. Appendix 1.11.2 derives a rough, partial measures of the welfare benefit of reduced tree cover loss through FSC certification, suggesting similar benefits from violation reporting as my computation of consumer valuation. But a lack of research prevents me from deriving a comprehensive measure of the overall benefits of violation reporting.[63]

In any case, label owners cannot fully monetize and redistribute the welfare benefits to incentivize certifiers to increase rigor through monetary compensation as long as there is no fully functioning market for the externalities. Therefore, label owners and their accreditation bodies threaten to suspend lenient certifiers' market access to increase minimum rigor. The following section investigates the effects of such accreditation suspensions.

1.6.2.2 Suspending the accreditation of lenient certifiers

This section investigates the effects of an indirect shift in certifiers' expected relative rigor through the global suspension of the most lenient certifiers. For each set of suspended certifiers, I also consider an equivalent, direct shift of those certifiers' rigor to the next most lenient certifier's rigor *in each market*.

Table 1.13 shows the effects on rigor and participation. The first two columns present the induced changes in the average minimum rigor level across markets, in numbers and percent. Each additional certifier suspension induces an average increase in market-level minimum rigor of roughly 10%. The last two columns show the predicted effects on participation. As in the first set of counterfactuals, there is a clear trade-off between rigor and participation.

The reduction in participation is larger with a suspension than it would be if FSC could shift a certifier's rigor without a suspension: I predict that the suspension of a certifier causes

[63]The benefit from reduced tree cover loss does not account for benefits from more frequent types of violation reports, such as the protection of workers and biodiversity. There is neither sufficient research about FSC's effect on such protection nor are the credible estimates of the social cost of biodiversity loss.

Table 1.13: Simulated changes in audit quality and participation

Counterfactual	Avg. minimum rigor across markets (mechanical change)		Participating FMUs (expected change)	
	in numbers	in percent	in numbers	in percent
(1) Targeting the most lenient certifier				
Accreditation suspension	0.04	9.73	-57.16	-5.92
Equivalent minimum rigor shift	0.04	9.73	-11.13	-1.15
(2) Targeting the two most lenient certifiers				
Accreditation suspension	0.08	20.51	-124.12	-12.85
Equivalent minimum rigor shift	0.08	20.51	-57.21	-5.92
(3) Targeting the three most lenient certifiers				
Accreditation suspension	0.11	29.25	-145.86	-15.10
Equivalent minimum rigor shift	0.11	29.25	-103.32	-10.70

Notes: Changes in counterfactual scenarios compared to the baseline. Accreditation suspension is implemented by removing the corresponding certifier from the choiceset. A direct shift of minimum rigor is implemented by shifting the targeted certifiers' expected relative rigor to the next most lenient certifier's rigor in each market.

1.4 to 5 times as many drop-outs as equivalent minimum rigor shifts. This suggests that switching cost and other certifier differences are jointly even more important for FMUs than certifiers' leniency.

Table 1.14: Expected changes in violation reports

Counterfactual	Conditional on certification	Due to drop-outs	Total	
	in numbers	in numbers	in numbers	in percent
(1) Targeting the most lenient certifier				
Accreditation suspension	62.36	-47.58	14.77	2.51
Equivalent minimum rigor shift	18.62	-14.34	4.28	0.73
(2) Targeting the two most lenient certifiers				
Accreditation suspension	120.59	-98.68	21.91	3.72
Equivalent minimum rigor shift	64.65	-45.26	19.39	3.29
(3) Targeting the three most lenient certifiers				
Accreditation suspension	145.39	-116.74	28.65	4.87
Equivalent minimum rigor shift	133.10	-86.56	46.54	7.90

Notes: Changes in counterfactual scenarios compared to the baseline. Accreditation suspension is implemented by removing the corresponding certifier from the choiceset. A direct shift of minimum rigor is implemented by shifting the targeted certifiers' expected relative rigor to the next most lenient certifier's rigor in each market. These numbers are computed under the assumption that consumers value the effect on violation reports conditional on certification.

Table 1.14 presents the predicted changes in the weighted sums of expected violation reports, proxies for quality. The first column shows the changes conditional on participation, and the second shows the reduction due to drop-outs. The latter follows directly from the reduction in participation, discussed in the last paragraph. The last two columns show the total effect in numbers and percent. The total effect is the effect on the unconditional weighted sum of expected violation reports, a proxy for quality among all once-certified FMUs. I predict positive total effects on such quality, even when all three of the most lenient certifiers are suspended.

In all three sets of suspensions, quality conditional on participation increases more than predicted by an equivalent minimum rigor shift. FMUs' preferences for certifier characteristics other than rigor and switching cost can explain that. Removing a certifier from the choice set forces FMUs to transfer to another certifier or drop out. Some FMUs then transfer to a certifier even more rigorous than the next most lenient due to preferences for characteristics other than rigor. In the cases of suspending the most or the two most lenient certifiers, the effect is more important than the amplified reduction in participation. The total effects on quality among all once-certified FMUs due to suspension exceed the quality effect of an equivalent minimum rigor shift. These results suggest that suspensions are not always a necessary evil compared to direct rigor shifts, but the former can sometimes outperform the latter.

Table 1.15: Expected change in surplus from FSC certification

Counterfactual	FMUs in MM USD	Targeted certifiers in MM USD	Untargeted certifiers in MM USD	Consumers in MM USD	Total in MM USD
(1) Targeting the most lenient certifier					
Accreditation suspension	-0.51	-1.53	1.20	23.91	23.06
Equivalent minimum rigor shift	-0.10	-0.21	0.11	16.01	15.81
(2) Targeting the two most lenient certifiers					
Accreditation suspension	-0.92	-2.82	1.88	16.42	14.56
Equivalent minimum rigor shift	-0.47	-1.08	0.47	26.94	25.86
(3) Targeting the three most lenient certifiers					
Accreditation suspension	-1.07	-3.48	2.36	7.96	5.77
Equivalent minimum rigor shift	-0.87	-2.38	1.46	49.14	47.34

Notes: Changes in counterfactual scenarios compared to the baseline. Accreditation suspension is implemented by removing the corresponding certifier from the choiceset. A direct shift of minimum rigor is implemented by shifting the targeted certifiers' expected relative rigor to the next most lenient certifier's rigor in each market.

Finally, Table 1.15 presents the predicted changes in FMUs' and certifiers' surplus and consumer valuation. Unsurprisingly, FMUs and targeted certifiers lose, while the other certifiers benefit from the simulated changes. Consumer valuation greatly exceeds these costs. The predicted effects on welfare are, therefore, positive in all cases. The limitations discussed in the previous section apply here as well. Under the assumptions made, the welfare benefits of suspending the most lenient certifiers may be in the tens of millions of US dollars. Throughout the counterfactuals, FMUs and certifiers lose more through suspensions than through equivalent shifts in minimum rigor, even though untargeted certifiers win more through suspensions. Consumer valuation and total welfare effects are larger for the suspensions of the most lenient certifier than for an equivalent rigor shift, but smaller in case of additional suspensions.

In Appendix 1.11.3, I show the results from equivalent counterfactual exercises that differ only in terms of keeping prices fixed. The results are similar to those discussed above. Price changes due to changes in market power are, thus, only a minor driver of the abovementioned effects. However, price changes reduce the positive quality and welfare effects of suspensions and, albeit to a much smaller extent, of equivalent rigor shifts.

Overall, the presented counterfactuals show that stricter accreditation has the potential to reduce negative production externalities and increase quality and welfare. Reduced participation decreases the positive effect. Suspensions reduce participation, FMUs' and certifiers' surplus more than do direct shifts in minimum rigor, if they were feasible. However, suspensions also have positive side effects and lead to greater welfare benefits in some cases.

1.7 Conclusion

Certification is used to check compliance with quality standards whose implementation is costly and not observed by consumers and investors. If firms maximize profits, basic theory would not predict them to comply more than they need to obtain certification. When all certifiers of a given standard, such as FSC, provide the same signal to the public, firms are expected to prefer more lenient certifiers since leniency reduces compliance costs. The results of this paper confirm this but also show that other factors may mitigate the mechanism. In the context of this paper, for example, a forest unit's certifier choice is very persistent due to high switching costs. That reduces the degree of shopping for leniency. Nevertheless, this paper shows that firms' willingness to pay for leniency provides bad incentives for certifiers that reduce the credibility of certification, unless counterbalanced by strict accreditation.

This phenomenon has not only led to calls for stricter accreditation, but also for reorganizing certification, for example, by randomly assigning certifiers and paying them through a fund

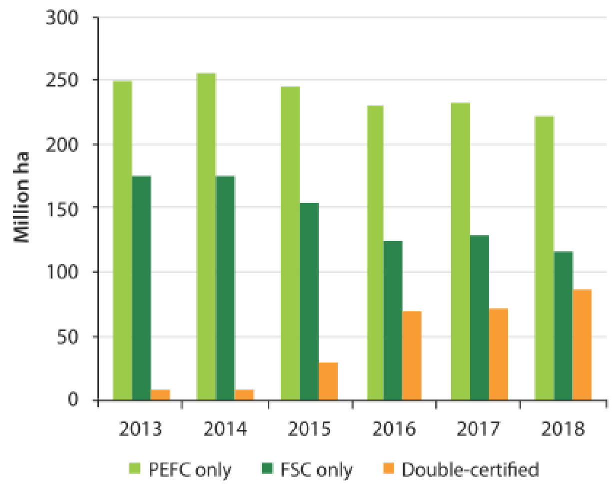
(Earthsight, 2021). Yet, the reason FSC is hesitant to do so might be the risk of increasing forest units' cost beyond their participation constraint. The framework and data of this paper are insufficient to evaluate the effects substantial structural changes such as financing certifiers through a fund would have. But my counterfactual analysis shows that stricter accreditation by suspension of lenient certifiers would indeed reduce participation and the surplus of forest firms and certifiers, as long as price premia do not change. Nevertheless, I predict that the overall effect of moderately stricter accreditation on quality and welfare would be positive.

FSC and any institution regulating voluntary certification face the trade-off between participation and quality of certification. Empirical estimates of demand parameters and measures of leniency are crucial to evaluate the scope for improvement within each context. This paper contributes to that by providing a framework that may be applied in other contexts, where audit results are available. The paper has limitations in the identification due to multiple selection issues and limited price data. A helpful extension would be to model and estimate the drivers of certifiers' choice of rigor in a dynamic context of suspension risk. Such an extension would require a longer time horizon and richer data to estimate region-level rigor. In addition, data on price premia and purchases of certified products by upstream firms and consumers could allow to estimate how benefits of FSC certification for FMUs vary with changes in the quality of certification and demand-side parameters.

Understanding the trade-off between participation and quality is also crucial for policy makers that want to take complementarities of voluntary certification and mandatory regulation into account. This paper shows that while there is scope for improvement in case of FSC, the certification scheme could not afford very large increases in audit rigor. To reduce negative production externalities beyond those limits, price premia would need to increase substantially or corresponding rules would need to be mandatory.

1.8 Appendix A: Institutional setting

Figure 1.5: FSC and the PEFC certified forests areas, 2013-2018 (UNECE/FAO 2019)



Note: Figure 1.5 shows that an increasing share of forests is certified both according to Programme for the Endorsement of Forest Certification (PEFC) and FSC standards. This phenomenon may be explained by the demand from different downstream firms that buy from those forests and may have focused their communication on one of the two standards. That only affects the analysis if forest units choose their certifier jointly for FSC and PEFC. But even then it should not affect the main results, as certifiers are probably of a similar rigor type in PEFC and FSC certification.

1.9 Appendix B: Data and descriptive evidence

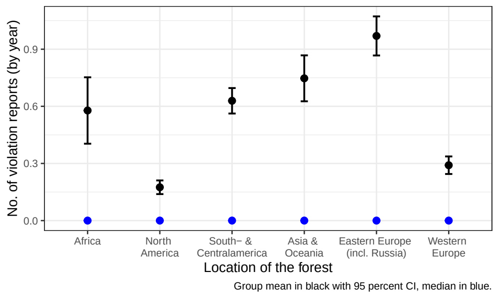
1.9.1 FMU panel

1.9.1.1 Audit data extraction

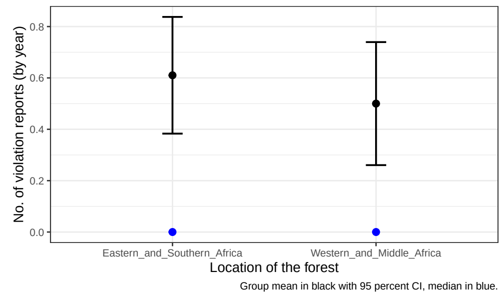
In this section, I describe the extraction of the full audit data collected for the FMU panel, i.e. without restriction to the regions that the analysis focuses on. The full dataset contains basic information on the audits conducted in 2015-2020 with FMUs whose Forest Management / Chain-of-Custody certificate was valid at some point between 2015 and 2020, to the extent that they are documented in the audit summaries published on the certificate profiles accessible through https://info.fsc.org/certificate.php, as demanded in FSC-STD-20-007b (V1-0) EN.

Scraping of the raw data

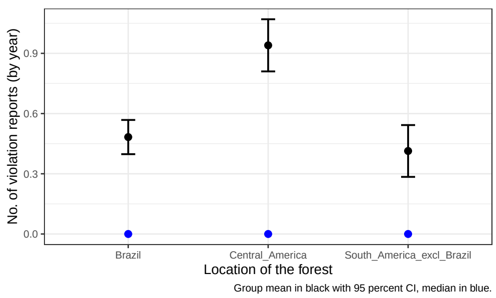
Figure 1.6: Major violations reports in FSC audits 2015-2019 across regions



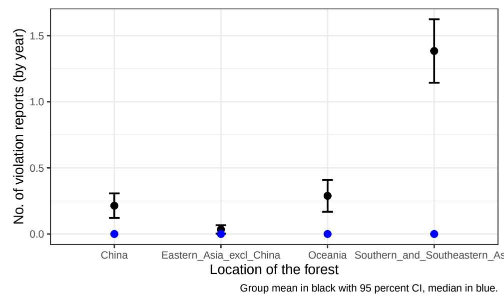
(a) Global comparison



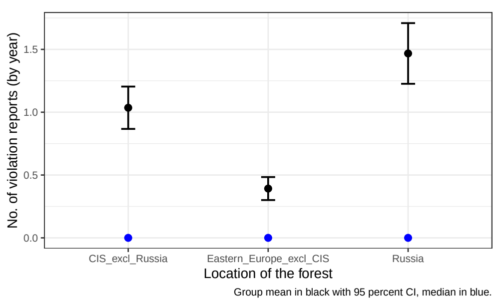
(b) Africa



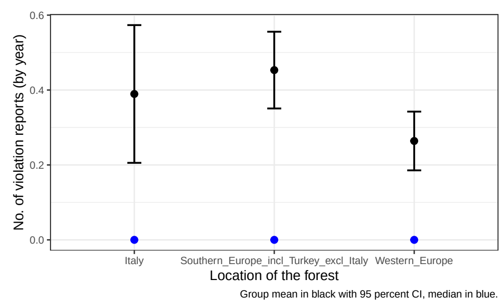
(c) Latin America



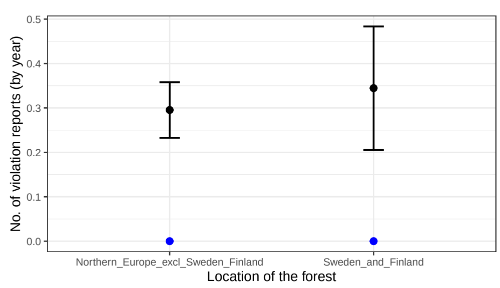
(d) Asia



(e) Europe (1)



(f) Europe (2)



(g) Europe (3)

Figure 1.7: Certificate list in FSC's public database (FSC 2023)

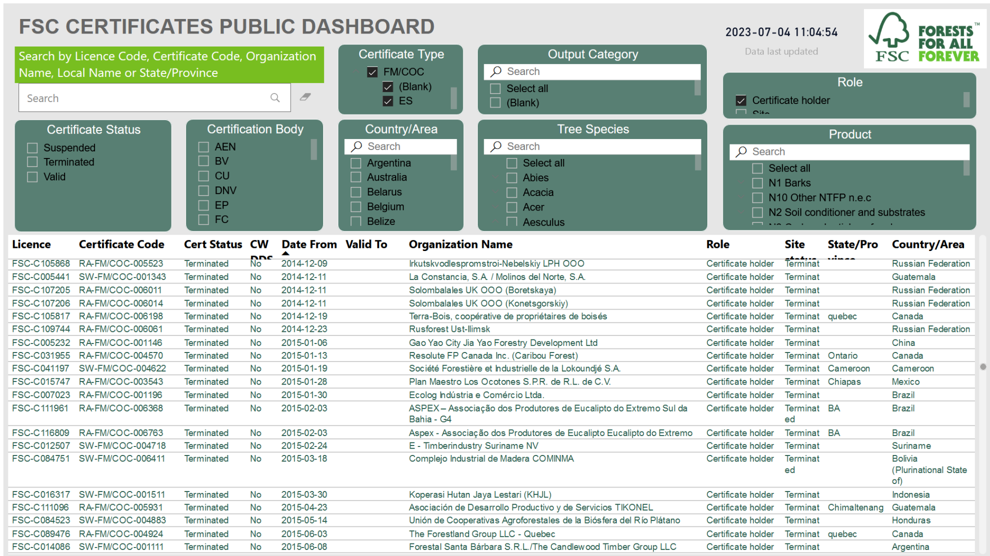
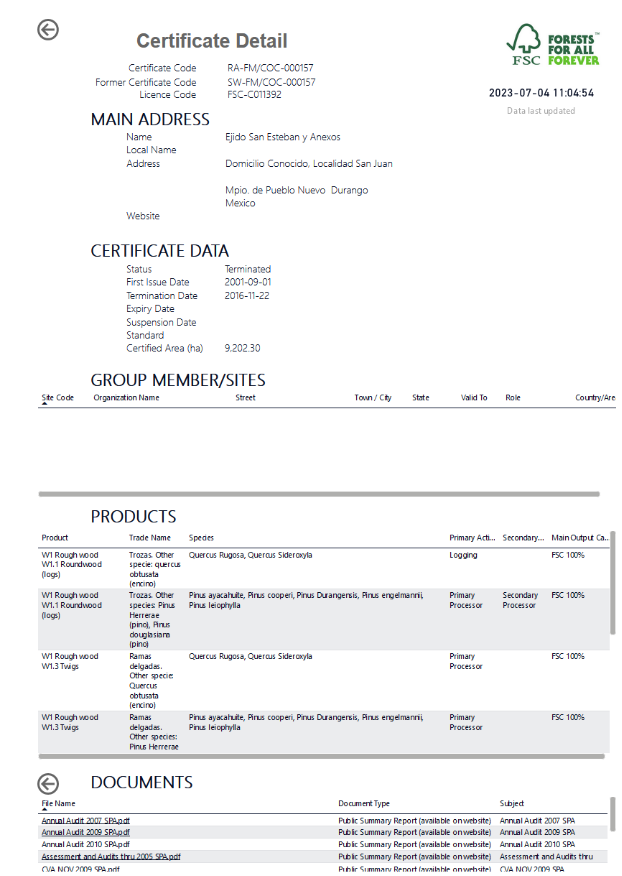


Figure 1.8: Certificate profile in FSC's public database (FSC 2023)



From the FSC Certificate Search, I scraped basic information about all FMUs and the links to their profiles. Basic information on the certificates and the files published on the profiles were downloaded first in June 2020[64] and then in September - November 2021.[65]

Filtering one report per audit

From all the files, the final dataset should only contain one summary for each audit. In case of several files referring to the same audit, the filtering process aimed to choose (1) the "oldest" or most original summary, i.e., the summary with the report date that is closest to the audit, and (2) the summary in the most common language from the perspective of the Researcher.

The motivation for choice (1) was that there were cases where corrected violations were no longer mentioned in later versions of an audit summary. The motivation for choice (2) was first to facilitate the correct extraction of variables by preferring languages that were either known to Researcher (English, French, German) or which were more prevalent (Spanish preferred over Portuguese, Russian over Ukrainian).

The files were filtered with algorithms built according to those rules by the following means:

1. Files with no audit summaries ("Public Summary Report"), according to the information on the certificate profile or clear indicators in the file name, were removed, such as member lists and species lists.
2. For other files, different versions in time or languages could be identified by certain suffixes used in the file names by certain certifiers, e.g., "_ENG.pdf" for English or "_V2.pdf" for a second version in time of the same report.
3. The language of the file was identified using language-detecting algorithms from the R packages *cld*2 and *cld*3. For two identical or almost identical file names, with one file being in English and the other in a local language, the one in the local language could then be removed. Languages are privileged in the following order: English, Spanish, French, German, Russian

[64]The certificate profiles were scraped the 04/06/2020, including the links to the files published on the certificate profiles. From those links, the files were downloaded from 04/06/2021 until 01/07/2021. The long time scope for that was taken to avoid overcharging the FSC portal.

[65]The list of certificates from the FSC Certificate Search, including the certificate profile links, was scraped on 30/08/2021. Information from the certificate profiles, most importantly the file links were scraped on 02/09/2021. Yet, due to an error noticed later, additional information from the certificate profiles, such as the certified forest area, was incompletely scraped, so the data from certificate profiles in this final, complete dataset comes from scraping the profiles the 11/11/2021.

4. Once audit and report dates were extracted, these were used to identify different summaries of the same audit. The identified certifier and audit type were also used as further confirmation of identical audits.

Variable extraction

Relevant variables were extracted from the audit summaries using automated extraction with regular expressions and manual addition wherever the former was too complicated or risky. Based on the extraction of the audit year, the dataset was restricted to reports on audits conducted from 2015 to 2020.[66] The information extracted from one audit summary concerns on the most recent audit covered in a report (see the discussion in the section on Completeness). The following subsections document choices made in the extraction.

The certifiers are identified from the certificate code in the document. If this is not found (majority of the files), it is identified from mentions of the certifier in the full name or sufficiently unambiguous abbreviation on the first page. Suppose this is not found (a few hundred cases, in particular, where the certifier only appears in a label in image format). In that case, the certifier is identified from patterns in the naming of the file or manually added.

Audits are classified as (Re)certification audits (Initial or Recertification), Annual surveillance audits, Audits to verify the correction of major violations, Other special audits. Audit types (3) and (4) are excluded in the analysis. The types are mostly identified from the title or corresponding table on the front page of the file. In other cases, clear patterns in the file name indicate the type. For a small remaining part, the types were added manually.

The report date that was aimed to extract is the most recent update date of the audit summary, as it should be published on the front pages. Where the algorithm finds no report date, dates from the file names are used if sufficiently clear. The audit years are extracted where possible from the front pages and, if not possible, looked up in the file, typically in the section of the audit schedule.

The violation reports were collected in the following steps. I aimed to include all upgraded major violation reports as new violation reports.

1. Where possible, with sufficiently clear identifiers, violation references are extracted into the violation list. The number of such references is counted as the number of

[66] Temporarily, also reports from 2014 were included to remove observations from 2014 that appear again in 2015, as will be explained for the major violation reports below.

violation reports. Upgraded violations that were clearly marked as such obtain the suffix “upgraded” in the violation list.

2. Where this is too complicated or risky but a section in the file states how many violations of each grade were found, the number of violations is extracted directly from there.
3. The rest of the violations are added manually.
4. For violations collected in step 1 and from certifiers where it is not always possible to exclude violations from previous audits, violations from previous audits were removed ex-post in a dedicated algorithm:

- First, by excluding clear violation references that appear in several audit summaries of the same CH are removed from the newer audit summary. Wherever the violation reference is not clear enough to be unique, we check manually whether the reference in the same document refers to the same violation or not (e.g., for CB 2 the format of the corresponding table often only allows extracting simple numbers of violations such as “1”.)
- Second, by excluding those violations whose patterns suggest that they were obtained in a previous year (CB 1(a) for example, typically denotes violations as “01/18” in 2018, which the algorithm would remove for an audit of the year 2019), unless if the reference had the suffix “upgraded”. For CB 1(a), CB 1(b) and Imaflora, where upgraded did not have clear markers and would thus be wrongly removed, all files were violation patterns from the previous year (e.g., “01/18” for an audit in 2019) were rechecked manually to add upgraded violations where necessary. For CB 7, where upgraded violations are not marked but not extracted in the algorithm above, all documents were rechecked, and upgraded violations were added where necessary.

5. Each file with duplicate violation references was checked manually to remove the duplicates in case they refer to the same NC. This may not be the case, either if the algorithm did not extract the complete violation reference[67] or if the certifier uses two identical references for different violations (recognized by having a separate box with a different description and concerned indicator of the NC).
6. In all cases where only the number of violations was extracted but was different from zero, the references were added later when adding the concerned indicator/norm.

For easily identifiable and common combinations of certifier and language, the indicator or

[67] E.g. if the rest of the reference appears in the next line after text in other columns in a table that is not actually formatted as a table as often for CB 2.

norm/reference text field was extracted automatized for each reported major violation, which was detectable with regular expressions. For all others, the text fields were added manually. The text fields also include the standard references for some certifiers but not for others. Yet, given that the numbering of national FSC standards shall follow the international Principles and Criteria, the indicator should allow to identify at least those. This is not true in some cases, potentially due to exceptions to the mentioned rule. Yet, the applicable national standards should be identifiable, given the information on the country and year of the audit and information on the FSC website on standard changes (https://fsc.org/en/document-centre). From the text field, the indicators were extracted using the standard pattern of the principle number being followed by a dot and the criterium number, potentially again followed by a dot and the indicator number (in some cases, the dot is further followed by a “p” for Principle or “c” for Criterium etc.).

1.9.1.2 Audit data accuracy and completeness

This section provides a quick evaluation of the extent to which the dataset matches the information published on the certificate profiles and in the audit summaries, as far as we have been able to assess it, given time constraints. Further possible assessments are suggested.

Completeness

The datasets of the FSC certificate search and, consequently of the certificate profiles scraped in 2020 and 2021 are complete: After scraping, they contained the same number of observations as found online (checked at the time): 3532 in 2021 and 3309 in 2020. After removing duplicates and certificates that were terminated before 2015, there are 2506 in 2021 and 2277 in 2020.

Possible reasons for the missing audit summaries in the final dataset include the following:

- Audit summary is not or no longer on the certificate profile (In many cases, we could see that after a certifier switch, the reports from the previous certifier are no longer available (even though they should have kept them))
- Unaccessible files (not downloadable / behind a login).
- Corrupted / not machine-readable files
- Filtered by the algorithm due to inaccurate information in the audit summary (typos in the audit dates etc.)
- Filtered by the algorithm due to inaccuracies in the algorithm itself

Given that there are exceptions to the necessity of an annual audit in the FSC standard and that certifiers are allowed to join reports of several annual surveillance audits in the same document, it is not straightforward to check whether audit summaries are missing. In most of the latter cases, a new version of the audit summary is published for each audit, so we always decided to extract only the information on the most recent audit covered in a report. Given this choice, information from former audits may be missed if the certifier removed the former version of the joint file of audit summaries.

Scrape 2020:

The total number of files of the certificate profiles at the time of the download was 22730. 27 files could not be accessed and downloaded, partly due to a wrong link or requiring a login. 158 were identified as duplicates and removed. The remaining files were read into a text format using the R package *readtext*. Excel files were read into as lists of data frames (one per sheet), using the R package *readxl*. For 191 downloaded files, no text could be read. Those files are corrupted.

Scrape 2021:

2021 we only downloaded the files under URLs that were not contained in 2020, 7578 in total. 694 could not be read in or were duplicates.

After removals of non-audit files, newer versions of the same audits and translations, we have in the merged dataset around 9000 audit summaries. There are around 2000 additional CH-year combinations (not included in the dataset) without an audit summary[68]. Yet, in a good part of cases, there are instead two audits in the previous year. The summary seems missing in others, potentially for one of the reasons noted above. Further checks would be necessary to ensure complete audit sequences.

It would be useful, for example, to check for 100 randomly chosen certificates, to what extent, and why audit reports were missing compared to the files published on the certificate profiles. In the current project, there was no time for this though.

Accuracy of extracted data

For 100 randomly chosen files, the accuracy of the extracted variables was checked manually (accuracy according to the information found by looking into the report). For the first set

[68] Excluded in the final dataset as that is organized at the audit summary level.

of variables, this was checked on subsamples drawn from the whole dataset, including some fraction of manually extracted.

Variable	Accurately assigned
Certifier	99%
File type	97%
Audit date	99%

For the second set of variables, this was checked on subsamples drawn from only those observations that were not added manually anyway. The approximate fraction of observations of the given variable in the whole dataset, which was added manually, is given in the second column. Assuming all manually added observations are correct, the fourth column computes the overall accuracy according to this check.

Variable	Fraction manually detected	Accurately assigned among obs. extracted with Regex	Total accurately assigned if one assumes 100% accuracy for manuals
Number of major violation reports	approx. 34%	97%	approx. 98%
Number of minor violation reports	approx. 26%	>91%*	>93%
Reported major violation indicators	approx. 36%	>91%*	>94%

*The sources of most mistakes in that test were corrected in the newest dataset version. The accuracy rate is likely much higher now, but there was no time left for another test.

1.9.1.3 Applicable rules for audit summaries

Public summaries of FSC Forest Management certification audit reports and annual surveillance updates (hereafter "audit summaries") are published according to a set of rules specified in FSC's standard for certifiers and accompanying guidelines (FSC, 2009a, FSC (2021b)).

The most relevant for the data collection and analysis can be summarized as follows:

1. **Content:** Among others, the following variables must be published in each audit summary: The name and contact details of the certifier and the certificate holder, the date the report was last updated, the name and location of the certified forest areas, the date of issue and expiry of the certificate, the dates, duration, and type of the field audit (annual surveillance, (re)certification or correction of violations), the (updated) certification decision and most importantly "a list of all non-conformities [violations] that the managers are required to correct to maintain their certification, including the time course within which corrective actions shall be taken" (See Art. 5, 7, 8 and Box 1 in FSC (2009a)).
2. **Publication date:** The summaries of the certification audits must be published before a certificate is issued / re-issued. Annual updates must be published within 90 days after the surveillance audit. (See Art. 3 in FSC (2009a))
3. **Deletion date:** The audit summaries shall remain in the public database, even after the suspension of a certificate holder (See question INT-STD-20-007b_05 in FSC (2021b)). FSC clarified this in 2020, so there might be inconsistencies before that.
4. **Language:** For certificates that cover more than 1,000 ha forest area and for which not all group member forest units are "small", i.e., less than 100ha each (Or designated as small in a formal procedure outlined Art. 2.2 FSC (2004)) the summaries must be available in one of FSC's official languages, which are. For smaller forests, the local/national language of the forest's location is sufficient (See Art. 2.1 in FSC (2009a))

Certifiers have to upload the audit summaries on a database that is managed by FSC and from which they are automatically published on a public certificate profile that can be accessed through a search engine.[69]

1.9.1.4 List of yearly country-level characteristics

- The numbers of FSC-affine upstream firms as upstream demand proxies (FSC, 2023b), including certified downstream firms, certified public projects, and retailers officially promoting FSC products.
- Availability of national FSC standards (FSC, 2023a)
- Trade values and volumes in different wood product categories (FAO, 2023)
- The percentage of publicly and privately owned forests per country in 2015 (FAO, 2020)

[69]The search engine is https://info.fsc.org/certificate.php. An example of a certificate profile accessed through the search engine is https://info.fsc.org/details.php?id=a023300000azGp6AAE&type=certificate

- Transparency International's Corruption Perceptions Index (Transparency International, 2021)
- Indicators of progress towards sustainable forest management (United Nations, 2021)
- GDP per capita, profit tax rates, currency conversion factors, and inflation rates (IMF, 2020; World Bank, 2021)
- Geographic characteristics, such as being landlocked, distances, and languages (Mayer and Zignago, 2011)

1.9.1.5 Variation in violation reports and FMU characteristics

Figure 1.9: Major violation reports in audits with at least one reported major violation

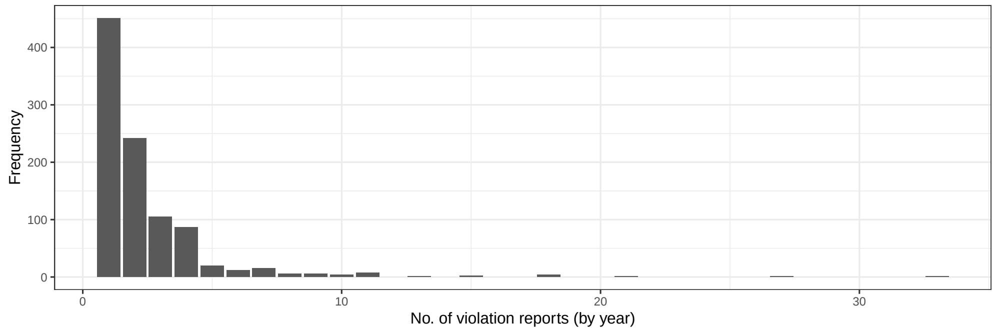
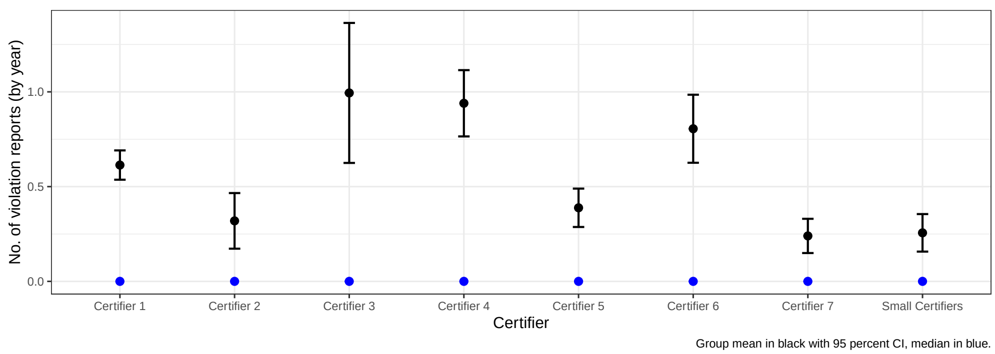


Figure 1.10: Major violation reports across certifiers



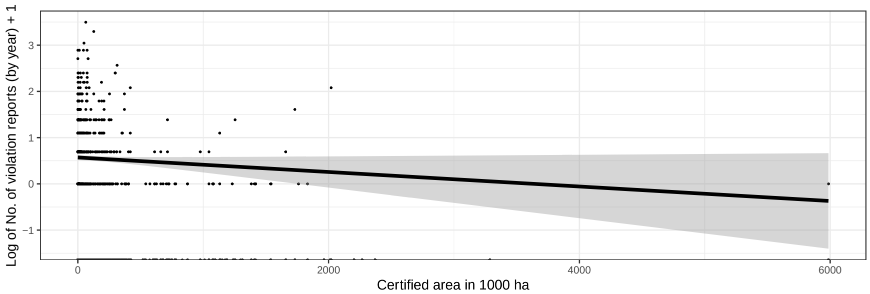
1.9.2 Market definition, choice sets and certifier-market panel

1.9.3 Price panel

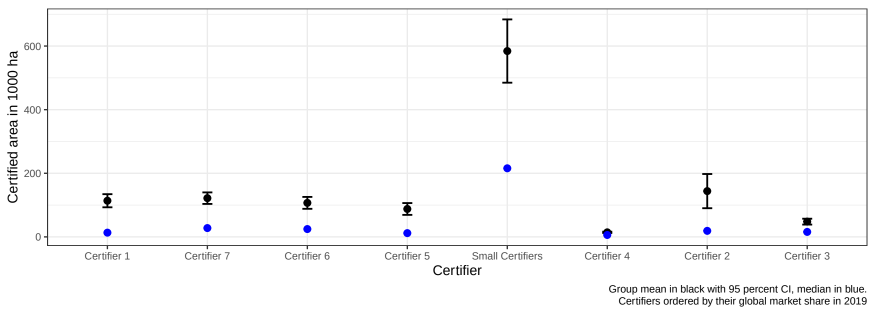
1.9.3.1 Survey design and response

The survey was sent by e-mail to all FMUs with a valid certificate in June 2020 for whom an e-mail-address could be found. I excluded terminated certified FMUs since they are likely to be less motivated to participate in a survey and might even have an incentive to give false answers out of negative sentiments towards FSC. From the 1756 valid certificates in June 2020, e-mail-addresses of FSC contact persons could be extracted from audit report summaries for 1456 and found through their websites for 109 unique certified FMUs. 101 certified FMUs have two or more certificates (up to 6) for different forest units which may be the case if a firm owns forests in very diverse locations. These were asked to respond for each certificate separately but did not follow the request. 12 separate certificates of the same certified FMUs were however added manually ex post based on the information three

Figure 1.11: Selection of FMUs of varying size into certifiers

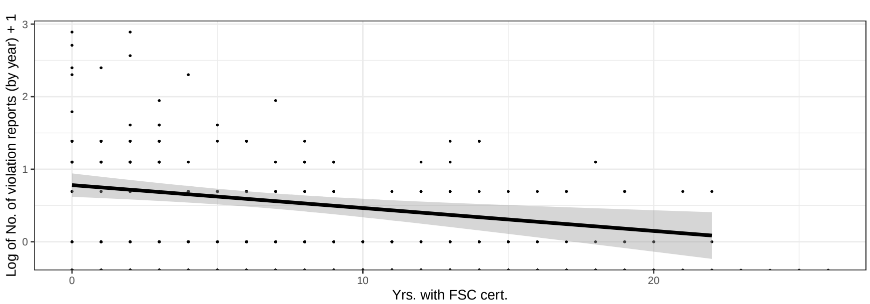


(a) Major violation reports per year in 2015-2019

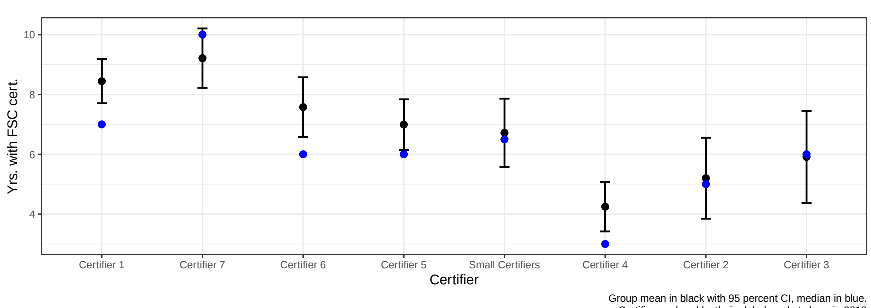


(b) Mean per certifier in 2015-2019

Figure 1.12: Selection of FMUs with varying certification experience into certifiers

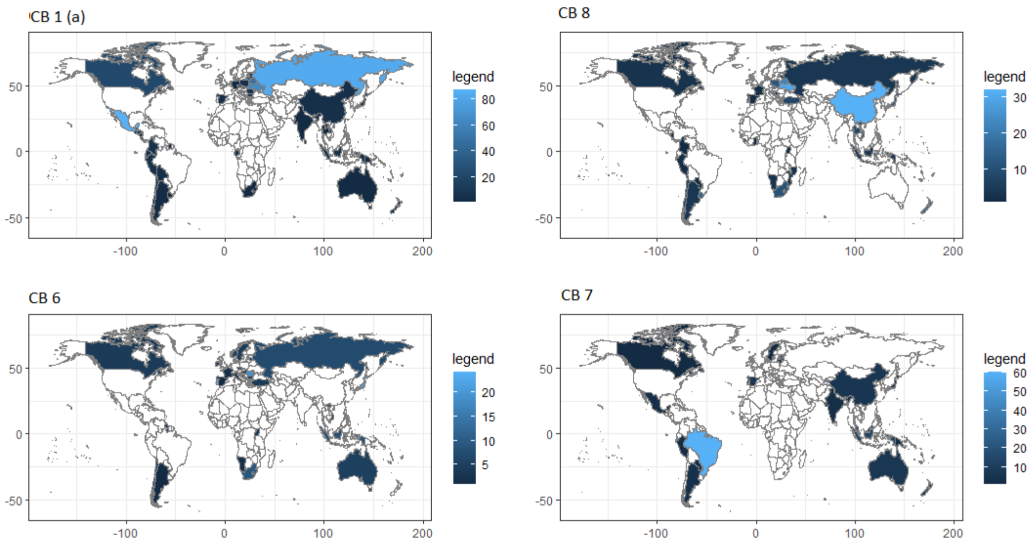


(a) Major violation reports in 2019

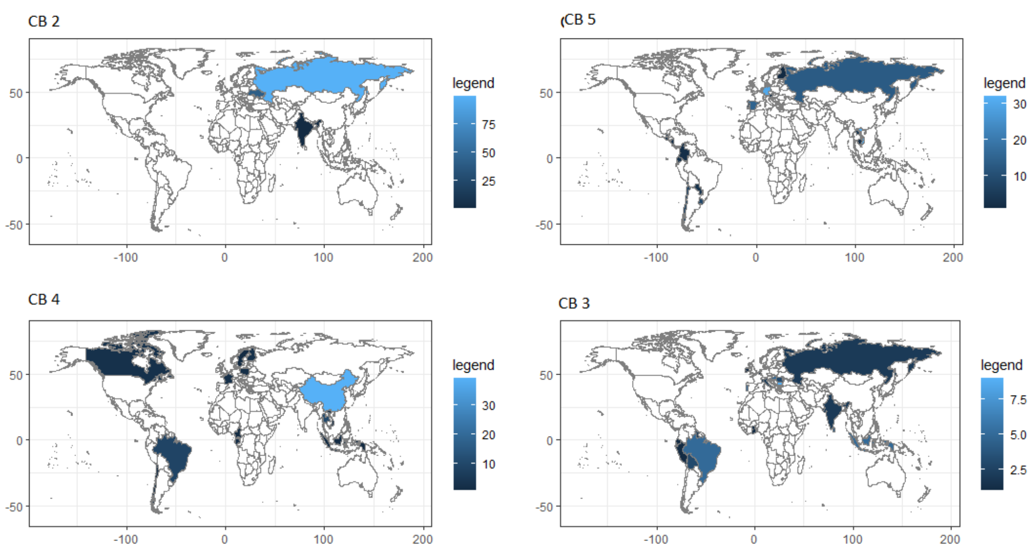


(b) Mean per certifier in 2019

Figure 1.13: Numbers of valid certificates in May 2020 issued by the largest certifiers (CB) in different countries



(a) Certifiers 1, 8, 6 and 7



(b) Certifiers 2, 5, 4 and 3

certified FMUs gave in comments about the multitude of certificates they have / have had in the past. Over 100 e-mail-addresses turned out to be out of date or incorrect, about 40 could be replaced by another contact that was accessible. For 66 cases, no other email address could be found.

The certified FMUs with valid e-mail-adress were invited by e-mail to participate in a short survey asking questions about the prices they had to pay for certification and a few characteristics of them that may be relevant to classify the prices It was conducted in an anonymous manner in order to encourage firms to participate in spite of the sensitivity of the issue.

335 certified FMUs submitted complete responses to the survey, a response rate of 21 Percent.[70] Complete means that they replied at least to all mandatory questions[71] and explicitly submitted their response at the end of the survey. In comparison, response rates in other, much smaller surveys of FSC certified FMUs range from about 15% to about 80% (Araujo et al., 2009; Galati et al., 2017; Jaung et al., 2016; Overdevest and Rickenbach, 2006). These surveys did not face the additional challenge of a multiplicity of languages and locations (not permitting a dual-mode survey) and questions about prices which are particularly sensitive. Moreover, the way in which I use the data from the survey in the main project does not require it to be perfectly representative.

From the remaining respondents,[72] I obtained 1122 informative price quotes or approximations, out of which 877 were prices quoted explicitly.[73] To understand the difference, note that

[70]In addition, there were 2 respondents that seem to be not part of the target population: They claimed that their forest had never been certified. This might be due to a mistake in the e-mail-address, an employee that is not aware of the certification or a lack of willingness to respond. They were thus excluded from the survey data.

[71]These can be recognized by a red * before the question, as seen in Figure 1.14 and the following.

[72]50 of the respondents had to be removed, 25 of them since an unknown fraction of their certification fee is paid by a donor, while the prices quoted by the other 25 were not informative or reliable for other reasons. 9 of them did not mention their certifier. (7 commented that the price quoted included multiple certificates, either multiple FSC certificates without saying how many or certificates from other standards. 6 did not enter any price quote while 1 entered Zeros everywhere.)

[73]2 of the included respondents' yearly price quotes were removed because the respondents did not note the certifier for the corresponding year. 11 concerned years in which the certificate holder was not certified and thus paid nothing. 59 price quotes concerned years before an extension of the certificate; these observations were excluded as important characteristics of the certificate such as the certified area were only provided with respect to the period after the latest extension. 4 observations concerned years in which the certificate holder quoted a price for multiple certificates together without specifying the number or mixing with a certificate from another standard. In 6 cases it was noted that they had a certifier which was not in the suggested list without specifying which one. And in 23 cases, a zero price was noted which might not be wrong if no audit was conducted in the year, but my analysis uses only years in which audits are conducted so that such "price

the survey first asked the respondent to type the price paid for certification in each year (in the following "direct price quotes"). In a second question, respondents were asked to use a slider to mark the approximate price (in 5000 USD intervals or the chosen currency's corresponding interval). Since also the direct price quotes can be approximations, I may use the approximate price quotes in the 245 cases in which no direct price quote was given. In my main specification, I focus however on the sample with the direct price quotes. The approximate price quotes allows for a robustness check.

The price quotes were then transfered into USD prices, using the currency conversion factor used by the World Bank (World Bank, 2020). They were further adjusted across time, using the US paper and wood pulp PPI, with 2015 as base year (FRED, 2020).

To analyze whether the survey data are reliable, I compare the distribution of respondents to the full sample of valid certificates and the price quotes of similar firms to check for consistency in the next sections. I also checked for consistency of denoted world regions, certifier availability in this region, the language in which the survey was conducted and the currencies denoted. Responses were consistent and many respondents raised additional confidence in their motivation to respond thouroughly by additional comments they made in writing.

1.9.3.2 Comparison with the main dataset

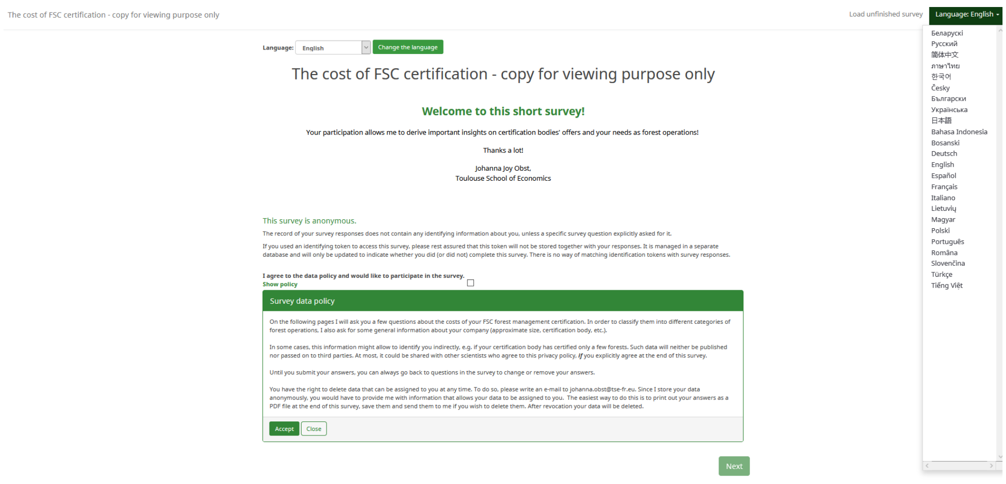
A comparison of the distribution of all characteristics that we observe both in the survey and in the certificate dataset suggests that the survey respondents do not differ much from the whole population of certified FMUs. This can be seen in Figures 1.19, 1.20, 1.22 and 1.21 comparing the distribution of respondents and all certified FMUs across geographic regions, experience with certification as well as size of the forest. Four differences seem nevertheless relevant.

First, Figures 1.19 show that some certifiers are more representated than others. Especially underrepresentated is CB 8with only 24 informative price quotes. Corresponding estimates will thus need to be interpreted with caution and I thus exclude the countries of the Commonwealth of Independent States in at least one version of my analysis, as it is the only region in which CB 8is present.

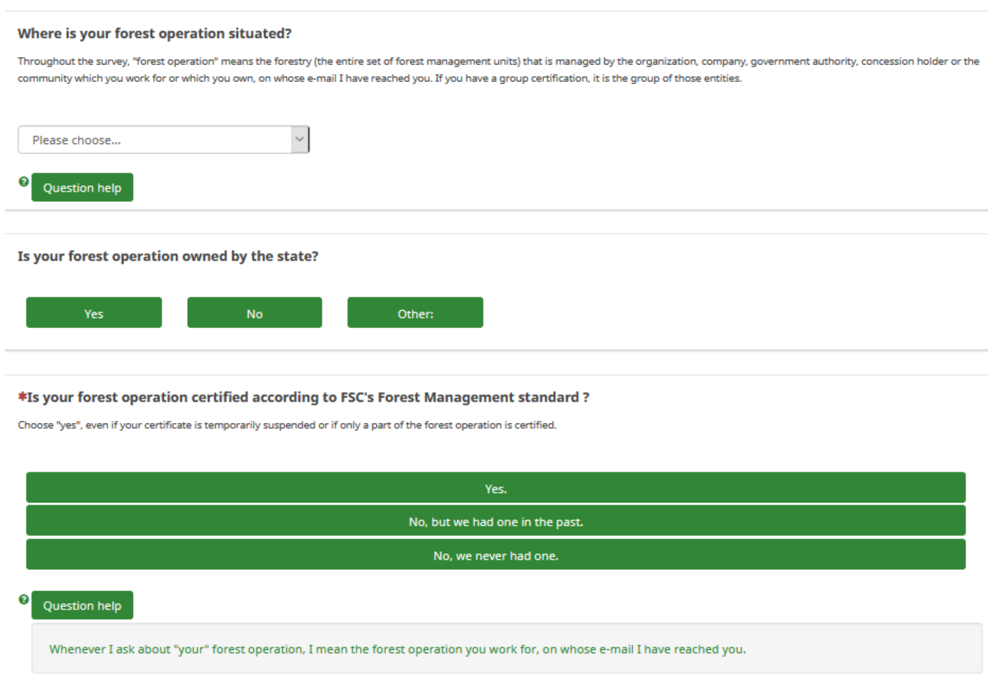
Second, respondents tend to have been certified for a slightly longer time than the whole

quotes" should not be taken into account.

Figure 1.14: Overview over the survey with valid FSC FM certificate holders (Oct.-Nov. 2020)

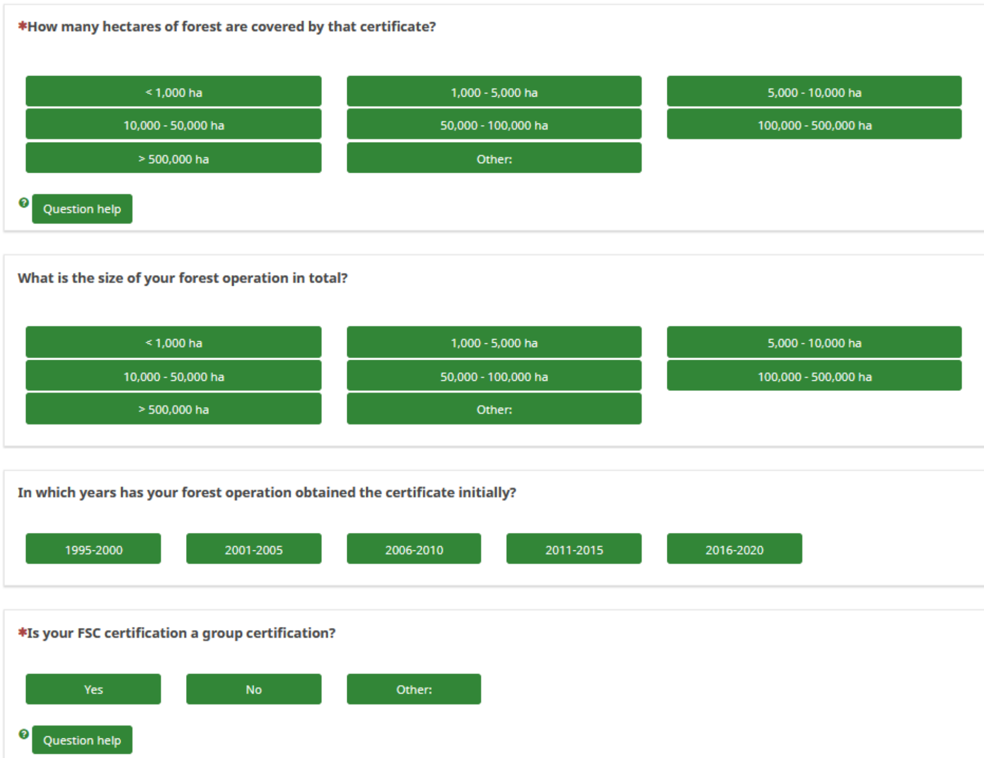


(a)



(b)

Figure 1.15: Overview over the survey with valid FSC FM certificate holders (Oct.-Nov. 2020)



*How many hectares of forest are covered by that certificate?

< 1,000 ha | 1,000 - 5,000 ha | 5,000 - 10,000 ha
10,000 - 50,000 ha | 50,000 - 100,000 ha | 100,000 - 500,000 ha
> 500,000 ha | Other:

Question help

What is the size of your forest operation in total?

< 1,000 ha | 1,000 - 5,000 ha | 5,000 - 10,000 ha
10,000 - 50,000 ha | 50,000 - 100,000 ha | 100,000 - 500,000 ha
> 500,000 ha | Other:

In which years has your forest operation obtained the certificate initially?

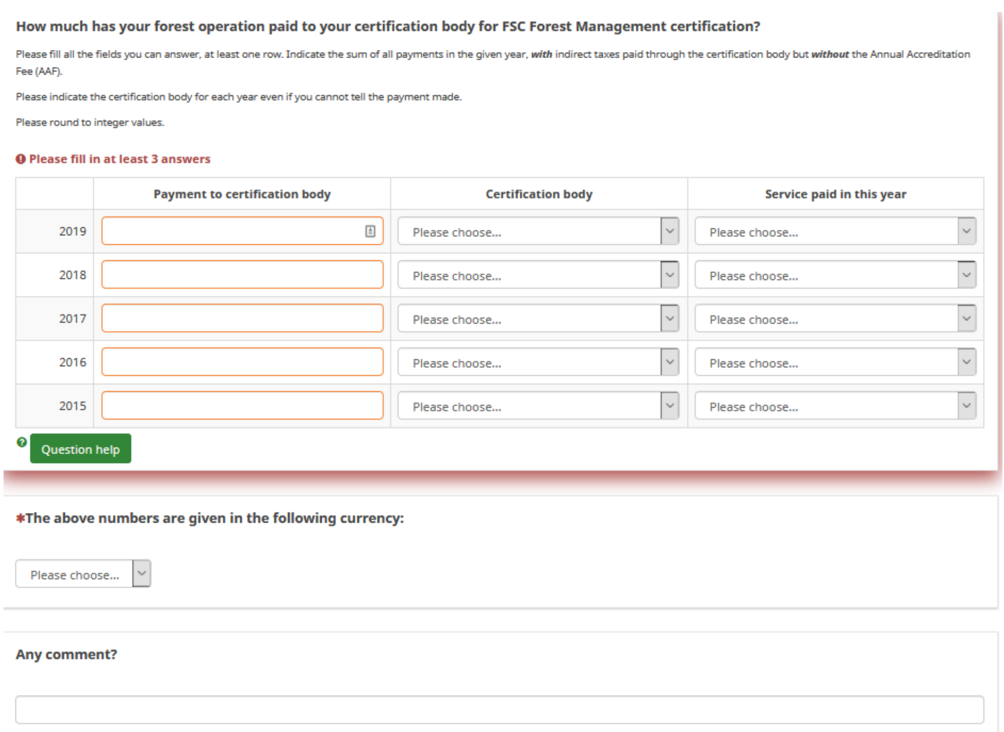
1995-2000 | 2001-2005 | 2006-2010 | 2011-2015 | 2016-2020

*Is your FSC certification a group certification?

Yes | No | Other:

Question help

(a)



How much has your forest operation paid to your certification body for FSC Forest Management certification?

Please fill all the fields you can answer, at least one row. Indicate the sum of all payments in the given year, *with* indirect taxes paid through the certification body but *without* the Annual Accreditation Fee (AAF).

Please indicate the certification body for each year even if you cannot tell the payment made.

Please round to integer values.

Please fill in at least 3 answers

	Payment to certification body	Certification body	Service paid in this year
2019		Please choose...	Please choose...
2018		Please choose...	Please choose...
2017		Please choose...	Please choose...
2016		Please choose...	Please choose...
2015		Please choose...	Please choose...

Question help

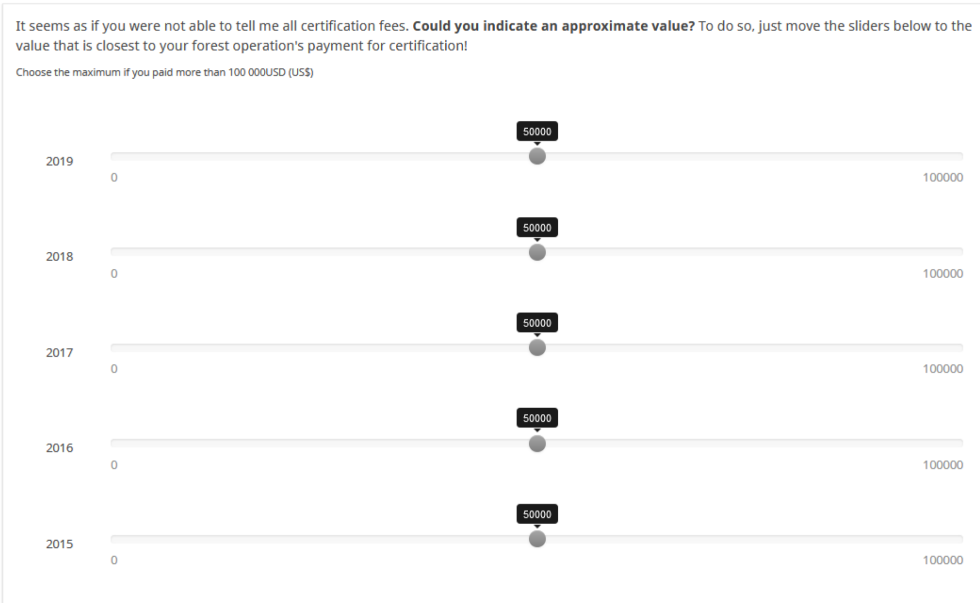
*The above numbers are given in the following currency:

Please choose...

Any comment?

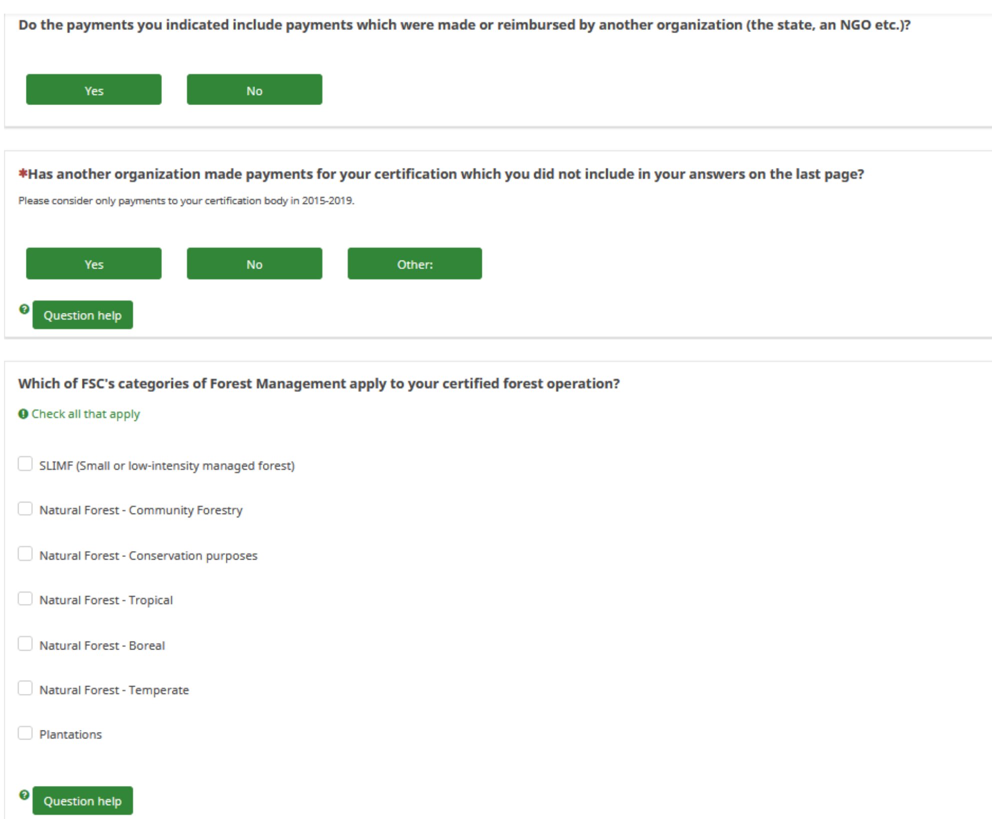
(b)

Figure 1.16: Overview over the survey with valid FSC FM certificate holders (Oct.-Nov. 2020)


It seems as if you were not able to tell me all certification fees. **Could you indicate an approximate value?** To do so, just move the sliders below to the value that is closest to your forest operation's payment for certification!

Choose the maximum if you paid more than 100 000USD (US$)

(a)


Do the payments you indicated include payments which were made or reimbursed by another organization (the state, an NGO etc.)?

Yes | No

***Has another organization made payments for your certification which you did not include in your answers on the last page?**

Please consider only payments to your certification body in 2015-2019.

Yes | No | Other:

Question help

Which of FSC's categories of Forest Management apply to your certified forest operation?

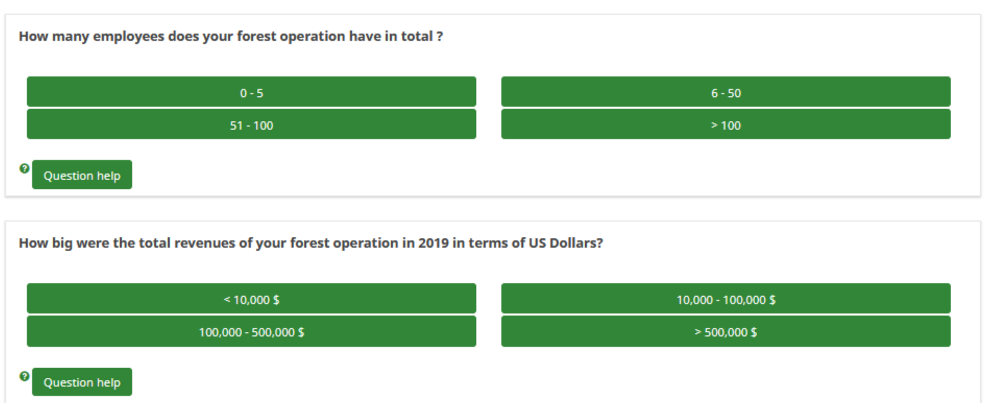
Check all that apply

- SLIMF (Small or low-intensity managed forest)
- Natural Forest - Community Forestry
- Natural Forest - Conservation purposes
- Natural Forest - Tropical
- Natural Forest - Boreal
- Natural Forest - Temperate
- Plantations

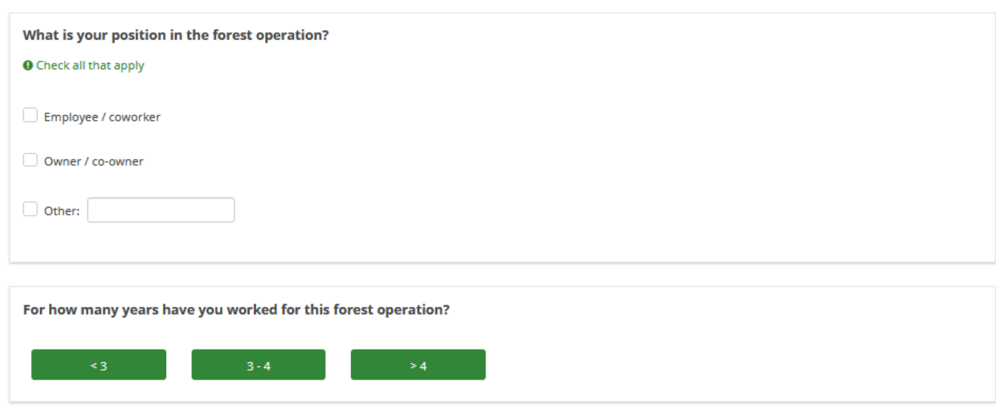
Question help

(b)

Figure 1.17: Overview over the survey with valid FSC FM certificate holders (Oct.-Nov. 2020)

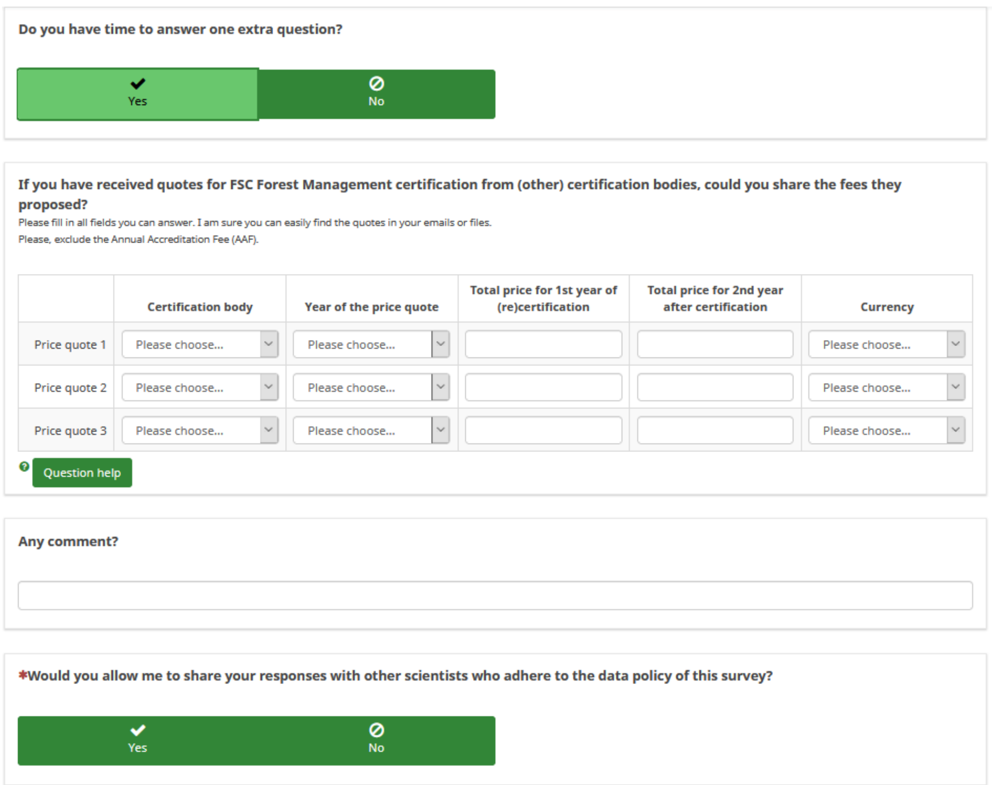


(a)



(b)

Figure 1.18: Overview over the survey with valid FSC FM certificate holders (Oct.-Nov. 2020)



Do you have time to answer one extra question?

Yes | No

If you have received quotes for FSC Forest Management certification from (other) certification bodies, could you share the fees they proposed?

Please fill in all fields you can answer. I am sure you can easily find the quotes in your emails or files.

Please, exclude the Annual Accreditation Fee (AAF).

	Certification body	Year of the price quote	Total price for 1st year of (re)certification	Total price for 2nd year after certification	Currency
Price quote 1	Please choose...	Please choose...			Please choose...
Price quote 2	Please choose...	Please choose...			Please choose...
Price quote 3	Please choose...	Please choose...			Please choose...

Question help

Any comment?

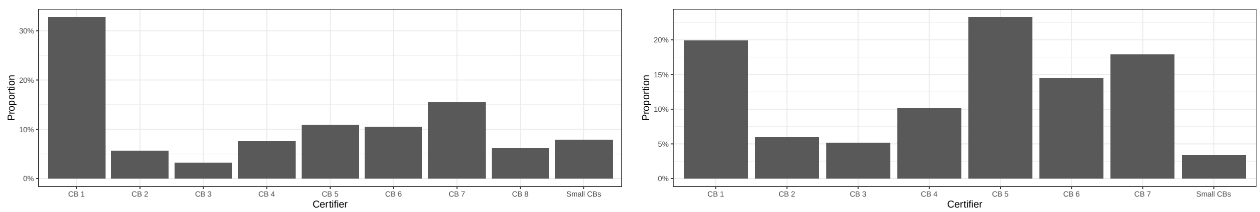
***Would you allow me to share your responses with other scientists who adhere to the data policy of this survey?**

Yes | No

population as Figures 1.21 show. This might be related to earlier certified FMUs being more intrinsically motivated for sustainability and more motivated to contribute to corresponding research and a survey. This is a possibility that will need to be considered in the interpretation of the results of my study later on.

Third, there are relatively more respondents from Western Europe and relatively less from Eastern Europe. Region dummies will help to control for this in the price estimation in the main project.

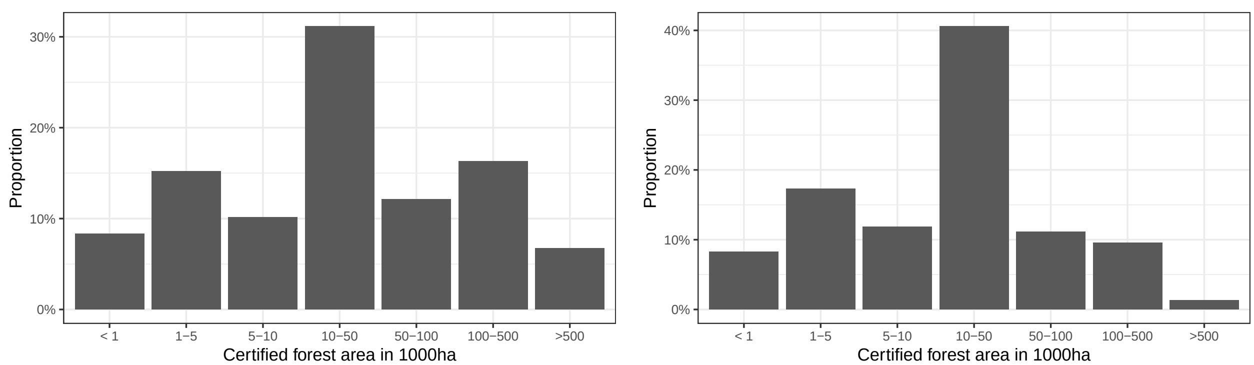
Figure 1.19: Distribution across certifiers



(a) Target population: FSC FM certificate holders, valid in June 2020

(b) Survey respondents

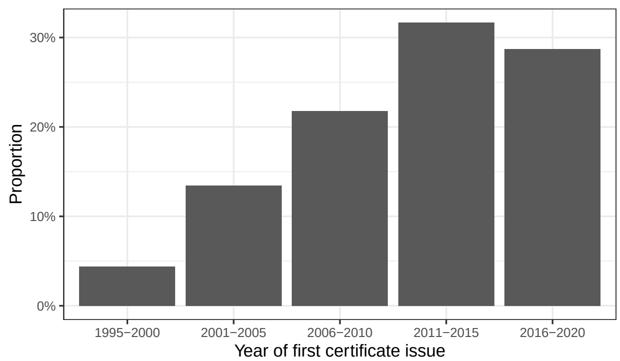
Figure 1.20: Distribution across forest size categories



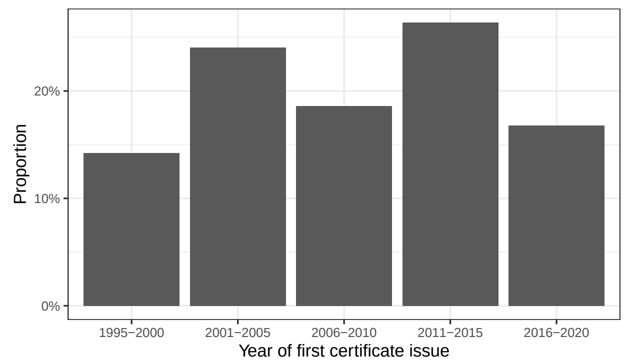
(a) Target population: FSC FM certificate holders, valid in June 2020

(b) Survey respondents

Figure 1.21: Distribution across initial certification year categories

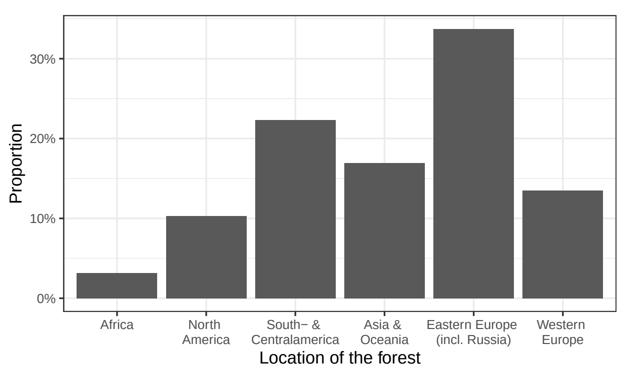


(a) Target population: FSC FM certificate holders, valid in June 2020

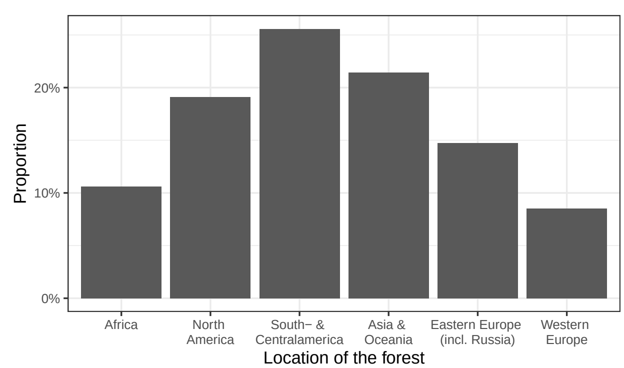


(b) Survey respondents

Figure 1.22: Distribution across regions



(a) Target population: FSC FM certificate holders, valid in June 2020



(b) Survey respondents

	Forest location	**Respondents**
1	Africa: - Eastern Africa	3
2	Africa: - Middle Africa	3
3	Africa: - Southern Africa	9
4	Africa: - Western Africa	1
5	Americas: - Caribbean	1
6	Americas: - Central America	17
7	Americas: - Northern America	31
8	Americas: - South America	49
9	Asia: - Central Asia	2
10	Asia: - Eastern Asia	18
11	Asia: - South-eastern Asia	23
12	Asia: - Southern Asia	3
13	Europe: - Eastern Europe (including Northern Asia)	53
14	Europe: - Northern Europe	24
15	Europe: - Southern Europe	23
16	Europe: - Western Europe	45
17	Oceania: - Australia and New Zealand	9
18	Oceania: - Melanesia, Micronesia & Polynesia	2
19	No answer	11

1.9.3.3 Prediction of unobserved prices

1.9.3.3.1 Accounting for selection into the survey

1. Following Heckman (1979), I model survey participation as the outcome of a Probit model, where the utility of survey participation is a function of predictors that are selected using repeated cross-validation with Lasso regularization:

$$
\begin{aligned}
u\_SurveyParticipation_{ijt} &= \underbrace{\boldsymbol{f}'_{it}\tilde{\boldsymbol{\rho}}^{\boldsymbol{fs}} + \boldsymbol{x}'_{jt}\tilde{\boldsymbol{\rho}}^{\boldsymbol{xs}}}_{E[u\_SurveyParticipation_{ijt}|boldsymbol f_{it},\boldsymbol{x_{jt}}]} + v_{ijt} \\
SurveyParticipation_{ijt} &= 1\{u\_SurveyParticipation_{ijt} > 0\} \qquad (1.24)
\end{aligned}
$$

2. I assume that the unobserved shocks to survey participation and prices, v_{ijt} and ϵ^p_{ijt} are jointly normal, with covariance σ_{21}. This implies that conditional mean of the price in the selected sample can then be written as:

$$
\begin{aligned}
&E\left[\log\left(\frac{p_{ijt}}{\log(\overline{area}_t)}\right)\Big|\boldsymbol{f_{it}}, \boldsymbol{x_{jt}}, SurveyPart_{ijt} = 1\right] \\
&= E\left[\log\left(\frac{p_{ijt}}{\log(\overline{area}_t)}\right)\Big|\boldsymbol{f_{it}}, \boldsymbol{x_{jt}}\right] + E\left[\epsilon^p_{ijt}|v_{ijt} > -\mathbf{f}'_{ijt}\tilde{\rho}\right] \\
&= E\left[\log\left(\frac{p_{ijt}}{\log(\overline{area}_t)}\right)\Big|\boldsymbol{f_{it}}, \boldsymbol{x_{jt}}\right] + \sigma_{21} InverseMillsRatio(\boldsymbol{f}'_{it}\tilde{\boldsymbol{\rho}}^{\boldsymbol{fs}} + \boldsymbol{x}'_{jt}\tilde{\boldsymbol{\rho}}^{\boldsymbol{xs}}) \qquad (1.25)
\end{aligned}
$$

   Not accounting for the second term would lead to omitted variable bias.

3. I construct the binary variable $SurveyParticipation_{ijt}$ for survey participation in the audit data. Given that the survey data is anonymized, the price quotes cannot be matched one by one to FMUs in the audit data. I therefore construct types of FMUs based on the characteristics collected in the survey and available in the FMU panel. These are the certifier, the year, the year of initial certification, the forest size, whether it has a group rather than an individual certificate, the region (defined as in the market) and whether the year requires a recertification rather than an annual surveillance audit.[74]

4. I estimate $\tilde{\rho}$ by probit regression of $SurveyParticipation_{ijt}$ and construct the Inverse Mills Ratio $InverseMillsRatio(\boldsymbol{f}'_{it}\tilde{\boldsymbol{\rho}}^{\boldsymbol{fs}} + \boldsymbol{x}'_{jt}\tilde{\boldsymbol{\rho}}^{\boldsymbol{xs}})$.

5. I add the Inverse Mills Ratio as an additional regressor to the pricing model, but do

[74] The year of initial certification and the forest size are given in categories in the survey data, as described in the data section of this paper. In the models of survey participation and pricing, I use the average by category and market in the FMU panel as a numerical value.

Table 1.19: Log of annual certification fee in 1000 USD per log area in 1000 ha, PPI adj.

	(1)	(2)	(3)
Certified area in 1000ha (lower bound by cat.) x log_country is landlocked_std	−0.061** (0.028)	−0.059** (0.029)	−0.075*** (0.025)
Internal dist. of country x log_country is landlocked_std	0.105*** (0.027)	0.105*** (0.027)	0.098*** (0.024)
Internal dist. of country x (Re-)Certification audit_std	0.073 (0.075)	0.075 (0.076)	0.077 (0.093)
No. of certified downstream firms in country in 2023 x (Re-)Certification audit_std	0.073 (0.044)	0.072 (0.045)	0.078* (0.046)
% of privately owned forest in country in 2016 x nat. avg. Longitude_std	0.229*** (0.062)	0.224*** (0.063)	0.241*** (0.062)
% of privately owned forest in country in 2016 x log_% of publicly owned forest in country in 2015_std	0.032 (0.066)	0.032 (0.065)	0.036 (0.067)
Forest nat. total profit tax (in %)_X___Forest nat. total profit tax (in %)sq_std	0.090 (0.079)	0.098 (0.079)	0.055 (0.085)
Forest nat. total profit tax (in %)_X___log_Yrs. with FSC cert._std	−0.111* (0.065)	−0.122* (0.070)	−0.106 (0.073)
Exported share of nat. production of sawnwood x nat. export from wood chips etc. in m3_std	−0.018 (0.071)	−0.021 (0.073)	−0.015 (0.073)
Exported share of nat. production of sawnwood x No. of promotional license holders in country by year_std	0.008 (0.040)	0.007 (0.040)	0.007 (0.041)
Exported share of nat. production of sawnwood x log_nat. import from fibre furnish in t_std	0.073 (0.055)	0.079 (0.058)	0.077 (0.059)
Exported share of nat. production of woodbased panels x log_Certified area in 1000ha (lower bound by cat.)_std	−0.216*** (0.063)	−0.222*** (0.064)	−0.211*** (0.065)
Exported share of nat. production of woodbased panels x (Re-)Certification audit_std	0.071** (0.028)	0.074** (0.030)	0.081 (0.073)
nat. export from fibreboard in m3 x Group certificate (vs. individual)_std	−0.021 (0.048)	−0.020 (0.048)	−0.023 (0.050)
nat. export from woodbased panels in m3 x No. of certified downstream firms in country in 2023 sq._std	0.042 (0.062)	0.037 (0.063)	0.020 (0.065)
nat. avg. latitude x log_Certified area in 1000ha (lower bound by cat.)_std	−0.199** (0.083)	−0.196** (0.084)	−0.197** (0.086)
nat. avg. latitude x log_country has an FSC national standard, by year_std_std	−0.012 (0.064)	−0.012 (0.064)	−0.004 (0.068)
nat. avg. Longitude x nat. net export from woodbased panels in 1000 USD_std	−0.041 (0.063)	−0.051 (0.065)	−0.031 (0.066)
nat. avg. Longitude x log_Certified area in 1000ha (lower bound by cat.)_std	−0.337*** (0.097)	−0.337*** (0.097)	−0.352*** (0.101)
nat. avg. Longitude x log_country is landlocked_std	0.040 (0.032)	0.042 (0.032)	0.051* (0.029)
nat. avg. Longitude x Classified as plantation_std	0.143*** (0.047)	0.140*** (0.046)	0.139*** (0.047)
Yrs. with FSC cert. sq. x Exported share of nat. production of sawnwood sq._std	−0.057*** (0.020)	−0.062*** (0.023)	−0.060*** (0.023)
Yrs. with FSC cert. sq. x nat. avg. latitude sq._std	−0.020 (0.056)	−0.020 (0.056)	−0.021 (0.057)
Yrs. with FSC cert. sq. x log_country has an FSC national standard, by year_std_std	−0.023 (0.047)	−0.026 (0.049)	−0.026 (0.049)
% of publicly owned forest in country in 2015 sq. x log_country has an FSC national standard, by year_std_std	−0.112*** (0.040)	−0.107** (0.042)	−0.108** (0.044)
Forest nat. total profit tax (in %)sq x (Re-)Certification audit_std	−0.011 (0.042)	−0.014 (0.043)	−0.004 (0.045)
Forest nat. total profit tax (in %)sq x Classified as plantation_std	0.034 (0.068)	0.041 (0.072)	0.068 (0.082)
Exported share of nat. production of woodbased panels sq. x nat. import from wood pulp in t sq._std	0.083 (0.065)	0.083 (0.066)	0.082 (0.068)
Exported share of nat. production of woodbased panels sq. x No. of promotional license holders in country by year sq._std	0.019 (0.045)	0.021 (0.046)	0.013 (0.047)
nat. export from fibreboard in m3 sq. x log_Exported share of nat. production of industrial roundwood_std	0.125*** (0.037)	0.129*** (0.039)	0.140*** (0.042)
Years since first FSC national standard_std sq. x nat. net export from wood chips etc. in 1000 USD sq._std	0.104** (0.053)	0.103* (0.053)	0.101* (0.054)
nat. net export from woodbased panels in 1000 USD sq. x (Re-)Certification audit_std	0.009 (0.030)	0.010 (0.030)	0.007 (0.036)
nat. net export from wood chips etc. in 1000 USD sq. x log_Certified area in 1000ha (lower bound by cat.)_std	−0.100** (0.050)	−0.101** (0.049)	−0.094* (0.051)
log_Certified area in 1000ha (lower bound by cat.) x log_% of privately owned forest in country in 2016_std	0.013 (0.074)	0.021 (0.074)	0.015 (0.073)
log_Certified area in 1000ha (lower bound by cat.) x log_Forest nat. total profit tax (in %)_std	−0.020 (0.066)	−0.028 (0.067)	−0.024 (0.067)
log_nat. export from wood chips etc. in m3 x (Re-)Certification audit_std	−0.001 (0.051)	−0.004 (0.052)	−0.012 (0.087)
log_nat. export from wood chips etc. in m3 x Group certificate (vs. individual)_std	0.050 (0.053)	0.049 (0.053)	0.061 (0.056)
log_country has an FSC national standard, by year_std x (Re-)Certification audit_std	−0.013 (0.050)	−0.013 (0.051)	−0.020 (0.051)
Group certificate (vs. individual) x Classified as plantation_std	0.063 (0.049)	0.061 (0.049)	0.058 (0.051)
Expected relative rigor x log_nat. export from fibreboard in 1000 USD per m3 (t-avg.)_std	0.025 (0.036)	0.026 (0.036)	0.043 (0.036)
Certifier market_group_no_sales_inclNone x nat. avg. latitude (t-avg.)_std	−0.009 (0.053)	−0.007 (0.054)	0.004 (0.055)
nat. export of wood chips in 1000 USD per cubic metre (t-avg.) x nat. export from wood chips etc. in m3 (t-avg.) sq._std	0.025 (0.029)	0.024 (0.029)	0.028 (0.029)
nat. export from paper products in 1000 USD per t (t-avg.) x log_ch_country_yr_Export_t_Wood_pulp (t-avg.)_std	0.054* (0.032)	0.056* (0.033)	0.055* (0.033)
Local currency unit per USD (t-avg.) x Country with certifier's closest office and country share the language (t-avg.) sq._std	0.054* (0.029)	0.054* (0.029)	0.059* (0.032)
nat. avg. latitude (t-avg.) x log_country is landlocked (t-avg.)_std	0.078 (0.063)	0.071 (0.065)	0.064 (0.067)
nat. avg. Longitude (t-avg.) x Certifier's market share among FSC certifiers sq._std	0.106* (0.061)	0.102* (0.061)	0.111* (0.062)
Certifier's years with accreditation sq. x country is landlocked (t-avg.) sq.	−0.004 (0.008)	−0.006 (0.009)	−0.004 (0.009)
avg. dist. to certifier's headquarter by market in 1000 km sq. x country is landlocked (t-avg.) sq._std	−0.160*** (0.062)	−0.147** (0.068)	−0.143** (0.071)
avg. dist. to certifier's headquarter by market in 1000 km sq. x nat. import from fibre furnish in t (t-avg.) sq._std	0.031 (0.044)	0.033 (0.045)	0.040 (0.047)
Expected relative rigor sq. x Exported share of nat. production of woodbased panels (t-avg.) sq._std	−0.035 (0.042)	−0.036 (0.042)	−0.025 (0.046)
Certifier market_group_no_sales_inclNone sq. x log_nat. export of wood chips in 1000 USD per cubic metre (t-avg.)_std	−0.022 (0.033)	−0.024 (0.033)	−0.024 (0.033)
log_Log of nat. score in Corruption Perceptions Index_fsc (t-avg.) x log_dist. to the certifier's closest office (t-avg.)_std	0.047 (0.054)	0.041 (0.055)	0.035 (0.056)
log_nat. export from woodbased panels in 1000 USD per m3 (t-avg.) x log_ch_country_yr_Export_t_Wood_pulp (t-avg.)_std	−0.005 (0.039)	−0.006 (0.039)	−0.002 (0.040)
Certified area in 1000ha (lower bound by cat.) (t-avg.) x _year2018_std	−0.041 (0.042)	−0.044 (0.042)	−0.045 (0.044)
Certified area in 1000ha (lower bound by cat.) (t-avg.) x cb_Certifier 5_std	0.008 (0.054)	0.010 (0.053)	0.012 (0.055)
nat. import from fibre furnish in t (t-avg.) x _year2019_std	0.003 (0.042)	−0.002 (0.043)	−0.004 (0.044)
Country with certifier's closest office and country share the language (t-avg.) x cb_Certifier 5_std	−0.113** (0.057)	−0.116** (0.057)	−0.112* (0.058)
nat. avg. latitude (t-avg.) x cb_Certifier 6_std	−0.156*** (0.044)	−0.160*** (0.044)	−0.172*** (0.046)
nat. avg. latitude (t-avg.) x cb_Certifier 7_std	0.067 (0.067)	0.065 (0.068)	0.065 (0.069)
nat. avg. Longitude (t-avg.) x cb_Certifier 5_std	0.040 (0.048)	0.039 (0.048)	0.040 (0.048)
avg. dist. to certifier's headquarter by market in 1000 km sq. x ___groupEastern_and_Southern_Africa___all_same_std	−0.014 (0.055)	−0.022 (0.058)	−0.024 (0.058)
country is landlocked (t-avg.) sq. x cb_Certifier 5_std	−0.042 (0.046)	−0.033 (0.050)	−0.035 (0.052)
log_Relative rigor (certifier constant) x cb_Certifier 3_std	−0.057** (0.024)	−0.056** (0.024)	−0.084*** (0.029)
log_Relative rigor (certifier constant) x cb_Certifier 4_std	0.040 (0.037)	0.043 (0.037)	0.040 (0.036)
log_Log of nat. score in Corruption Perceptions Index_fsc (t-avg.) x _year2019_std	0.016 (0.033)	0.013 (0.034)	0.026 (0.034)
log_nat. export from roundwood in 1000 USD per m3 (t-avg.) x _year2019_std	0.049 (0.033)	0.052 (0.034)	0.048 (0.034)
log_nat. import from paper products in t (t-avg.) x _year2018_std	−0.016 (0.032)	−0.016 (0.032)	−0.018 (0.033)
Control functions:			
Small certifiers			0.055 (0.102)
Certifier 1			−0.081 (0.118)
Certifier 2			0.159 (0.151)
Certifier 3			0.154 (0.136)
Certifier 4			−0.020 (0.157)
Certifier 5			−0.046 (0.127)
Certifier 6			0.066 (0.144)
Certifier 7			−0.122 (0.399)
Small certifiers			
Inverse Mills Ratio from probit of survey part.		−0.058 (0.121)	−0.030 (0.126)
Constant	1.104*** (0.037)	1.196*** (0.189)	1.130*** (0.200)
Observations (choice situations)	379	379	379
R^2	0.772	0.772	0.776
Adjusted R^2	0.722	0.722	0.719
Residual Std. Error	0.410 (df = 311)	0.410 (df = 310)	0.412 (df = 302)
F Statistic	15.679*** (df = 67; 311)	15.409*** (df = 68; 310)	13.728*** (df = 76; 302)

Notes: Standard errors are robust to heteroskedasticity and serial correlation. They do not account for the variance of the generated regressors used as the control function, yet. Abbreviations: std. (standardized), sq. (squared)., nat. (national), avg. (average), t-avg. (avg. by market), dist. (distance), cat. (category), fun. (function)$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

not include its predictions in the predictions of prices for the audit data.

A positive estimate of σ_{21} would suggest that a positive price shock is associated with a higher shock to the utility of survey participation, and vice versa. In other words, FMUs that pay higher prices are more likely to participate in the survey. This is what I expect since FMUs that are able to pay higher prices than other firms, are likely to have a less restrictive budget constraint for certification, which is likely to be correlated with having time resources to participate in a survey. Another potential driver of a positive association between prices and survey participation is that FMUs' with higher intrinsic motivation for certification are more likely to be willing to pay higher prices for higher quality and are at the same time also more likely to participate in the survey. Such a positive estimate σ_{21} of would suggest that without controlling for selection into the survey, I would overestimate prices, and vice versa. In Table 1.19, I do find a positive, but insignificant estimate of σ_{21}.

1.10 Appendix C: Estimation and results

1.10.1 Stage 3: Violation reporting

1.10.1.1 Selection of baseline violation predictors

To select relevant predictors $\boldsymbol{f}_{it}^{v}$ in the model of violation reporting, I take the following steps.

1. I start by including potentially relevant certifier-invariant audit and FMU characteristics, for example certificate types, categories of forest types and tree species, country-level variables on the trade of various wood products, corruption levels, national FSC standards, number of FSC-certified retailers by country, etc.
2. I remove variables with less than 10% correlation with violation reports or a correlation with a p-value of more than 0.5.
3. I include squares and logs of the remaining variables (if numeric) and include all possible interactions (leaving around 32,000 predictors)
4. I remove near-constant and collinear predictors (leaving around 10,800 predictors).
5. I remove predictors with less than 10% correlation with violation reports or a correlation of less tgab 50% with a p-value of more than 0.1 (leaving around 2,300 predictors).
6. I remove predictors that are more than 90% correlated with other predictors (leaving around 300 predictors).
7. I select relevant predictors by estimating model (1.12) with PPML using LASSO

regularization, excluding the certifier variant variables $\boldsymbol{x^v_{ijt}}$ and certifier fixed effects from the penalization. I use a slightly higher penalty factor than suggested by cross-validation since I do not predict much out-of-sample in certifier-invariant characteristics. The in-sample R-Squared, adapted for the count data model following (Cameron and Windmeijer, 1996), is 0.28.

1.10.1.2 Bias correction for count data models with multinomial selection

This section provides further detail on the derivation of a control function for sample selection in the Stage 3 model.

v_{ijt} are only observed from the certifier j chosen by i in t, i.e. conditional on the choices $y_{ikt} \forall k \in J_t$ and their determinants $\boldsymbol{V_{it}}$:

$$\mathbb{E}[v_{ijt}|\boldsymbol{y_{it}}, \boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}] = \exp(\mu^v_{ijt}) \sum_{k \in \{1,...,J\}} y_{ikt} \mathbb{E}[\exp(\eta_i)|y_{ikt} = 1, \boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}] \quad (1.26)$$

This introduces sample selection bias since $\mathbb{E}[\exp(\eta_i)|\boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, y_{ikt} = 1, \boldsymbol{V_{it}}]$ will not necessarily be constant across the chosen alternatives k. This is the case if η_i is correlated with the differences between the taste shocks $\epsilon^u_{i1t} - \epsilon^u_{ijt}, ...; \epsilon^u_{iJt} - \epsilon^u_{ijt}$ from the demand model. A particularly relevant case of that is if firms that are less compliant due to factors not included in $\boldsymbol{f^v_{it}}$ select into less rigorous certifiers.

To account for the sample selection, I derive and include a control function. Following Heckman (1979), the idea is to control for the mean of the unobservable $\exp(\eta_i)$ within the selected sample by specifying the joint distribution of η_i and the unobservable terms of the demand model. The problem differs from the standard control function approach in two ways. First, the choice is multinomial. Second, the model is a count data model so that the conditional mean of $\exp(\eta_i)$ is not additive. To deal with that, I combine the approaches by Lee (1983) for multinomial selection problems and by Terza (1998) for count data models with binary selection. To the best of my knowledge, I am the first to do so.

Following Lee (1983), I rewrite the multinomial selection in terms of maximum order statistics:

$$y_{ijt} = 1 \text{ iff } V_{ijt} \geq e_{ijt} \text{ where } e_{ijt} \equiv max_{k \neq j}(V_{ikt} + \epsilon^u_{ikt} - \epsilon^u_{ijt}) \quad (1.27)$$

The marginal distribution of e_{ijt} is such that $F_j(V_{ijt}) = s_{ijt}$, the conditional choice probability defined in the paper. e_{ijt} is transformed into a standard normal random variable by defining

$$e^*_{ijt} \equiv G_j(e_{ijt}) \equiv \Phi^{-1}(F_j(e_{ijt})) \tag{1.28}$$

e^*_{ijt} and η_i are jointly normally distributed with zero means, variances 1 and σ_η and correlation coefficient ρ_j. Since $G_j()$ is a strictly increasing function, (1.27) and (1.28) translate into $y_{ijt} = 1$ iff $e^*_{ijt} \leq G_j(V_{ijt})$.

By the Law of Iterated Expectation,

$$\begin{aligned}&\mathbb{E}[\exp(\eta_i)|y_{ikt} = 1, \boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}]\\ &= \mathbb{E}\Big[\mathbb{E}[\exp(\eta_i)|e^*_{ijt}, \boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}]|e^*_{ijt} \leq G_j(V_{ijt}), \boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}\Big]\end{aligned} \tag{1.29}$$

Terza (1998) (Appendix A) shows the following for two random variables η_i and e^*_{ijt} that satisfy the same assumptions as in my case (ϵ and v in his paper), in particular a joint normal distribution with zero means and covariance matrix $\Sigma = \begin{bmatrix} \sigma_\eta & \sigma_\eta\rho_j \\ \sigma_\eta\rho_j & 1 \end{bmatrix}$:

$$\mathbb{E}[\exp(\eta_i)|e^*_{ijt}, \boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}] = \exp(\rho_j\sigma_\eta e^*_{ijt} + \sigma^2_\eta/2(1-\rho^2_j)) \tag{1.30}$$

Consequently

$$\begin{aligned}&\mathbb{E}\Big[\mathbb{E}[\exp(\eta_i)|e^*_{ijt}, \boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}]|e^*_{ijt} \leq G_j(V_{ijt}), \boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}\Big]\\ &= \exp(\sigma^2_\eta/2(1-\rho^2_j))\mathbb{E}\Big[\exp(\theta_j e^*_{ijt})|e^*_{ijt} \leq G_j(V_{ijt}), \boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}\Big]\end{aligned} \tag{1.31}$$

where $\theta_j \equiv \rho_j\sigma_\eta$

The next steps differ follow equation (9) in Terza (1998), with the difference that my condition is $e^*_{ijt} \leq G_j(V_{ijt})$, exploiting that $\Phi(-x) = 1 - \Phi(x)$:

$$
\begin{aligned}
&\mathbb{E}\Big[\exp(\theta_j e^*_{ijt})|e^*_{ijt} \leq G_j(V_{ijt}), \boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}\Big] \\
&= \frac{\mathbb{E}\Big[\exp(\theta_j e^*_{ijt}) \wedge e^*_{ijt} \leq G_j(V_{ijt})|\boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}\Big]}{\Phi(G_j(V_{ijt}))} \\
&= \frac{\int_{-\infty}^{G_j(V_{ijt})} \exp(\theta_j e^*_{ijt})\phi(e^*_{ijt})de^*_{ijt}}{\Phi(G_j(V_{ijt}))} \\
&= \frac{\int_{-\infty}^{G_j(V_{ijt})} \exp(\theta_j e^*_{ijt}) \exp(-\frac{(e^*_{ijt})^2}{2})de^*_{ijt}}{\sqrt{2\pi}\Phi(G_j(V_{ijt}))} \\
&= \frac{\int_{-\infty}^{G_j(V_{ijt})} \exp(-\frac{(e^*_{ijt}-\theta_j)^2}{2} + \frac{\theta_j^2}{2})de^*_{ijt}}{\sqrt{2\pi}\Phi(G_j(V_{ijt}))} \\
&= \frac{\exp(\frac{\theta_j^2}{2}) \int_{-\infty}^{G_j(V_{ijt})} \exp(-\frac{(e^*_{ijt}-\theta_j)^2}{2})de^*_{ijt}}{\sqrt{2\pi}\Phi(G_j(V_{ijt}))} \\
&= \frac{\exp(\frac{\theta_j^2}{2})\sqrt{2\pi}\Phi(G_j(V_{ijt}) - \theta_j)}{\sqrt{2\pi}\Phi(G_j(V_{ijt}))}
\end{aligned} \tag{1.32}
$$

Plugging this into (1.30), one obtains

$$
\mathbb{E}[\exp(\eta_i)|y_{ikt} = 1, \boldsymbol{x^v_{ijt}}, j, \boldsymbol{f^v_{it}}, \boldsymbol{V_{it}}] = \exp(\sigma^2_\eta/2)\frac{\Phi(G_j(V_{ijt}) - \theta_j)}{\Phi(G_j(V_{ijt}))} \tag{1.33}
$$

If $\theta_j = 0$, η_i and e_{ijt} are independent, then the control function $\frac{\Phi(G_j(V_{ijt})-\theta_j)}{\Phi(G_j(V_{ijt}))}$ equals 1 and is thus irrelevant. Otherwise the estimates of $\boldsymbol{\gamma}$ and $\boldsymbol{r}$ would be inconsistent without inclusion of the control function. If θ_j is positive, forestries choosing j have a higher η_i, i.e. more true violations, than the average forestry with the same characteristics, and vice versa. A negative correlation of θ_j with r_j would thus suggest selection on unobservables of less compliant firms tend to select into less rigorous certifiers.

The model with the control function is estimated via Poisson Pseudo Maximum Likelihood as in Egger et al. (2011). Estimates of $G_j(V_{ijt})$ are obtained from the standard normal quantiles of the conditional choice probabilities estimated in a first stage version of the demand model, that replaces violation reports with its predictors.[75]

[75] I estimate a simplified version of the demand model (1.5), but without including expected relative rigor, baseline violations, and prices explicitly. Certifier-market constants capture variation in expected relative rigor across regions, years, and certifiers. Interactions of certifier dummy variables with observed FMU characteristics capture variation in baseline violations.

1.10.1.3 Conditional exogeneity of the accreditation body's inspection assignments

FSC's accreditation body assigns audits without rigid rules but taking several factors into account:[76]

1. Their main focus is to inspect where they suspect leniency. The model controls for various certifier-invariant and certifier-variant determinants of leniency. ω_2 identifies the (expected) effect of inspections on violation reporting if the accreditation body does not assign inspections based on additional information about leniency (which FMUs observe as well).
2. They tend to inspect more in regions in which they expect lower compliance by FMUs. My rich set of controls should capture this variation.
3. They inspect natural forests more often than plantations since they expect greater welfare effects of improved violation reporting in those forests. The plantation dummy in the model controls for that.
4. They tend to inspect larger certifiers more often, but also new entrants and certifiers in markets with new entrants. That is, accreditation inspections may vary with certifiers' market shares. That variation with market shares does not cause endogeneity of accreditation inspections, unless it correlates with determinants of (expected) violation reports. Such correlation is precisely what the observable controls and the control function take care of. As noted above, the control function is particularly reliable for across-market variation, including variation in market shares. I am, thus, not concerned about endogeneity in that dimension.
5. Conditional on the mentioned factors, the accreditation body may assign inspections randomly or based on practical considerations that are unlikely to correlate with violation reports, such as the location of their assessors.

The fourth and fifth factor generate exogenous variation in inspection assignments that identifies ω_2.

[76] This information is based on ASI (2021) and an interview with the accreditation body's staff.

1.10.1.4 Results

Figure 1.23 presents the main results from the estimation of the model of violation reporting: the implied ranking of certifiers in terms of rigor type $\hat{r}_j$. The figure plots the estimates as well as their 95% confidence interval. The orange dot and line come from estimation with control function that accounts for potential selection bias, the blue line is from the estimation without that control function. The certifiers are ranked according to the point estimate from the estimation with this control function.

Figure 1.23: Estimates of rigor types from the Stage 3 model, with and without control function to correct for sample selection bias

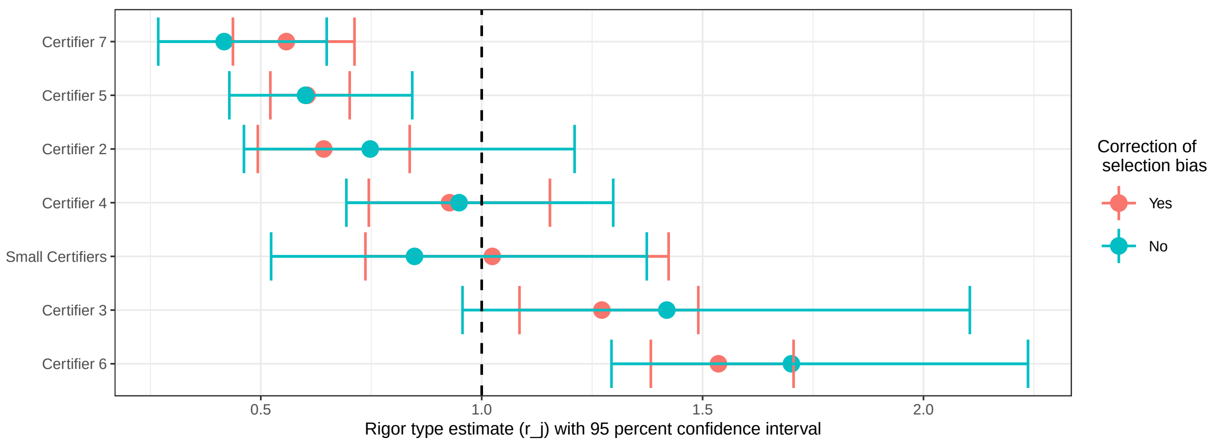


Table 1.20: No. of violation reports (by year)

	Without control function (1)	*With control function* (2)
$Log(r_j)$, rigor type:		
Certifier 2	-0.291 (0.246)	-0.442^{***} (0.135)
Certifier 3	0.350^{*} (0.201)	0.241^{***} (0.081)
Certifier 4	-0.053 (0.160)	-0.076 (0.112)
Certifier 5	-0.509^{***} (0.172)	-0.503^{***} (0.075)
Certifier 6	0.531^{***} (0.140)	0.429^{***} (0.054)
Certifier 7	-0.875^{***} (0.226)	-0.584^{***} (0.125)
Small certifiers	-0.165 (0.246)	0.024 (0.168)
$\boldsymbol{\omega}$*:*		
Audit inspected by accreditation body	0.776^{***} (0.156)	0.793^{***} (0.054)
avg. dist. to certifier's headquarter by market in 1000 km	-0.034^{**} (0.017)	-0.033^{***} (0.007)
Forest nat. CPI inflation rate_X___nat. net export from roundwood in 1000 USD_std	0.015 (0.093)	0.006 (0.037)
nat. forest area annual net change rate (Percent) x nat. certified forest area (1000 ha)_std	-0.075 (0.132)	-0.080^{*} (0.044)
nat. forest area annual net change rate (Percent) x nat. net export from forest products in 1000 USD sq._std	-0.046 (0.097)	-0.029 (0.026)
nat. forest area annual net change rate (Percent) x log_Certified area in 1000 ha_std	-0.108 (0.081)	-0.111^{***} (0.031)
nat. avg. Longitude x nat. export from fibre furnish in 1000 USD per t sq._std	0.041 (0.061)	0.048 (0.032)
nat. net export from industrial roundwood in 1000 USD x nat. net export from wood chips etc. in 1000 USD_std	0.214 (0.135)	0.212^{*} (0.119)
nat. net export from paper products in 1000 USD x Forest has gymnosperms (clade of plants)_std	-0.226^{***} (0.065)	-0.237^{***} (0.030)
nat. net export from roundwood in 1000 USD x nat. forest area annual net change rate (Percent) sq._std	-0.214^{**} (0.094)	-0.209^{***} (0.049)
nat. net export from wood chips etc. in 1000 USD x log_Years since first FSC national standard_std_std	-0.082 (0.097)	-0.090^{**} (0.041)
nat. net export from forest products in 1000 USD x nat. export from industrial roundwood in 1000 USD per m3_std	0.034 (0.118)	0.048 (0.038)
Forest nat. total profit tax (in %)_X___nat. avg. latitude sq._std	-0.100 (0.114)	-0.114^{*} (0.058)
Forest nat. total profit tax (in %)_X___log_Yrs. with FSC cert._std	-0.101 (0.090)	-0.091^{**} (0.036)
Above-ground biomass stock in nat. forests (t per ha) x log_Yrs. with FSC cert._std	-0.037 (0.105)	-0.024 (0.037)
No. of certified downstream firms in country in 2023 x log_Yrs. with FSC cert._std	-0.094 (0.092)	-0.092^{*} (0.048)
No. of promotional license holders in country by year x nat. net export from sawnwood in 1000 USD sq._std	-0.108 (0.098)	-0.118^{**} (0.055)
nat. export from woodbased panels in 1000 USD per m3 x Forest has gymnosperms (clade of plants)_std	-0.232^{***} (0.047)	-0.226^{***} (0.024)
Exported share of nat. production of industrial roundwood x log_No. of certified public projects in country by year_std	-0.159 (0.156)	-0.156^{*} (0.094)
Forest nat. CPI inflation ratesq x nat. net export from wood chips etc. in 1000 USD sq._std	0.199^{***} (0.064)	0.205^{***} (0.052)
nat. avg. latitude sq. x nat. export from sawnwood in 1000 USD per m3 sq._std	-0.055 (0.101)	-0.047 (0.048)
Certified area in 1000 ha sq. x Local currency unit per USD sq._std	-0.028 (0.024)	-0.027^{***} (0.004)
nat. score in Corruption Perceptions Index sq. x Forest nat. total profit tax (in %)sq_std	0.012 (0.106)	-0.014 (0.052)
nat. GDP PC, PPP (2017 1K USD) sq. x Forest nat. total profit tax (in %)sq_std	-0.024 (0.178)	0.024 (0.093)
nat. GDP PC, PPP (2017 1K USD) sq. x log_Local currency unit per USD_std	-0.119 (0.088)	-0.114^{**} (0.055)
Fraction of certified forest's products that is rough wood sq. x log_nat. export from sawnwood in 1000 USD per m3_std	-0.131^{**} (0.053)	-0.128^{***} (0.021)
No. of certified public projects in country by year sq. x Exported share of nat. production of woodbased panels sq._std	-0.071 (0.115)	-0.066^{*} (0.039)
nat. export from roundwood in 1000 USD per m3 sq. x log_No. of certified public projects in country by year_std	-0.052 (0.056)	-0.052 (0.037)
nat. export from sawnwood in 1000 USD per m3 sq. x log_Certified area in 1000 ha_std	0.119 (0.085)	0.117^{***} (0.027)
log_Certified area in 1000 ha x log_Local currency unit per USD_std	0.024 (0.087)	0.026 (0.032)
log_Yrs. with FSC cert. x log_% of publicly owned forest in country in 2015_std	-0.015 (0.106)	-0.004 (0.040)
log_Yrs. with FSC cert. x Forest has angiosperms (clade of plants)_std	-0.035 (0.074)	-0.022 (0.039)
log_Forest nat. total profit tax (in %) x log_nat. certified forest area (1000 ha)_std	-0.031 (0.113)	-0.020 (0.040)
log_No. of certificate members x log_nat. export from roundwood in 1000 USD per m3_std	0.072 (0.045)	0.076^{***} (0.019)
Forest has gymnosperms (clade of plants) x Certified forest has a website_std	0.184^{***} (0.051)	0.179^{***} (0.027)
nat. forest area annual net change rate (Percent) x _ch_forest_typeNatural_std	-0.104^{**} (0.048)	-0.100^{***} (0.021)
No of certified forest's products x _cMexico_std	0.019 (0.056)	0.012 (0.018)
No. of promotional license holders in country by year x _cMexico_std	0.078 (0.050)	0.078^{***} (0.015)
No. of certificate members sq. x _cMexico_std	0.046^{***} (0.010)	0.047^{***} (0.003)
% of publicly owned forest in country in 2015 sq. x _ch_forest_typeNatural_std	0.097^{*} (0.053)	0.097^{***} (0.017)
log_Local currency unit per USD x _ch_forest_typeNatural_std	-0.029 (0.063)	-0.027 (0.027)
log_No. of certificate members x _Southern_and_Southeastern_Asia_std	-0.016 (0.089)	-0.019 (0.014)
log_No of certified forest's products x _Central_America_std	0.040 (0.056)	0.041 (0.030)
(Re-)Certification audit x _ch_forest_typeNatural_std	0.144^{***} (0.041)	0.118^{***} (0.013)
θ_j from control functions:		
Certifier 1		-0.061 (0.040)
Certifier 2		-0.256^{***} (0.091)
Certifier 3		-0.222^{***} (0.057)
Certifier 4		-0.083 (0.064)
Certifier 5		-0.080 (0.054)
Certifier 6		-0.228^{***} (0.033)
Certifier 7		0.545^{***} (0.174)
Small certifiers		0.270 (0.219)
Constant	-0.700^{***} (0.161)	-0.715^{***} (0.065)
Observations (choice situations)	3,810	3,810

Notes: Standard errors are robust to heteroskedasticity and serial correlation. They do not account for the variance of the generated regressors used as the control function, yet. Abbreviations: std. (standardized), sq. (squared)., nat. (national), avg. (average), t-avg. (avg. by market), dist. (distance), cat. (category), fun. (function)$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

1.10.1.5 Discussion: Potential explanations of certifiers' rigor differences

1.10.1.5.1 Summary Why should certifiers have different rigor types? Different ownership types and intrinsic motivation of managers or staff might explain variation in types. Intrinsically motivated staff might also be willing to work for a lower salary than staff that is not motivated by their activity's cause (Preston, 1989). Certifiers with more intrinsically motivated staff might, thereby, be more cost-efficient for the same level of rigor. In this way, such certifiers could survive despite firms' disutility of higher rigor by offering lower prices.

However, different rigor types can appear and survive even if the rigor type is a strategic long-run choice of for-profit certifiers without intrinsic motivation. This paper's model nests this possibility, even though I do not model it explicitly. The next section discusses this in more detail. Intuitively, a profit-maximizing certifier sets its rigor type by equalizing the marginal cost and the marginal benefits of increasing the rigor type. The cost of increasing rigor can come from increased audit costs and from reduced revenues if FMUs prefer leniency. The size of the reduction in revenues may vary across certifiers due to differentiation in other characteristics, potentially explaining heterogeneity in rigor types. Certifiers' benefits from increasing rigor arise from a reduced likelihood of accreditation suspension, which increases the expected future profits. This fact can rationalize the heterogeneity in rigor types due to heterogeneity in profits. Classic reputation theory suggests that larger certifiers with higher profits have reason to be more concerned about losing market access and may, thus, be more rigorous than smaller certifiers (Klein and Leffler, 1981). However, such a pattern is not found here. Certifiers 2 and 7 are very large certifiers both in FSC and other markets. They are the only two publicly traded FSC certifiers. The overall correlation of certification companies' total revenue in 2020 with the relative rigor estimates is negative and statistically significant. Given the low number of certifiers, this finding cannot be interpreted as a pattern in the opposite direction, but, rather, as motivation for further research. An effect of rigor on the likelihood of accreditation suspensions can still rationalize the heterogeneity of rigor if certifiers differ in the degrees to which they discount future profits.

1.10.1.5.2 Rigor as a strategic choice This paper does not consider rigor a strategic variable that certifiers can choose freely in each market. In particular, the model treats the effect of accreditation inspections and the loss of quality control with further distance from the headquarters as deterministic. It fixes other differences in rigor between certifiers to a ratio r_j. Nevertheless, the model allows some discussion of certifiers' potential strategic choice of r_j.

In each market, the profit the certifier expects is $\mathbb{E}[\pi_{jt}|c_{jt}, \bar{c}_t, s_{jt}, j \in J_t]$, defined in model 1.8, if they are accredited. A marginal decrease in r_j would raise that conditionally expected profit in each market through an effect on revenues and an effect on the marginal cost of certification:

$$\frac{\partial \mathbb{E}[\pi_{jt}|c_{jt},\bar{c}_t,s_{jt},j\in J_t]}{\partial r_j} = N_t\bar{c}_t\left[\underbrace{\left(\frac{\partial k_{jt}}{\partial r_j}-1\right)c_{jt}s_{jt}+(k_{jt}-1)\,c_{jt}\frac{\partial s_{jt}}{\partial r_j}}_{\text{Effect on revenues}}+\underbrace{(k_{jt}-1)\frac{\partial c_{jt}}{\partial r_j}s_{jt}}_{\text{Effect on cost}}\right] \tag{1.34}$$

where $s_{jt} = s_{jt}(\boldsymbol{r})$ and $k_{jt} = k_{jt}\Big(s_{jt}(\boldsymbol{r}), s_{jt|FSC}(\boldsymbol{r})\Big)$. I expect reduced rigor to decrease the marginal cost of certification but do not quantify that effect due to a lack of identifying variation. But the demand model estimates allow quantifying the main marginal effect on revenues. Table 1.21 shows the increase in conditionally expected profits due to a decrease in the rigor type r_j by one percentage point. It shows that the willingness to pay for leniency creates strong incentives for certifiers to lower their rigor.

Table 1.21: Certifiers' marginal revenue benefits from leniency

Statistic	Median	Mean	Min	Max
Total	432,641.80	747,577.80	118,426.50	1,793,817.00
By market	82.81	148.11	4.09	791.72

Notes: Predicted reduction in conditional expected profits due to demand effects of an increase of the rigor type by 1 p.p. in 1000 USD

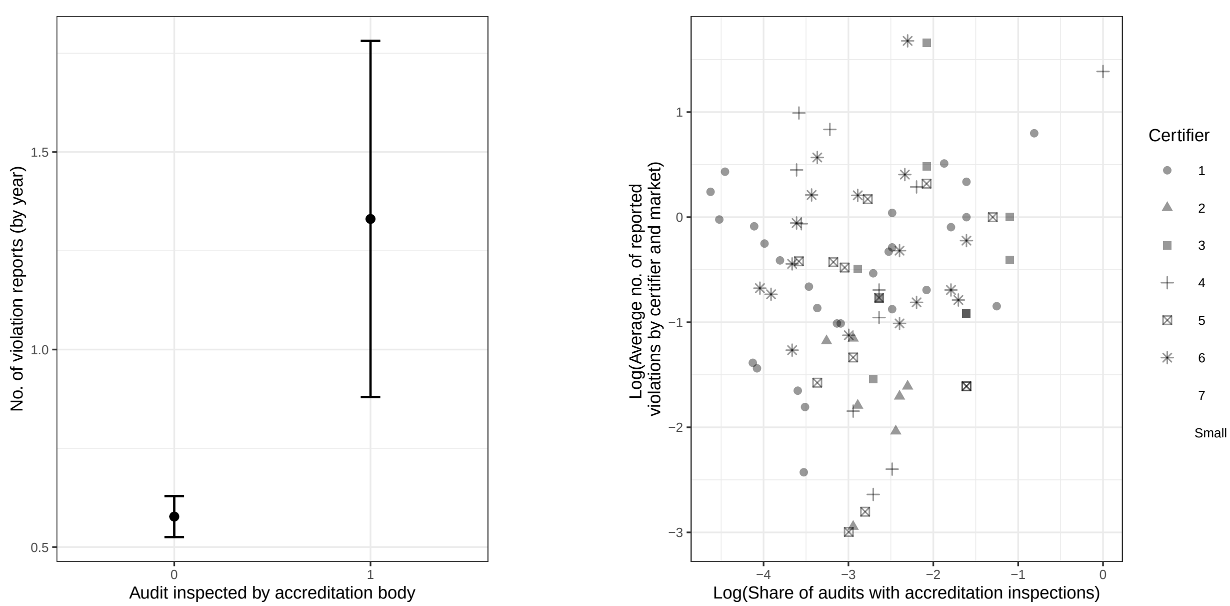
If certifiers choose their rigor types strategically, the cost of decreasing the rigor type below the levels identified as optimal must compensate for the benefits. First, an increased likelihood of accreditation suspension could cause such a cost of leniency. That alone could rationalize the heterogeneity in rigor across certifiers due to heterogeneity in profits. Classic reputation theory suggests that larger certifiers that profit more have reason to be more concerned about losing market access and may thus be more rigorous (Klein and Leffler, 1981). However, there is no pattern in that sense here. Certifiers 2 and 7 are very large certifiers both in FSC and other markets. They are the only two publicly traded certifiers. The overall correlation of certification companies' total revenue in 2020 with the relative rigor estimates is negative and statistically significant. Given the few certifiers, one cannot interpret this as a pattern in the opposite direction, but it encourages further research. An effect of rigor on the likelihood of accreditation suspensions could still rationalize the heterogeneity in rigor if certifiers differ in the degree to which they discount future profits. A second potential explanation for differences in rigor is that certifier staff that is more intrinsically motivated might also be willing to work for a lower salary than staff that is not motivated by their activity's cause (Preston, 1989). Certifiers with more intrinsically motivated staff might thereby be more cost-efficient for the same level of rigor so that they can survive despite firms' disutility of higher rigor by offering lower prices and reducing risks of losing market access due to low rigor in the future. In line with this hypothesis, the certifiers with lower costs according to the supply estimates also tend to be more rigorous.

1.10.2 Stage 2: Demand

1.10.2.1 Identifying variation in the likelihood of accreditation inspections

To identify FMUs' willingness to pay for leniency, this paper exploits variation in the share of audits inspected by the accreditation body, by certifier and market, henceforth the *likelihood of accreditation inspections*. Figure 1.24 (a) shows that the average number of violation reports in audits with accreditation inspections is significantly higher than without. Targeting high-risk markets and FMUs might confound that effect. The empirical model aims to control for that. Figure 1.24 (b) plots the likelihood of accreditation inspections in logs for positive observations. The plot shows substantial variation across certifiers and markets, which I exploit in the empirical model.

Figure 1.24: Association of average violation reports with accreditation inspections



(a) Across audits with and without accreditation inspection (Means with 95 percent confidence interval)

(b) Across certifier-market shares of accreditation inspections for non-zero observations

1.10.2.2 Coefficient tables

Table 1.22: Preference estimates at the certifier-market-level

	OLS	*2SLS*
	(1)	(2)
$\bar{\alpha}$: Price at certifier-market-level in 1K USD	−0.069*** (0.003)	−0.770*** (0.178)
$\bar{\beta}^r$: Expected relative rigor	−1.367*** (0.046)	−17.366*** (1.132)
$\boldsymbol{\beta^x}$:		
First year certifier is available	0.294*** (0.019)	1.698*** (0.195)
Average distance to certifier's headquarter by market in 1000 km	−0.025*** (0.001)	−0.139*** (0.016)
$\boldsymbol{\xi}$:		
Certifier 2	−1.274*** (0.015)	−3.492*** (0.130)
Certifier 3	−0.483*** (0.012)	0.709*** (0.089)
Certifier 4	−0.818*** (0.011)	−0.624*** (0.139)
Certifier 5	−0.928*** (0.012)	−3.756*** (0.210)
Certifier 6	−0.621*** (0.013)	3.398*** (0.415)
Certifier 7	−1.063*** (0.011)	−3.476*** (0.147)
Certifier Small Certifiers	−0.326*** (0.016)	0.683*** (0.151)
FSC x Brazil x 2015	0.101*** (0.014)	−0.427*** (0.099)
FSC x Brazil x 2016	0.139*** (0.016)	−0.425*** (0.138)
FSC x Brazil x 2017	−0.502*** (0.024)	−1.240*** (0.135)
FSC x Brazil x 2018	0.364*** (0.018)	−0.150 (0.142)
FSC x Brazil x 2019	0.359*** (0.036)	0.530*** (0.060)
FSC x Central_America x 2015	−0.115*** (0.016)	−0.727*** (0.192)
FSC x Central_America x 2016	0.436*** (0.020)	−0.162 (0.154)
FSC x Central_America x 2017	−0.055** (0.024)	−0.626*** (0.169)
FSC x Central_America x 2018	0.949*** (0.015)	0.011 (0.236)
FSC x Central_America x 2019	0.390*** (0.023)	−0.167 (0.141)
FSC x Eastern_Asia_excl_China x 2015	−0.176*** (0.046)	−0.468*** (0.116)
FSC x Eastern_Asia_excl_China x 2016	0.243*** (0.025)	−0.010 (0.106)
FSC x Eastern_Asia_excl_China x 2017	1.105*** (0.029)	0.826*** (0.104)
FSC x Eastern_Asia_excl_China x 2018	0.707*** (0.029)	1.858*** (0.222)
FSC x Eastern_Asia_excl_China x 2019	−0.914*** (0.018)	−0.654*** (0.146)
FSC x Eastern_Europe_excl_CIS x 2015	−0.616*** (0.012)	0.937*** (0.191)
FSC x Eastern_Europe_excl_CIS x 2016	0.224*** (0.026)	0.195*** (0.051)
FSC x Eastern_Europe_excl_CIS x 2017	0.146*** (0.024)	0.296*** (0.064)
FSC x Eastern_Europe_excl_CIS x 2018	0.481*** (0.020)	0.023 (0.068)
FSC x Eastern_Europe_excl_CIS x 2019	0.223*** (0.020)	0.509*** (0.079)
FSC x Eastern_and_Southern_Africa x 2015	0.668*** (0.029)	−0.613*** (0.200)
FSC x Eastern_and_Southern_Africa x 2016	−0.199*** (0.034)	−1.035*** (0.152)
FSC x Eastern_and_Southern_Africa x 2017	−0.493*** (0.025)	−1.303*** (0.157)
FSC x Eastern_and_Southern_Africa x 2018	0.720*** (0.042)	0.150 (0.188)
FSC x Eastern_and_Southern_Africa x 2019	1.040*** (0.040)	0.395*** (0.124)
FSC x Northern_America x 2015	−0.184*** (0.023)	−0.147*** (0.045)
FSC x Northern_America x 2016	−0.778*** (0.017)	−0.865*** (0.055)
FSC x Northern_America x 2017	−1.032*** (0.015)	−1.097*** (0.045)
FSC x Northern_America x 2018	−0.654*** (0.020)	−0.916*** (0.138)
FSC x Northern_America x 2019	−0.970*** (0.016)	−0.716*** (0.042)
FSC x Oceania x 2015	−0.434*** (0.034)	1.115*** (0.320)
FSC x Oceania x 2016	0.679*** (0.033)	1.022*** (0.085)
FSC x Oceania x 2017	−0.476*** (0.051)	0.157* (0.082)
FSC x Oceania x 2018	0.637*** (0.036)	0.863*** (0.082)
FSC x Oceania x 2019	0.323*** (0.042)	1.170*** (0.192)
FSC x South_America_excl_Brazil x 2015	−0.162*** (0.015)	−0.916*** (0.175)
FSC x South_America_excl_Brazil x 2016	−0.104*** (0.018)	0.129 (0.137)
FSC x South_America_excl_Brazil x 2017	0.095*** (0.018)	−0.424** (0.187)
FSC x South_America_excl_Brazil x 2018	0.228*** (0.018)	−0.548*** (0.196)
FSC x South_America_excl_Brazil x 2019	0.157*** (0.019)	−0.240** (0.112)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2016	−0.904*** (0.026)	−1.192*** (0.037)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2017	−0.493*** (0.026)	0.911*** (0.140)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2018	−0.019* (0.011)	0.077* (0.043)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2019	−0.112*** (0.019)	0.453*** (0.095)
FSC x Southern_and_Southeastern_Asia x 2015	0.234*** (0.012)	0.564*** (0.067)
FSC x Southern_and_Southeastern_Asia x 2016	0.542*** (0.017)	0.987*** (0.117)
FSC x Southern_and_Southeastern_Asia x 2017	−0.470*** (0.016)	−0.466*** (0.040)
FSC x Southern_and_Southeastern_Asia x 2018	1.001*** (0.013)	1.472*** (0.059)
FSC x Southern_and_Southeastern_Asia x 2019	1.031*** (0.016)	3.344*** (0.513)
Constant	−1.554*** (0.026)	7.999*** (0.913)
Observations	34,794	34,794
Adjusted R^2	0.649	−1.006
Residual Std. Error (df = 34733)	0.484	1.156
F Statistic	1,072.535*** (df = 60; 34733)	
Wu-Hausman stat.		723.996*** (df=2,34731)
Weak IV stat. (Expected relative rigor)		343.283*** (df=2,34733)
Weak IV stat. (Price at certifier-market-level in 1K USD)		77.239*** (df=2,34733)

Notes: Heteroscedasticity robust standard errors. Standard errors are not yet corrected for the use of generated regressors and regressands. The dependent variable are the estimates of mean utility by certifier and market, obtained from Maximum Likelihood estimation of the certifier-market constants in the nested logit model of FMUs' certifier choice. The outside option in the nested logit includes both dropping out of FSC certification or not getting certified yet. Dropping out is permanent in the data, such that only the first year after drop-out is included. In the 2SLS regression here, the outside option is excluded since its mean utility is normalized to zero for all markets in the nested logit. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 1.23: First stage of IV: Price prediction

	(1)	(2)
***β^x*:**		
First year certifier is available	−0.154*** (0.024)	−0.131*** (0.025)
Average distance to certifier's headquarter by market in 1000 km	0.148*** (0.003)	0.147*** (0.003)
Excluded instruments:		
Close-by rivals in terms of certifier's experience		−0.119*** (0.006)
Close-by rivals in terms of certifier's experience x Forest country's longitude (mean by market)		−0.000** (0.000)
***ξ*:**		
Certifier 2	0.222*** (0.010)	0.345*** (0.012)
Certifier 3	−0.866*** (0.012)	−0.729*** (0.014)
Certifier 4	0.905*** (0.011)	1.212*** (0.019)
Certifier 5	−0.137*** (0.017)	0.004 (0.018)
Certifier 6	1.085*** (0.024)	1.209*** (0.023)
Certifier 7	0.192*** (0.016)	0.347*** (0.019)
Small Certifiers	0.532*** (0.011)	0.791*** (0.020)
FSC x Brazil x 2015	−0.461*** (0.045)	−0.614*** (0.046)
FSC x Brazil x 2016	−0.584*** (0.045)	−0.673*** (0.045)
FSC x Brazil x 2017	−0.619*** (0.045)	−0.703*** (0.046)
FSC x Brazil x 2018	−0.774*** (0.045)	−0.858*** (0.046)
FSC x Brazil x 2019	−0.304*** (0.045)	−0.388*** (0.046)
FSC x Central_America x 2015	−1.112*** (0.044)	−1.324*** (0.046)
FSC x Central_America x 2016	−0.850*** (0.045)	−1.062*** (0.047)
FSC x Central_America x 2017	−0.959*** (0.047)	−1.172*** (0.049)
FSC x Central_America x 2018	−1.308*** (0.044)	−1.521*** (0.046)
FSC x Central_America x 2019	−0.760*** (0.048)	−0.973*** (0.050)
FSC x Eastern_Asia_excl_China x 2015	−0.181* (0.094)	−0.298*** (0.094)
FSC x Eastern_Asia_excl_China x 2016	−0.186** (0.095)	−0.302*** (0.094)
FSC x Eastern_Asia_excl_China x 2017	−0.164* (0.093)	−0.281*** (0.093)
FSC x Eastern_Asia_excl_China x 2018	−0.153 (0.095)	−0.270*** (0.095)
FSC x Eastern_Asia_excl_China x 2019	0.398*** (0.129)	0.275** (0.128)
FSC x Eastern_Europe_excl_CIS x 2015	0.098* (0.052)	−0.041 (0.052)
FSC x Eastern_Europe_excl_CIS x 2016	0.136*** (0.052)	−0.002 (0.053)
FSC x Eastern_Europe_excl_CIS x 2017	0.120** (0.052)	−0.018 (0.053)
FSC x Eastern_Europe_excl_CIS x 2018	−0.091* (0.048)	−0.148*** (0.049)
FSC x Eastern_Europe_excl_CIS x 2019	0.346*** (0.053)	0.196*** (0.055)
FSC x Eastern_and_Southern_Africa x 2015	−0.565*** (0.097)	−0.819*** (0.097)
FSC x Eastern_and_Southern_Africa x 2016	−0.649*** (0.097)	−0.895*** (0.097)
FSC x Eastern_and_Southern_Africa x 2017	−0.737*** (0.095)	−0.983*** (0.095)
FSC x Eastern_and_Southern_Africa x 2018	−0.870*** (0.096)	−1.116*** (0.096)
FSC x Eastern_and_Southern_Africa x 2019	−0.483*** (0.102)	−0.729*** (0.101)
FSC x Northern_America x 2015	−0.235*** (0.045)	−0.307*** (0.046)
FSC x Northern_America x 2016	−0.252*** (0.045)	−0.324*** (0.046)
FSC x Northern_America x 2017	−0.203*** (0.045)	−0.276*** (0.046)
FSC x Northern_America x 2018	−0.824*** (0.043)	−0.897*** (0.044)
FSC x Northern_America x 2019	0.114** (0.047)	0.042 (0.048)
FSC x Oceania x 2015	1.538*** (0.106)	1.425*** (0.108)
FSC x Oceania x 2016	−0.242*** (0.075)	−0.355*** (0.077)
FSC x Oceania x 2017	−0.097 (0.076)	−0.210*** (0.078)
FSC x Oceania x 2018	−0.307*** (0.073)	−0.419*** (0.075)
FSC x Oceania x 2019	0.761*** (0.096)	0.648*** (0.098)
FSC x South_America_excl_Brazil x 2015	−0.933*** (0.049)	−0.952*** (0.051)
FSC x South_America_excl_Brazil x 2016	−1.023*** (0.049)	−1.043*** (0.050)
FSC x South_America_excl_Brazil x 2017	−1.115*** (0.048)	−1.134*** (0.050)
FSC x South_America_excl_Brazil x 2018	−1.081*** (0.050)	−1.100*** (0.051)
FSC x South_America_excl_Brazil x 2019	−0.606*** (0.051)	−0.625*** (0.052)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2016	−0.001 (0.049)	−0.001 (0.049)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2017	0.019 (0.047)	0.133*** (0.047)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2018	−0.205*** (0.046)	−0.088* (0.046)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2019	0.248*** (0.050)	0.296*** (0.049)
FSC x Southern_and_Southeastern_Asia x 2015	0.284*** (0.049)	0.333*** (0.052)
FSC x Southern_and_Southeastern_Asia x 2016	0.627*** (0.049)	0.676*** (0.053)
FSC x Southern_and_Southeastern_Asia x 2017	0.005 (0.048)	0.055 (0.051)
FSC x Southern_and_Southeastern_Asia x 2018	0.112** (0.049)	0.162*** (0.052)
FSC x Southern_and_Southeastern_Asia x 2019	2.745*** (0.062)	2.794*** (0.066)
Constant	2.173*** (0.035)	2.410*** (0.038)
Observations	34,794	34,794
Adjusted R^2	0.610	0.611
Residual Std. Error	0.833 (df = 34735)	0.831 (df = 34733)
F Statistic	937.949*** (df = 58; 34735)	913.239*** (df = 60; 34733)

Notes: First stage of 2SLS regression of estimates of mean utility by certifier and market (obtained from MLE of the certifier-market constants in the nested logit demand model) on prices and other characteristics. The dependent variable in the first stage is the endogeneous certifier-market level prediction of prices. Standard errors are heteroscedasticity and serial correlation robust, but not yet corrected for the use of generated regressors and regressands.$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 1.24: First stage of IV: Prediction of expected relative rigor

	(1)	(2)
β^x:		
First year certifier is available	0.094*** (0.010)	0.095*** (0.009)
Average distance to certifier's headquarter by market in 1000 km	−0.014*** (0.000)	−0.013*** (0.000)
Excluded instruments:		
Close-by rivals in terms of certifier's experience		0.013*** (0.001)
Close-by rivals in terms of certifier's experience x Forest country's longitude (mean by market)		−0.000*** (0.000)
ξ:		
Certifier 2	−0.148*** (0.001)	−0.163*** (0.001)
Certifier 3	0.113*** (0.001)	0.095*** (0.001)
Certifier 4	−0.028*** (0.001)	−0.062*** (0.002)
Certifier 5	−0.171*** (0.001)	−0.188*** (0.001)
Certifier 6	0.204*** (0.001)	0.188*** (0.001)
Certifier 7	−0.159*** (0.001)	−0.178*** (0.001)
Small Certifiers	0.040*** (0.002)	0.004* (0.002)
FSC x Brazil x 2015	−0.013*** (0.002)	−0.002 (0.002)
FSC x Brazil x 2016	−0.010*** (0.003)	−0.008*** (0.003)
FSC x Brazil x 2017	−0.019*** (0.002)	−0.018*** (0.002)
FSC x Brazil x 2018	0.002 (0.002)	0.003 (0.002)
FSC x Brazil x 2019	0.024*** (0.003)	0.026*** (0.003)
FSC x Central_America x 2015	0.010*** (0.002)	0.023*** (0.002)
FSC x Central_America x 2016	−0.000 (0.002)	0.012*** (0.002)
FSC x Central_America x 2017	0.006*** (0.002)	0.019*** (0.002)
FSC x Central_America x 2018	−0.001 (0.002)	0.011*** (0.002)
FSC x Central_America x 2019	−0.001 (0.002)	0.011*** (0.002)
FSC x Eastern_Asia_excl_China x 2015	−0.010*** (0.003)	0.019*** (0.003)
FSC x Eastern_Asia_excl_China x 2016	−0.008*** (0.003)	0.021*** (0.003)
FSC x Eastern_Asia_excl_China x 2017	−0.010*** (0.003)	0.019*** (0.003)
FSC x Eastern_Asia_excl_China x 2018	0.079*** (0.012)	0.108*** (0.012)
FSC x Eastern_Asia_excl_China x 2019	−0.001 (0.003)	0.027*** (0.003)
FSC x Eastern_Europe_excl_CIS x 2015	0.093*** (0.009)	0.111*** (0.009)
FSC x Eastern_Europe_excl_CIS x 2016	−0.008*** (0.002)	0.010*** (0.002)
FSC x Eastern_Europe_excl_CIS x 2017	0.004** (0.002)	0.022*** (0.002)
FSC x Eastern_Europe_excl_CIS x 2018	−0.025*** (0.003)	−0.015*** (0.003)
FSC x Eastern_Europe_excl_CIS x 2019	0.003 (0.002)	0.022*** (0.002)
FSC x Eastern_and_Southern_Africa x 2015	−0.055*** (0.005)	−0.027*** (0.005)
FSC x Eastern_and_Southern_Africa x 2016	−0.024*** (0.003)	0.004 (0.003)
FSC x Eastern_and_Southern_Africa x 2017	−0.018*** (0.004)	0.010*** (0.004)
FSC x Eastern_and_Southern_Africa x 2018	0.003 (0.002)	0.031*** (0.003)
FSC x Eastern_and_Southern_Africa x 2019	−0.019*** (0.003)	0.009** (0.004)
FSC x Northern_America x 2015	0.013*** (0.002)	0.006*** (0.002)
FSC x Northern_America x 2016	0.006** (0.002)	−0.001 (0.002)
FSC x Northern_America x 2017	0.005** (0.002)	−0.002 (0.002)
FSC x Northern_America x 2018	0.020*** (0.002)	0.013*** (0.002)
FSC x Northern_America x 2019	0.011*** (0.002)	0.004** (0.002)
FSC x Oceania x 2015	0.029*** (0.004)	0.060*** (0.004)
FSC x Oceania x 2016	0.032*** (0.003)	0.062*** (0.004)
FSC x Oceania x 2017	0.044*** (0.004)	0.074*** (0.004)
FSC x Oceania x 2018	0.028*** (0.005)	0.058*** (0.005)
FSC x Oceania x 2019	0.020*** (0.003)	0.050*** (0.003)
FSC x South_America_excl_Brazil x 2015	−0.006*** (0.002)	−0.020*** (0.002)
FSC x South_America_excl_Brazil x 2016	0.059*** (0.004)	0.046*** (0.004)
FSC x South_America_excl_Brazil x 2017	0.016*** (0.003)	0.003 (0.003)
FSC x South_America_excl_Brazil x 2018	−0.001 (0.002)	−0.015*** (0.002)
FSC x South_America_excl_Brazil x 2019	0.002 (0.002)	−0.012*** (0.002)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2016	−0.018*** (0.002)	−0.018*** (0.002)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2017	0.087*** (0.005)	0.075*** (0.005)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2018	0.015*** (0.003)	0.003 (0.003)
FSC x Southern_Europe_incl_Turkey_excl_Italy x 2019	0.024*** (0.004)	0.020*** (0.003)
FSC x Southern_and_Southeastern_Asia x 2015	0.008*** (0.002)	0.026*** (0.003)
FSC x Southern_and_Southeastern_Asia x 2016	0.000 (0.002)	0.018*** (0.003)
FSC x Southern_and_Southeastern_Asia x 2017	0.000 (0.002)	0.018*** (0.003)
FSC x Southern_and_Southeastern_Asia x 2018	0.025*** (0.002)	0.042*** (0.003)
FSC x Southern_and_Southeastern_Asia x 2019	0.024*** (0.003)	0.042*** (0.003)
Constant	0.502*** (0.002)	0.477*** (0.002)
Observations	34,794	34,794
Adjusted R^2	0.884	0.886
Residual Std. Error	0.056 (df = 34735)	0.055 (df = 34733)
F Statistic	4,553.911*** (df = 58; 34735)	4,500.315*** (df = 60; 34733)

Notes: First stage of 2SLS regression of estimates of mean utility by certifier and market (obtained from MLE of the certifier-market constants in the nested logit demand model) on prices and other characteristics. The dependent variable in the first stage is the endogeneous certifier-market level prediction of expected relative rigor. Standard errors are heteroscedasticity and serial correlation robust, but not yet corrected for the use of generated regressors and regressands.$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

1.10.2.3 Robustness and tests

I test the relevance of the differentiation instruments used in the demand model in two ways. First, Table 1.7 shows the statistics and significance levels of tests for weak instruments, as provided by the *ivreg* package (Zeileis et al., 2023). For both prices and expected relative rigor, these tests allow me to reject the null hypothesis that the differentiation instruments are jointly insignificant in the first stage, i.e., in predicting the endogenous variable conditional

Table 1.25: Estimates of heterogeneity in preferences across FMU types (Nested Logit)

$\tilde{\alpha}$: Price at certifier-market-level in 1K USD X ...	
Cost factor	0.020*** (0.008)
$\tilde{\boldsymbol{\beta}}^{\boldsymbol{r}}$: Expected relative rigor X ...	
Expected baseline violations	0.082 (0.097)
≥ 5 violations reports expected from most rigorous certifier	0.519 (1.290)
Forest chose same certifier last year	2.496*** (0.345)
Forest chose same certifier last year x Yrs. with FSC cert.	0.051*** (0.014)
Log of forest's country's score in Corruption Perceptions Index x FSC	0.399 (0.316)
Forest has angiosperms (clade of plants) x FSC	0.167 (0.190)
Forest has gymnosperms (clade of plants) x FSC	0.513*** (0.145)
Forest has myrtaceae (clade of plants) x FSC	0.141 (0.147)
Country's export value of wood chips in 1000 USD per cubic metre	−0.273*** (0.097)
Certifier's experience in forest's first certification year	0.046*** (0.017)
Forest chose same certifier last year x Recertification year x FSC	−0.076 (0.166)
FSC Certifier has office in forest's country	0.549*** (0.141)
Is first FSC certifier (entry cost)	−2.457*** (0.281)
$\boldsymbol{\xi}$:	
Group certificate (vs. individual) X ...	
Certifier 1	−0.046 (0.149)
Certifier 2	−0.373 (0.328)
Certifier 3	−0.098 (0.287)
Certifier 4	0.205 (0.254)
Certifier 5	−0.004 (0.206)
Certifier 6	0.135 (0.191)
Certifier 7	−0.309 (0.280)
Small Certifiers	−0.598** (0.289)
Forest is classified as plantation X ...	
Certifier 1	−0.568*** (0.193)
Certifier 2	0.801*** (0.300)
Certifier 3	−0.220 (0.351)
Certifier 4	0.631*** (0.232)
Certifier 5	−0.481* (0.265)
Certifier 6	0.345 (0.212)
Certifier 7	−0.009 (0.254)
Small Certifiers	−0.317 (0.276)
λ:	
Within FSC nest correlation	0.464*** (0.063)
Certifier-market FE	Yes
Observations (choice situations)	6,250
Log Likelihood	−2,744.536

Notes: Maximum Likelihood estimation of nested logit choice model with R package mlogit. Heteroscedasticity and serial correlation robust standard errors. Standard errors are not yet corrected for the use of generated regressors. The outside option includes both dropping out of FSC certification or not getting certified yet. Dropping out is permanent in the data, such that only the first year after drop-out is included. The mean utility of the outside option is normalized to zero for all markets.$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

on all other controls. Second, I follow Stock and Yogo (2005) to test the relevance of the instruments in a way that is adapted to the case of multiple endogenous variables. That is, I test for the joint significance of the differentiation instruments in both first stages jointly, using the adapted critical values derived by Stock and Yogo (2005). The corresponding F test statistic is 177.577, which is much larger than 8.2, the critical value for 2 endogenous variables and 2 instruments (Hansen, 2021). I, thus, reject the null hypothesis of weak instruments. A caveat is that both tests assume homoskedasticity. However, to my knowledge, no adapted heteroskedasticity-robust test zxists for weak instruments in the case of multiple endogenous variables (Andrews, 2018).

Table 1.26: Preference estimates at the certifier-market-levelwith region and year instead of market (region-year) fixed effects

	OLS	*2SLS*
	(1)	(2)
$\bar{\alpha}$: Price at certifier-market-level in 1K USD	-0.077^{***} (0.004)	-1.098^{***} (0.195)
$\bar{\beta}^r$: Expected relative rigor	-1.319^{***} (0.054)	-13.160^{***} (0.748)
$\boldsymbol{\beta^x}$:		
First year certifier is available	0.448^{***} (0.019)	1.659^{***} (0.171)
Average distance to certifier's headquarter by market in 1000 km	-0.024^{***} (0.001)	-0.126^{***} (0.010)
$\boldsymbol{\xi}$:		
Certifier 2	-1.235^{***} (0.017)	-2.540^{***} (0.086)
Certifier 3	-0.490^{***} (0.014)	-0.052 (0.131)
Certifier 4	-0.786^{***} (0.013)	-0.317^{**} (0.148)
Certifier 5	-0.931^{***} (0.015)	-3.105^{***} (0.145)
Certifier 6	-0.615^{***} (0.015)	2.818^{***} (0.327)
Certifier 7	-1.025^{***} (0.014)	-2.417^{***} (0.091)
Certifier Small Certifiers	-0.321^{***} (0.018)	1.007^{***} (0.192)
Central_America	0.176^{***} (0.014)	0.058 (0.035)
Eastern_and_Southern_Africa	0.205^{***} (0.025)	-0.496^{***} (0.105)
Eastern_Asia_excl_China	0.114^{***} (0.028)	0.413^{***} (0.079)
Eastern_Europe_excl_CIS	-0.005 (0.017)	0.280^{***} (0.054)
Northern_America	-0.796^{***} (0.014)	-1.198^{***} (0.100)
Oceania	0.096^{***} (0.024)	1.682^{***} (0.266)
South_America_excl_Brazil	-0.076^{***} (0.012)	-0.134^{***} (0.045)
Southern_and_Southeastern_Asia	0.334^{***} (0.012)	1.296^{***} (0.161)
Southern_Europe_incl_Turkey_excl_Italy	-0.402^{***} (0.015)	0.230^{**} (0.093)
2016	0.022^{**} (0.010)	-0.002 (0.020)
2017	-0.262^{***} (0.010)	-0.231^{***} (0.022)
2018	0.402^{***} (0.009)	0.265^{***} (0.036)
2019	0.235^{***} (0.011)	1.081^{***} (0.158)
Constant	-1.375^{***} (0.035)	6.705^{***} (0.710)
Observations	34,794	34,794
Adjusted R^2	0.484	-1.482
Residual Std. Error (df = 34769)	0.575	1.262
F Statistic	$1,359.602^{***}$ (df = 24; 34769)	
Wu-Hausman stat.		454.37^{***} (df=2,34767)
Weak IV stat. (Expected relative rigor)		396.568^{***} (df=2,34769)
Weak IV stat. (Price at certifier-market-level in 1K USD)		49.082^{***} (df=2,34769)

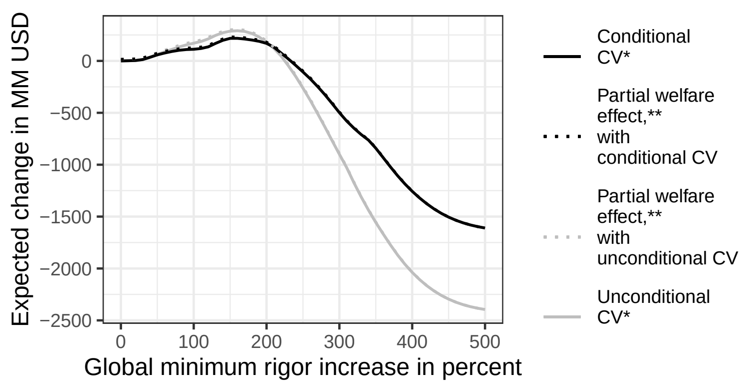
Notes: Heteroscedasticity robust standard errors. Standard errors are not yet corrected for the use of generated regressors and regressands. The dependent variable are the estimates of mean utility by certifier and market, obtained from Maximum Likelihood estimation of the certifier-market constants in the nested logit model of FMUs' certifier choice. The outside option in the nested logit includes both dropping out of FSC certification or not getting certified yet. Dropping out is permanent in the data, such that only the first year after drop-out is included. In the 2SLS regression here, the outside option is excluded since its mean utility is normalized to zero for all markets in the nested logit.$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

1.11 Appendix D: Counterfactual Analysis

1.11.1 Robustness of the measure of consumer valuation

In the counterfactual analysis, I assume that consumer valuation is a linear function of the amount of certified wood and the weighted sum of violation reports conditional on certification. This assumption can be debated in two dimensions. First, there is no empirical evidence to microfound the functional form. The marginal benefits for consumers and society of additional violation reports may as well decrease for higher quality, such that the plotted welfare estimates would overestimate true benefits. Second, the assumption implies that consumer valuation depends on the amount of certified wood and certification quality. This seems to be the most plausible assumption since this type of conditional valuation relates more to the traditional way of comparing products of varying quality. Nevertheless, other heuristics of valuing certification are possible.

Figure 1.25: Expected changes following direct shifts in minimum rigor



Notes: *Consumer valuation of FSC certification, assuming valuation at 10 percent of wood prices for 2017 baseline quality and constant wood prices. **Sum of CV, FMU and certifier surplus if price premia at baseline CV. Excl. externalities not valued by consumers.

In Figure 1.25, I consider the possibility of consumer valuation varying with the quality among all once-certified FMUs, i.e., the total impact of FSC in the industry, *vsum*, but not directly with the amount of certified wood. In other words, consumer valuation might work more like donations rather than price premia. Figure 1.25 shows changes in such unconditional consumer valuation and corresponding welfare effects in gray. For comparison, the black lines show the results using conditional consumer valuation presented in the paper. The effects are qualitatively similar. Unconditional consumer valuation seems to amplify the changes suggested by conditional consumer valuation mostly.

1.11.2 Back-of-the-envelope calculation of welfare benefit of reduced tree cover loss through FSC violation reports

An estimate of FSC's effect on deforestation allows me to obtain an idea of the benefits from corresponding violations: for the Congo Basin, Tritsch et al. (2020) estimate an average reduction in deforestation by 514 ha per FMU from 2000 to 2010 due to FSC certification. CO_2 emissions from deforestation vary largely across contexts. If the difference in CO_2 emissions between forested and deforested areas in the Congo Basin is similar to that in Brazil, I can use the average from there, 80 tCO_2/ha (Souza-Rodrigues, 2019). Using a carbon cost of 31 USD/tCO_2 (Nordhaus, 2017), the total social benefit of FSC certification per year and FMU would be 127,472 USD. Using the average number of violation reports per FMU and the fraction of those relating to tree cover loss in a randomly drawn sample, one violation report's carbon benefit is estimated at roughly 2 million USD, while this paper's measure of consumer valuation per violation report is 2.79 million USD. The derived estimate of carbon benefits should be interpreted cautiously, as the calculation is based on a mix of different sources.

1.11.3 Suspending the accreditation of lenient certifiers - The role of price changes

Table 1.27: Expected changes in violation reports

Counterfactual	Conditional on certification in numbers	Due to drop-outs in numbers	Total in numbers	Total in percent
(1) Targeting the most lenient certifier				
Accreditation suspension				
Effect through				
- Certifier choices and prices	62.36	-47.58	14.77	2.51
- Certifier choices only	63.90	-46.06	17.83	3.03
Equivalent minimum rigor shift				
Effect through				
- Certifier choices and prices	18.62	-14.34	4.28	0.73
- Certifier choices only	19.02	-14.45	4.56	0.78
(2) Targeting the two most lenient certifiers				
Accreditation suspension				
Effect through				
- Certifier choices and prices	120.59	-98.68	21.91	3.72
- Certifier choices only	122.68	-92.14	30.54	5.19
Equivalent minimum rigor shift				
Effect through				
- Certifier choices and prices	64.65	-45.26	19.39	3.29
- Certifier choices only	65.23	-44.33	20.89	3.55
(3) Targeting the three most lenient certifiers				
Accreditation suspension				
Effect through				
- Certifier choices and prices	145.39	-116.74	28.65	4.87
- Certifier choices only	146.18	-106.97	39.21	6.66
Equivalent minimum rigor shift				
Effect through				
- Certifier choices and prices	133.10	-86.56	46.54	7.90
- Certifier choices only	134.63	-80.55	54.08	9.19

Notes: Changes in counterfactual scenarios compared to baseline. Accreditation withdrawal is implemented by removing the corresponding certifier from the choiceset. Direct shift of minimum rigor is implemented by shifting the targeted certifiers to the next most lenient certifier's rigor in each market.

Table 1.28: Expected change in surplus from FSC certification

Counterfactual	FMUs in MM USD	Targeted certifiers in MM USD	Untargeted certifiers in MM USD	Consumers in MM USD	Total in MM USD	Total in percent
(1) Targeting the most lenient certifier						
Accreditation suspension						
Effect through						
- Certifier choices and prices	-0.51	-1.53	1.20	23.91	23.06	150.96
- Certifier choices only	-0.49	-1.53	1.18	31.76	30.92	202.42
Equivalent minimum rigor shift						
Effect through						
- Certifier choices and prices	-0.10	-0.21	0.11	16.01	15.81	103.51
- Certifier choices only	-0.10	-0.21	0.11	16.34	16.14	105.69
(2) Targeting the two most lenient certifiers						
Accreditation suspension						
Effect through						
- Certifier choices and prices	-0.92	-2.82	1.88	16.42	14.56	95.33
- Certifier choices only	-0.84	-2.82	1.83	45.40	43.58	285.27
Equivalent minimum rigor shift						
Effect through						
- Certifier choices and prices	-0.47	-1.08	0.47	26.94	25.86	169.27
- Certifier choices only	-0.45	-1.11	0.49	31.09	30.02	196.55
(3) Targeting the three most lenient certifiers						
Accreditation suspension						
Effect through						
- Certifier choices and prices	-1.07	-3.48	2.36	7.96	5.77	37.76
- Certifier choices only	-0.97	-3.48	2.28	42.28	40.11	262.58
Equivalent minimum rigor shift						
Effect through						
- Certifier choices and prices	-0.87	-2.38	1.46	49.14	47.34	309.92
- Certifier choices only	-0.80	-2.44	1.44	75.32	73.51	481.24

Notes: Changes in counterfactual scenarios compared to baseline. Accreditation withdrawal is implemented by removing the corresponding certifier from the choiceset. Direct shift of minimum rigor is implemented by shifting the targeted certifiers to the next most lenient certifier's rigor in each market.

Table 1.29: Simulated changes in audit quality and participation

Counterfactual	Avg. minimum rigor across markets (mechanical change)		Participating FMUs (expected change)	
	in numbers	in percent	in numbers	in percent
(1) Targeting the most lenient certifier				
Accreditation suspension				
Effect through				
- Certifier choices and prices	0.04	9.73	-57.16	-5.92
- Certifier choices only	0.04	9.73	-53.47	-5.54
Equivalent minimum rigor shift				
Effect through				
- Certifier choices and prices	0.04	9.73	-11.13	-1.15
- Certifier choices only	0.04	9.73	-11.24	-1.16
(2) Targeting the two most lenient certifiers				
Accreditation suspension				
Effect through				
- Certifier choices and prices	0.08	20.51	-124.12	-12.85
- Certifier choices only	0.08	20.51	-109.26	-11.31
Equivalent minimum rigor shift				
Effect through				
- Certifier choices and prices	0.08	20.51	-57.21	-5.92
- Certifier choices only	0.08	20.51	-54.60	-5.65
(3) Targeting the three most lenient certifiers				
Accreditation suspension				
Effect through				
- Certifier choices and prices	0.11	29.25	-145.86	-15.10
- Certifier choices only	0.11	29.25	-127.62	-13.22
Equivalent minimum rigor shift				
Effect through				
- Certifier choices and prices	0.11	29.25	-103.32	-10.70
- Certifier choices only	0.11	29.25	-92.67	-9.60

Notes: Changes in counterfactual scenarios compared to baseline. Accreditation withdrawal is implemented by removing the corresponding certifier from the choiceset. Direct shift of minimum rigor is implemented by shifting the targeted certifiers to the next most lenient certifier's rigor in each market.

Bibliography

Alecci, S. (2023). Environmental auditors approve green labels for products linked to deforestation and authoritarian regimes. https://www.icij.org/investigations/deforestation-inc/auditors-green-labels-sustainability-environmental-harm/. Last visited 2023-10-03.

Alix-Garcia, J. and Wolff, H. (2014). Payment for Ecosystem Services from Forests. *Annual Review of Resource Economics*, 6:361–380.

Andrews, I. (2018). Weak Instruments and What to Do about Them. https://www.nber.org/lecture/2018-methods-lecture-isaiah-andrews-weak-instruments-and-what-do-about-them. Last visited 2023-08-02.

Araujo, M., Kant, S., and Couto, L. (2009). Why Brazilian companies are certifying their forests? *Forest Policy and Economics*, 11(8):579–585.

ASI (2019). Procedure on Surveillance & Sampling - ASI-PRO-20-105-Surveillance & Sampling-V6.4. https://asi-login.my.salesforce.com/sfc/p/#A0000000aGza/a/1H000000HqeR/i9Ry6Lmd8ybwkdMnw1BNjQU0TbSKoq2tObMHfOHrwgA. Last visited 2022-04-09.

ASI (2020). Find a CAB. https://www.asi-assurance.org/s/find-a-cab. Last visited 2021-09-23.

ASI (2021). Procedure for Witness and Compliance Assessments - ASI-PRO-20-111-Witness & Compliance Assessments-V2.3. https://asi-login.my.salesforce.com/sfc/p/#A0000000aGza/a/120000000GKD/fJyFDEdbdnpj_qHotX7mJRDcjYsq1cbW2jXHgVGmqPE. Last visited 2022-09-04.

ASI (2022). Assurance Services International. https://www.asi-assurance.org. Last visited 2022-08-02.

Assunção, J., McMillan, R., Murphy, J., and Souza-Rodrigues, E. (2023). Optimal environmental targeting in the Amazon rainforest. *The Review of Economic Studies*, 90(4):1608–1641.

Auriol, E. and Schilizzi, S. G. (2015). Quality signaling through certification in developing countries. *Journal of Development Economics*, 116:105–121.

Balboni, C., Burgess, R., and Olken, B. A. (2021). The Origins and Control of Forest Fires in the Tropics. Working Paper.

Bar, T. and Zheng, Y. (2019). Choosing certifiers: Evidence from the British Retail Consortium food safety standard. *American Journal of Agricultural Economics*, 101(1):74–88.

Becker, B. and Milbourn, T. (2011). How did increased competition affect credit ratings? *Journal of Financial Economics*, 101(3):493–514.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 63(4):841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262.

Bizzotto, J. and Harstad, B. (2020). The Choice of Certifier in Endogenous Markets. *SSRN Electronic Journal*.

Blackman, A., Goff, L., and Rivera Planter, M. (2018). Does eco-certification stem tropical deforestation? Forest Stewardship Council certification in Mexico. *Journal of Environmental Economics and Management*, 89:306–333.

Blackman, A., Raimondi, A., and Cubbage, F. (2017). Does forest certification in developing countries have environmental benefits? Insights from Mexican corrective action requests. *International Forestry Review*, 19(3):247–264.

Blackman, A. and Rivera, J. (2011). Producer-Level Benefits of Sustainability Certification. *Conservation Biology*, 25(6):1176–1185.

Bolton, P., Freixas, X., and Shapiro, J. (2012). The credit ratings game. *The Journal of Finance*, 67(1):85–111.

Briesch, R. A., Dillon, W. R., and Blattberg, R. C. (2008). Treating Zero Brand Sales

Observations in Choice Model Estimation: Consequences and Potential Remedies. *Journal of Marketing Research (JMR)*, 45(5):618–632.

Burgess, R., Hansen, M., Olken, B. A., and Potapov, P. (2012). The political economy of deforestation in the tropics. *The Quarterly Journal of Economics*, 127(4):1707–1754.

Cai, Z. and Aguilar, F. X. (2013). Meta-analysis of consumer's willingness-to-pay premiums for certified wood products. *Journal of Forest Economics*, 19(1):15–31.

Cameron, A. C. and Windmeijer, F. (1996). R-Squared Measures for Count Data Regression Models with Applications to Health-Care Utilization. *Journal of Business & Economic Statistics*, 14(2):209–220.

Chen, J., Innes, J., and Tikina, A. (2010). Private cost-benefits of voluntary forest product certification. *International Forestry Review*, 12(1):1–12.

Chu, J., Fisman, R., Tan, S., and Wang, Y. (2021). Hometown Ties and the Quality of Government Monitoring: Evidence from Rotation of Chinese Auditors. *American Economic Journal: Applied Economics*, 13(3):176–201.

Clark, M. and Kozar, J. (2011). Comparing Sustainable Forest Management Certifications Standards: A Meta-analysis. *Ecology and Society*, 16(1).

Conniff, R. (2018). Greenwashed Timber: How Sustainable Forest Certification Has Failed. https://e360.yale.edu/features/greenwashed-timber-how-sustainable-forest-certification-has-failed. Last visited 2023-07-03.

Cubbage, F., Moore, S., Henderson, T., and Araujo, M. M. F. C. (2009). Costs and benefits of forest certification in the Americas. In *Natural Resources: Management, Economic Development and Protection*, pages 155–183. Nova Publishers.

Darby, M. R. and Karni, E. (1973). Free competition and the optimal amount of fraud. *The Journal of Law and Economics*, 16(1):67–88.

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.

D'Haultfoeuille, X., Durrmeyer, I., and Février, P. (2019). Automobile prices in market equilibrium with unobserved price discrimination. *The Review of Economic Studies*, 86(5):1973–1998.

Dragusanu, R., Giovannucci, D., and Nunn, N. (2014). The Economics of Fair Trade. *Journal of Economic Perspectives*, 28(3):217–236.

Dranove, D. and Jin, G. Z. (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature*, 48(4):935–63.

Duflo, E., Greenstone, M., Pande, R., and Ryan, N. (2013a). Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India. *The Quarterly Journal of Economics*, 128(4):1499–1545.

Duflo, E., Greenstone, M., Pande, R., and Ryan, N. (2013b). What does reputation buy? Differentiation in a market for third-party auditors. *American Economic Review: Papers and Proceedings*, 103(3):314–19.

Duflo, E., Greenstone, M., Pande, R., and Ryan, N. (2018). The Value of Regulatory Discretion: Estimates From Environmental Inspections in India. *Econometrica*, 86(6):2123–2160.

Earthsight (2020). FLATPÅCKED FÖRESTS. https://www.earthsight.org.uk/flatpackedforests-en#group-Summary-rayWKL9K1I. Last visited 2022-01-31.

Earthsight (2021). Open letter: FSC is no longer fit for purpose and must urgently reform. https://www.earthsight.org.uk/news/blog-open-letter-fsc-no-longer-fit-for-purpose-and-must-urgently-reform. Last visited 2022-07-18.

Egger, P., Larch, M., Staub, K. E., and Winkelmann, R. (2011). The Trade Effects of Endogenous Preferential Trade Agreements. *American Economic Journal: Economic Policy*, 3:113–143.

European Commission (2018). EU Law on Organic Production: An Overview. https://ec.europa.eu/agriculture/organic/eu-policy/eu-legislation/brief-overview_en. Last visited 2021-09-23.

European Union (2023). REGULATION (EU) 2023/1115 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 31 May 2023 on the making available on the Union market and the export from the Union of certain commodities and products associated with deforestation and forest degradation and repealing Regulation (EU) No 995/2010. *Official Journal of the European Union*, L 150:206–247.

FAO (2018). State of the world's forests 2018 - Forests pathways to sustainable development. Technical report, FAO, Rome.

FAO (2020). Global Forest Resources Assessment 2020: Main report. Technical report, FAO, Rome.

FAO (2023). Forestry Production and Trade. https://www.fao.org/faostat/en/#data/FO. Last visited 2023-04-15.

Federal Ministry for Economic Cooperation and Development (2023). Green Button | Grüner Knopf. https://www.gruener-knopf.de/en. Last visited 2023-08-12.

FRED (2020). Producer Price Index by Commodity: Pulp, Paper, and Allied Products: Wood Pulp. https://fred.stlouisfed.org/series/WPU0911. Last visited 2021-09-23.

Frey, G., Cubbage, F., Ha, T., Davis, R., Carle, J., Thon, V., and Dzung, N. (2018). Financial analysis and comparison of smallholder forest and state forest enterprise plantations in Central Vietnam. *International Forestry Review*, 20(2):181–198.

FSC (2004). SLIMF Eligibility Criteria (FSC-STD-01-003 (Version 1-0) EN). https://fsc.org/sites/default/files/2021-07/FSC-STD-01-003%20SLIMF%20Eligibility%20Criteria.pdf. Last visited 2021-09-23.

FSC (2009a). Forest management evaluations addendum – Forest certification public summary reports (FSC-STD-20-007b (V1-0) EN). https://connect.fsc.org/document-centre/documents/resource/253. Last visited 2023-10-03.

FSC (2009b). Forest management evaluations (FSC-STD-20-007 (V3-0) EN). https://connect.fsc.org/document-center/documents/0bbbebcf-81af-4e89-9c88-3d6e4ed6581d. Last visited 2021-10-13.

FSC (2010). Transfer of FSC Certificates and License Agreements (FSC-PRO-20-003 (V1-0) EN). https://connect.fsc.org/document-centre/documents/resource/236. Last visited 2023-05-17.

FSC (2015). FSC principles and criteria for Forest Stewardship (FSC-STD-01-001 V5-2 EN).

https://connect.fsc.org/current-processes/fsc-std-01-001-v5-2-fsc-principles-and-criteria-pc-forest-stewardship. Last visited 2021-10-15.

FSC (2015). General requirements for FSC accredited certification bodies (FSC-STD-20-001 V4-0 EN). https://connect.fsc.org/document-centre/documents/resource/280. Last visited 2021-10-13.

FSC (2016a). Annual Administration Fee (AAF) (FSC-POL-20-005 V2-5 EN). https://fsc.org/en/document-centre. Last visited 2020-04-26.

FSC (2016b). Cameroon Recognizes FSC Forest Management Certification as FLEGT-Compliant: A Welcome Precedent. https://fsc.org/sites/default/files/2019-04/Cameroon%20first%20country%20to%20recognise%20FSC%20FLEGT%20compliant%20%281%29.pdf.

FSC (2017). FSC Glossary of Terms (FSC-STD-01-002). https://connect.fsc.org/document-centre/documents/resource/207. Last visited 2022-03-22.

FSC (2018a). FSC Global Market Survey Report 2018. https://fsc.org/en/media/5314. Last visited 2021-09-23.

FSC (2018b). The Share of Sustainable Wood: Data on FSC's Presence in Global Wood Production. https://web.archive.org/web/20220224113603/https://fsc.org/en/newsfeed/the-share-of-sustainable-wood-data-on-fscs-presence-in-global-wood-production. Last visited 2022-02-24.

FSC (2019). FSC Evaluates the Credibility and Integrity of its Assurance System. https://fsc.org/en/newsfeed/fsc-evaluates-the-credibility-and-integrity-of-its-assurance-system. Last visited 2021-09-23.

FSC (2020). Public Certificate Search. https://info.fsc.org/certificate.php#result. Last visited 2021-09-23.

FSC (2021a). FSC statement on Earthsight report 2021. https://fsc.org/en/newscentre/integrity-and-disputes/fsc-statement-on-earthsight-report-2021. Last visited 2023-08-11.

FSC (2021b). Interpretations of the normative framework - Forest Management. https://fsc.org/en/document-centre. Last visited 2021-05-31.

FSC (2021c). Interpretations of the normative framework - General. https://fsc.org/en/document-centre. Last visited 2022-07-26.

FSC (2023a). Document Centre. https://fsc.org/en/document-centre. Last visited 2022-07-26.

FSC (2023b). Public Certificate Search. https://info.fsc.org/certificate.php#result. Last visited 2023-06-20.

FSC and IPSOS (2023). FSC Consumer Recognition Study - Global Report June 2023. https://fsc.org/sites/default/files/2023-06/FSC%20Consumer%20Recognition%20Study%20-%20Global%20Report%20June%202023.pdf. Last visited 2023-09-27.

FSC Watch (2020). FSC Watch. https://fsc-watch.com/. Last visited 2020-05-31.

Galati, A., Gianguzzi, G., Tinervia, S., Crescimanno, M., and Veca, D. S. L. M. (2017). Motivations, adoption and impact of voluntary environmental certification in the Italian Forest based industry: The case of the FSC standard. *Forest Policy and Economics*, 83:169–176.

Gandhi, A. and Houde, J.-F. (2019). Measuring substitution patterns in differentiated products industries. NBER Working Paper 26375, National Bureau of Economic Research.

Gavrilut, I., Halalisan, A.-F., Giurca, A., and Sotirov, M. (2015). The Interaction between FSC Certification and the Implementation of the EU Timber Regulation in Romania. *Forests*, 7(12):3.

Goodman, R. C., Harman Aramburu, M., Gopalakrishna, T., Putz, F. E., Gutiérrez, N., Mena Alvarez, J. L., Aguilar-Amuchastegui, N., and Ellis, P. W. (2019). Carbon emissions and potential emissions reductions from low-intensity selective logging in southwestern Amazonia. *Forest Ecology and Management*, 439:18–27.

Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Applications to Poisson Models. *Econometrica*, 52(3):701.

Greenpeace International (2021). Destruction: Certified. Technical report. Last visited 2021-11-18.

Gutierrez Garzon, A. R., Bettinger, P., Siry, J., Abrams, J., Cieszewski, C., Boston, K., Mei, B., Zengin, H., and Yeşil, A. (2020). A Comparative Analysis of Five Forest Certification Programs. *Forests*, 11(8):863.

Hansen, B. (2021). Econometrics. https://www.ssc.wisc.edu/bhansen/econometrics/. Last visited 2021-09-23.

Hastie, T., Friedman, J., and Tibshirani, R. (2001). *The Elements of Statistical Learning*. Springer Series in Statistics. Springer, New York, NY.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica : Journal of the Econometric Society*, 47(1):153–161.

Holopainen, J., Toppinen, A., and Perttula, S. (2015). Impact of European Union Timber Regulation on Forest Certification Strategies in the Finnish Wood Industry Value Chain. *Forests*, 6(12):2879–2896.

Hubbard, T. N. (2002). How do consumers motivate experts? Reputational incentives in an auto repair market. *The Journal of Law and Economics*, 45(2):437–468.

Hui, X., Saeedi, M., Spagnolo, G., and Tadelis, S. (2023). Raising the Bar: Certification Thresholds and Market Outcomes. *American Economic Journal: Microeconomics*, 15(2):599–626.

IMF (2020). World Economic Outlook Database. https://www.imf.org/en/Publications/WEO/weo-database/2020/April. Last visited 2023-06-20.

ISO (2023). ISO 9000 family - Quality management. https://www.iso.org/iso-9001-quality-management.html. Last visited 2023-10-03.

Jaung, W., Putzel, L., Bull, G. Q., Kozak, R., and Elliott, C. (2016). Forest Stewardship Council certification for forest ecosystem services: An analysis of stakeholder adaptability. *Forest Policy and Economics*, 70:91–98.

Jiang, J. X., Stanford, M. H., and Xie, Y. (2012). Does it matter who pays for bond ratings? Historical evidence. *Journal of Financial Economics*, 105(3):607–621.

Kalonga, S. K., Midtgaard, F., and Klanderud, K. (2016). Forest certification as a policy option in conserving biodiversity: An empirical study of forest management in Tanzania. *Forest Ecology and Management*, 361:1–12.

Kitchoukov, E., Stoyanov, N., Simeonova-Zarcin, T., and Stoyanov, T. (2019). Profitability of Forest Certification – Case Study Analysis. *Economic Alternatives*, (4):607–614.

Klein, B. and Leffler, K. B. (1981). The role of market forces in assuring contractual performance. *Journal of Political Economy*, 89(4):615–641.

Kollert, W. and Lagan, P. (2007). Do certified tropical logs fetch a market premium? *Forest Policy and Economics*, 9(7):862–868.

Landsberger, H. A. (1958). *Hawthorne Revisited: Management and the Worker, Its Critics, and Developments in Human Relations in Industry.* ERIC.

Lee, L.-F. (1983). Generalized econometric models with selectivity. *Econometrica : Journal of the Econometric Society*, 51(2):507–512.

Ludwig, G., Tronicke, C., Köck, W., and Gawel, E. (2014). Rechtsrahmen der Bioökonomie in Mitteldeutschland - Bestandsaufnahme und Bewertung. UFZ Discussion Papers 22, Leipzig: Helmholtz-Zentrum für Umweltforschung - UFZ.

Mathis, J., McAndrews, J., and Rochet, J.-C. (2009). Rating the raters: Are reputation concerns powerful enough to discipline rating agencies? *Journal of Monetary Economics*, 56(5):657–674.

Mayer, T. and Zignago, S. (2011). Notes on CEPII's distances measures: The GeoDist database. CEPII Working Paper 2011-25.

McFadden, D. (1977). Modeling the choice of residential location. Cowles Foundation Discussion Papers 710.

Miteva, D. A., Loucks, C. J., and Pattanayak, S. K. (2015). Social and Environmental Impacts of Forest Management Certification in Indonesia. *PLoS ONE*, 10(7):e0129675.

NEPCon (2019). FSC TM Forest Management certification - Service information. www.nepcon.org. Last visited 2019-12-02.

Nordhaus, W. D. (2017). Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, 114(7):1518–1523.

Overdevest, C. and Rickenbach, M. G. (2006). Forest certification and institutional governance:

An empirical study of forest stewardship council certificate holders in the United States. *Forest Policy and Economics*, 9(1):93–102.

Owari, T. and Sawanobori, Y. (2007). Analysis of the certified forest products market in Japan. *Holz als Roh- und Werkstoff*, 65(2):113–120.

Paluš, H., Krahulcová, M., and Parobek, J. (2021). Assessment of Forest Certification as a Tool to Support Forest Ecosystem Services. *Forests*, 12(3):300.

Panlasigui, S., Rico-Straffon, J., Pfaff, A., Swenson, J., and Loucks, C. (2018). Impacts of certification, uncertified concessions, and protected areas on forest loss in Cameroon, 2000 to 2013. *Biological Conservation*, 227:160–166.

Preston, A. E. (1989). The Nonprofit Worker in a For-Profit World. *Journal of Labor Economics*, 7(4):438–463.

Rafael, G. C., Fonseca, A., and Jacovine, L. A. G. (2018). Non-conformities to the Forest Stewardship Council (FSC) standards: Empirical evidence and implications for policy-making in Brazil. *Forest Policy and Economics*, 88:59–69.

Rico-Straffon, J., Wang, Z., Panlasigui, S., Loucks, C. J., Swenson, J., and Pfaff, A. (2023). Forest concessions and eco-certifications in the Peruvian Amazon: Deforestation impacts of logging rights and logging restrictions. *Journal of Environmental Economics and Management*, 118:102780.

Santos Silva, J. M. C. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.

Shim, E. and Sudit, E. F. (1995). How manufacturers price products. *Strategic Finance*, 76(8).

Simonet, G., Subervie, J., Ezzine-de Blas, D., Cromberg, M., and Duchelle, A. E. (2019). Effectiveness of a REDD+ Project in Reducing Deforestation in the Brazilian Amazon. *American Journal of Agricultural Economics*, 101(1):211–229.

Sims, K. R. and Alix-Garcia, J. M. (2017). Parks versus PES: Evaluating direct and incentive-based land conservation in Mexico. *Journal of Environmental Economics and Management*, 86:8–28.

Souza-Rodrigues, E. (2019). Deforestation in the Amazon: A Unified Framework for Estimation and Policy Analysis. *The Review of Economic Studies*, 86(6):2713–2744.

Stahl, K. and Strausz, R. (2017). Certification and Market Transparency. *Review of Economic Studies*, 84(4):1842–1868. https://search-ebscohost-com-s.biblio-dist.ut-capitole.fr/login.aspx?direct=true&db=eoh&AN=1686895&lang=fr&site=ehost-live.

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear IV regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, pages 80–108. eds Donald W.K. Andrews and James H. Stock, Cambridge University Press.

Subervie, J. and Vagneron, I. (2013). A Drop of Water in the Indian Ocean? The Impact of GlobalGap Certification on Lychee Farmers in Madagascar. *World Development*, 50:57–73.

Sugiura, K. and Oki, Y. (2018). Reasons for Choosing Forest Stewardship Council (FSC) and Sustainable Green Ecosystem Council (SGEC) Schemes and the Effects of Certification Acquisition by Forestry Enterprises in Japan. *Forests*, 9(4):173.

Terza, J. V. (1998). Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects. *Journal of Econometrics*, 84(1):129–154.

Train, K. E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Transparency International (2021). Corruption Perception Index. https://images.transparencycdn.org/images/CPI_FULL_DATA_2021-01-27-162209.zip. Last visited 2021-09-23.

Tritsch, I., Le Velly, G., Mertens, B., Meyfroidt, P., Sannier, C., Makak, J.-S., and Houngbedji, K. (2020). Do forest-management plans and FSC certification help avoid deforestation in the Congo Basin? *Ecological Economics*, 175(106660).

UNECE and FAO (2023). Prices. https://unece.org/forests/prices. Last visited 2023-08-13.

UNECE/FAO (2019). Forest Products Annual Market Review, 2018-2019. Technical report, United Nations Economic Commission for Europe (UNECE) and the Food and Agriculture Organization of the United Nations (FAO).

United Nations (2021). Global SDG Indicators Database. https://unstats.un.org/sdgs/indicators/database/. Last visited 2021-09-23.

United Nations (2023). UNSD — Methodology. https://unstats.un.org/unsd/methodology/m49/. Last visited 2023-10-03.

Villalobos, L., Coria, J., and Nordén, A. (2018). Has forest certification reduced forest degradation in Sweden? *Land Economics*, 94(2):220–238.

Visseren-Hamakers, I. J. and Pattberg, P. (2013). We Can't See the Forest for the TreesThe Environmental Impact of Global Forest Certification Is Unknown. *GAIA-Ecological Perspectives for Science and Society*, 22(1):25–28.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, Mass, 2nd ed edition.

World Bank (2020). DEC alternative conversion factor (LCU per US$). https://data.worldbank.org/indicator/PA.NUS.ATLS. Last visited 2021-09-23.

World Bank (2021). World Bank Open Data. https://data.worldbank.org/. Last visited 2021-09-01.

Zeileis, A., Fox, J., and Kleber, A. (2023). Ivreg: Instrumental-Variables Regression by '2SLS', '2SM', or '2SMM', with Diagnostics. https://zeileis.github.io/ivreg/. Last visited 2023-09-26.

Chapter 2

What are the Limits to Private Certification? Evidence from an Attempt to Protect Intact Forests

Kenneth Houngbedji[1]

Johanna Isman [2]

Maria Plakhtieva[3]

Liam Wren-Lewis[4,5]

[1]Laboratoire d'Economie de Dauphine / IRD
[2]Toulouse School of Economics
[3]University of Manchester
[4]Paris School of Economics, INRAE
[5]This project has benefited from support by EUR grant ANR-17-EURE-0001 and CEPREMAP.

2.1 Introduction

Voluntary certification schemes, where companies voluntarily submit themselves to a set of rules monitored by a third party, can play an important role in the regulation of company behaviour (Dragusanu et al., 2014). This is especially true when firms operate in countries with weak state capacity or when the benefits of compliance accrue internationally. Such certification schemes are becoming increasingly important as globalisation outpaces governmental regulation, and economists are beginning to gather evidence on their effectiveness (e.g. Boudreau (2021); Dragusanu et al. (2022)). However, it is unclear to what extent such schemes can substitute for government regulation and whether they can incentivise firms to undertake actions that may substantially restrict their operations.

This paper studies the effectiveness of a flagship private certification scheme in a context where there is serious concern that it might have reached its limits. We look at a recent modification of the rules within the Forest Stewardship Council (FSC). The literature highlights FSC as one of the world's most ambitious and successful transnational non-governmental regulatory schemes (Gutierrez Garzon et al., 2020; Moog et al., 2015). NGOs and companies jointly founded FSC in 1993 after over a decade of failed interstate efforts to improve global forestry management. In 2018, it covered approximately one-tenth of the global wood production (FAO, 2018). In 2014, under pressure from international NGOs including Greenpeace, FSC members passed Motion 65, which required the FSC to strictly regulate the activities concessionaires could undertake in Intact Forest Landscapes (IFLs). IFLs are areas of high conservation value without remotely detectable signs of human intervention.

Those in favour of the motion argued that it would be effective in protecting remaining natural forests. Others, however, argued that the new IFL policy could do more harm than good since it would make it extremely difficult for concessionaires to operate in IFLs, thus reducing FSC-certification rates (Karsenty, 2019). Kleinschroth et al. (2019), for instance, points out that "it has become apparent that the IFL definition collides with forestry practices in several regions [...] The activities of logging companies are thus not compatible with, or are otherwise severely constrained by, the IFL status if such companies are certified under FSC standards". Rotherham (2016) argued that the IFL policy already had led to reduced certification. If that was true, the overall impact of the policy change on global forests could be harmful.

We aim to answer two specific questions. First, did the policy change lead concessions with

IFLs to certify less? Second, did the policy change reduce IFL degradation, in particular through reduced tree-cover loss within certified IFLs? We employ difference-in-differences (DiD) analysis at the concession level to answer the questions.

In the future, we also aim to quantify the overall effect of the policy on IFL, accounting jointly for effects in IFLs that remained certified and those that changed their certification status. But studying the two outlined questions separately is crucial to analyse the potential trade-off between participation and stringency in designing voluntary certification schemes. Understanding the selection into or out of certification is necessary to understand any effect (or lack of) on forest-level outcomes.

This paper focuses on forestry *concessions*. Concessions are publicly owned forests for which management is leased to a typically private company. Concessions account for the vast majority of the FSC-certified forest area and allow the construction of a control group since, in the countries studied, data on concessions is publicly provided independently of the certification status.

To answer the first question, we compare certification choices amongst those concessions without IFLs to similar concessions that contain a large area of IFL. We do not require that certification rates of the two types of concessions were the same prior to the policy change but that the trends would have remained parallel without the policy change. Such an assumption is helped by the fact that the precise definition of an IFL is somewhat arbitrary (Haurez et al., 2017; Lee, 2009; Venier et al., 2018), and concessionaires were thus likely ignorant of the share of IFL in their concession prior to the policy change. Regarding outcomes, we look at the decision over whether to certify or remain certified with the FSC.

To answer the second question, we compare tree-cover loss and overall degradation in IFLs that are in FSC-concessions with non-FSC-certified concessions. Since we want to abstract from selection effects in this part of the analysis, we compare concessions that were certified before 2014 to those that were not. Such a comparison allows to establish a lower bound of a potential reduction in IFL degradation in certified concessions. The critical assumption here is that any change in relative deforestation rates in the years after FSC announced the policy change can be attributed to the policy. We investigate this assumption by comparing concessions certified without IFL in 2013 to uncertified concessions (irrespective of the area of IFL they had in 2013). We consider both the change in IFL extent over time (Potapov et al., 2017) and tree cover loss excluding loss due to fire (Hansen et al., 2013; Tyukavina et al., 2022). Since the data on tree cover loss contain periods before and after the FSC policy change, we can undertake an event-study analysis to look for changes around the policy change.

Geographically, we focus on countries with a substantial amount of IFLs, namely Canada, Russia, the countries of the Congo Basin and, to some extent, Brazil. We match spatial data for the concessions in these countries to certificate data.

We do not find entirely conclusive evidence as to whether the likelihood of terminating the certificate of areas with IFL increased with respect to certified areas without IFL. We do find suggestive evidence for that in Russia. Additional checks are needed to conclude whether trends in certification in concessions without IFL in Russia were comparable to those of concessions with IFL before the policy change. Looking at audit reports, we find that terminations of certificates in a few cases follow the reporting of violations of FSC's new IFL policy.

We find suggestive evidence that the policy led to a reduction in IFL degradation rates in certified areas. For a joint analysis of anticipation and implementation effects, the 95% confidence interval suggests that the policy reduced the annual share of degraded IFL by at least 0.7 percentage points, on average. That is a substantial reduction, compared to the baseline degradation rate of 3 percent per year. Looking at the periods in which the policy was announced and anticipated versus the period of its implementation separately, we find significant effects in both periods, but most of the effect already appearing in the anticipation period. In the paper, we argue that the estimated effects are most likely a lower bound of the actual effects within certified areas. However, the mechanism remains unclear, as we do not observe a reduction in tree cover loss in parallel, rather the contrary.

The paper is structured as follows. Section 2.2 describes the context, and section 2.3 the data. In section 2.4, we define treatment and control groups and describe the development of relevant variables in these groups over time. Section 2.5 describes the empirical strategy in more detail. In section 2.6, we present the results. Section 2.7 provides some robustness checks, and section 2.8 concludes.

2.2 Background

2.2.1 Intact forest landscapes

Intact forest landscapes (IFLs) are at least 500 km^2 large, unfragmented ecosystems within a forest zone without remotely detected signs of human activity (FSC, 2020e; Potapov et al., 2008; Zhuravleva et al., 2013). They are vital for biodiversity and climate change mitigation (Watson et al., 2018). The concept and methodology to identify IFL was only developed in the 2000s, using advancements in satellite imagery.

The global IFL area decreased by 7% from 2000 to 2013, at an increasing rate (Potapov et al., 2017). In 2013-2020, the annual IFL area loss increased by 28% (Greenpeace International et al., 2021). Russia, Brazil and Canada accounted for almost two-thirds of the global IFL area in 2000 and for about 52% of the reduction from 2000-2013. Tropical regions, including Brazil, account for 60% (Potapov et al., 2017). For these reasons and data availability, this study includes Russia, Brazil, Canada, and the countries of the Congo Basin - Cameroon, Gabon, the Republic of Congo and the Central African Republic.

2.2.2 Protection of intact forest landscapes through forest certification

Voluntary certification allows companies to signal to consumers that the wood and non-timber products they sell come from a sustainably managed forest. Such signals are particularly relevant for companies selling to countries with strong consumer interest in sustainability (Ulybina and Fennell, 2013) or for publicly traded firms that want to show compliance with ESG standards.

The membership association Forest Stewardship Council (FSC) has set the most demanding private standard for forest management (Gutierrez Garzon et al., 2020). Among other things, it requires companies to reduce logging to sustainable rates, often lower than those required by governments, use logging methods that cause less forest degradation and protect workers with proper equipment (Cerutti et al., 2011; FSC, 2015). To obtain certification, companies have to apply to an FSC certification body. The certification body audits the company and its forest annually and reports violations of compliance with the FSC standard. The applying company must correct these violations within three months or a year[6] to obtain or maintain

[6]That depends on the severity of the violation.

the certification.

Since 1998, FSC required companies to identify and maintain 'high conservation values' in forests. In 2012, FSC specified categories of high conservation values, including the category of 'landscape-level ecosystems and mosaics' (Kleinschroth et al., 2019). The latter came close to the definition of IFL but did not include them explicitly, and the rules for maintaining these areas were somewhat unclear. In autumn 2014, FSC's General Assembly voted to include IFLs explicitly in the categories of high conservation values and, among other things, to protect the 'core area' of IFLs such that their intactness is ensured (PPECF, 2014). This IFL policy came officially into effect on the 1st of January 2017 and is the only widespread standard that explicitly requires companies to preserve IFLs (FSC, 2018; Greenpeace International, 2021). By default, FSC defined the 'core area' to be protected as an area that covers at least 80% of the IFL remaining on the 1st of January 2017 in each forest management unit.[7] In 2018, FSC introduced some flexibility by allowing 'limited industrial activity' in a 'very limited portion of the core area' if this produced 'clear, substantial, additional, long-term conservation and social benefits' (FSC, 2018).[8] In 2021, national FSC standards allowing further exceptions from the 80% rule came into effect in Cameroon, Republic of Congo, Gabon and Russia (FSC, 2020a,b,c,d). These new standards allow reducing the threshold to 50% or 30% if other additional measures for the conservation of the IFL are taken, such as aiming to make the 30% a legally protected area. One rule that remains and that national standard cannot change is that any IFL shall not be reduced below the 500 km^2 threshold, such that the area does not lose the IFL property (FSC, 2022). The IFL policy includes a range of requirements, all jointly aiming to reduce loss of IFLs (FSC, 2022) - the focus of our analysis. Even if opposition against the 80% default rule might have weakened its implementation, the policy change sent a clear message that FSC aims to become stricter about IFL preservation. That message and the uncertainty regarding the exact criteria likely shaped concessions' certification and IFL preservation decisions.

2.2.3 Forestry concessions

A large share of the forest area, in particular in Canada, the Congo Basin and Russia, is public land rented out as concessions (Tritsch et al., 2020; Ulybina and Fennell, 2013; Williams

[7]The IFL map from 2016 is thus the appropriate baseline map (The IFL Mapping Team, 2021).

[8]The 'very limited portion' shall not be more than 0.5% of the certified forest unit in one year and not more than 5% in total. If an IFL only partially covers a forest unit, this is more than 0.5% and 5% of the IFL core area, respectively.

et al., 2021). In Brazil, the share of public forest land is small (Sist et al., 2021). The tenants of concessions are principally logging companies but can also be companies that produce non-timber products such as maple syrup. While the government regulates forest management in concessions, tenants have some degree of freedom in their management practices. Some tenants may not apply government regulations fully where enforcement is lax, but others apply more demanding environmental standards, for example, to obtain FSC certification. While the countries in our study consider environmental values at least to some extent in forest management regulation, preservation of IFL specifically has not been on the public policy agenda in the years considered (Ptichnikov, 2021; Williams et al., 2021).

2.3 Data

To conduct our analysis, we employ several sets of data, including global maps of IFLs, maps of forest concessions in countries of interest, the database of FSC certification, and remotely sensed measures of forest loss. We merge maps of forest concessions to form a global map of forest management units and intersect them with the data on FSC certification, allowing us to track tenants' certification decisions.

Global IFL maps are available from Potapov et al. (2008) for the years 2000, 2013, 2016, and 2020. Those maps are spatial databases built upon a combination of medium-resolution satellite imagery (most importantly Landsat data), high-resolution imagery from GoogleEarth$^{(TM)}$, and road and settlement data from open access sources. The concept and the mapping method rely on a methodology which a group of renowned researchers (Potapov et al., 2008; Zhuravleva et al., 2013) and non-governmental organizations[9] developed. The researchers designed criteria to locate IFL areas in a globally applicable and easily replicable way: criteria for areas minimally influenced by humans and size criteria compatible with the IFL definition mentioned in Section 2.2.

We obtain maps of the concessions and their tree cover in Brazil, Canada, Russia, and the countries of the Congo Basin from various sources. Canadian data is taken from the Global Forest Watch (GFW) portal.[10]. Available data provides information on 'forest tenures' in Canada, agreements between an industrial forest user and the provincial government to extract

[9] Greenpeace, University of Maryland, World Resource Institute, and Transparent World

[10] Availabe online We have compared the Canadian data from GFW with separate datasets from some provincial government sources (British Columbia, Alberta, Ontario, and others) and concluded that datasets contain the same information.

timber from a defined geographic area. We treat 'forest tenures' as concessions in Canada. For now, we exclude British Columbia from our analysis due to some unresolved issues in the data. The World Wildlife Fund (WWF) Russia provided maps of Russian forests.[11] We obtained quarter-level data covering rental agreements of all forests in Russia, which is based on the governmental data about all active, renewed, and expired rental agreements between the Russian government and private firms. Rental agreements are signed for 10 to 49 consecutive years, and the snapshot of the data we received covers the whole period of interest (2014 - 2020). We define a concession in Russia as a set of quarters managed by the same company in the same geographical region. For the Congo Basin, we used detailed information on logging concessions produced by Tritsch et al. (2020) based on official land-tenure data released by the OFAC and World Resources Institute (WRI) in the "Congo Basin Forest Atlases". The Brazilian concessions data is taken from the Ministry of Agriculture.[12] Concessions were first implemented in Brazil in 2006 (Azevedo-Ramos et al. (2015)) as a tool to control a rampaging timber industry and provide a source of accessible timber for sustainable harvest. There are only 18 concessions in Brazil.

Data on FSC certificates and audits of certified forests in 2014-2020 is provided from Isman (2023), based on the data and audits reports published by FSC (2020f). Data on earlier certificates was extracted from the FSC online database. Certificates were manually matched to the concessions they cover based on geographical and company information in FSC audit reports (FSC, 2020f). One certificate may cover several concessions, but a concession is rarely partially certified or covered by multiple certificates. That is, concessions typically certify fully or not at all. In most of the countries in our study, the boundaries of each concession are clearly defined by administrative decisions and broadly in line with certification decisions. In rare cases where concessions appear only partially certified, we record the concession as certified if more than 50% of the area is certified.

2.4 Descriptives and definition of comparison groups

2.4.1 Certification choices

To study the effect of the IFL policy on certification choices, we compare certification decisions over time in concessions that had at least 30% of the area covered by IFL in 2013-2016 to

[11]Online version of the maps can be found on their website https://hcvf.ru/ru.
[12]Services and Information of Brazil https://www.gov.br/pt-br.

those that had no IFL in 2013-2016. The IFL policy affects the certification choice of all concessions in the treated group throughout 2013-2021, as it increases the cost of obtaining or maintaining certification by requiring concessions to protect a large share of the IFL in their forests to get or remain certified.

Table 2.1: Concessions excluded from the analysis of certification choices

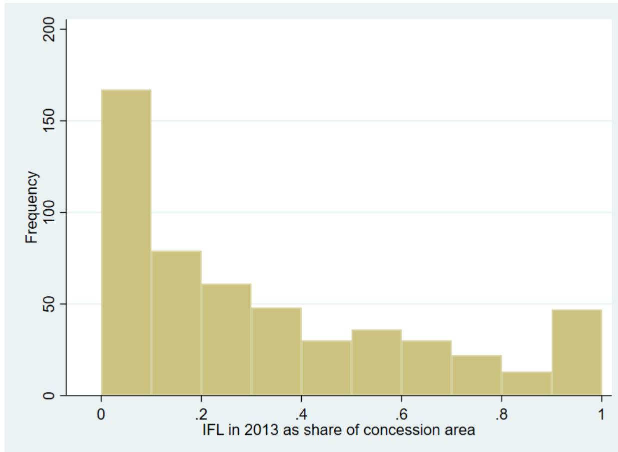
	Full sample		Excl. new concessions		... & excl. those with < 30% IFL	
	(1)	(2)	(3)	(4)	(5)	(6)
	Treated	Control	Treated	Control	Treated	Control
% with rental agreement after 2011	53	47				
% with IFL 2013 nonzero but < 30 % of area			57	0		
% with with IFL 2013 lost until 2016					13	0
Observations	1153	6080	543	3206	236	3206

In the main specification, we focus on concessions rented by the tenant in place in 2016-2020 since at least 2011. As Table 2.1 shows, we thus exclude roughly half of the concessions with IFL and less than half without IFL. 543 treated and 3206 control concessions remain for the analysis.

We restrict the sample that way for two reasons and check for robustness in Section 2.7. First, we want to avoid confounding changes in the differences between certification rates in forests with and without IFL due to changes in the tenant and rental arrangement around the policy change. We often do not know how the forest was used before the rental agreement, if it was even a concession. Moreover, the benefits of certification vary across tenants, as they may be motivated by the branding or upstream buyers of the tenant. Therefore, we exclude concessions rented after 2013, even in the analysis of IFL preservation below. Second, in the analysis of certification choices, we even exclude concessions rented after 2011 since we want to keep the panel balanced while excluding all years before the rental agreement. We want to exclude those years since we lack relevant information for several Russian certificates terminated before 2011, so we cannot match them to concessions. If we included years before 2011 in the analysis, we could count concessions as uncertified, which were certified in that year. Excluding all concessions rented out after 2011 avoids that while preserving a balanced panel.

We further exclude concessions that had less than 30% of their area covered by IFL in 2013 since the burden of complying with the policy is likely weaker for these concessions (Williams et al., 2021). This exclusion removes roughly half of the remaining treated concessions. Figure

Figure 2.1: Histogram of IFL area in 2013 as share of the concession area



2.1 shows that the share of IFL among concessions with higher IFL shares is relatively evenly distributed.

Finally, we exclude the 9% of the remaining concessions that had more than 30% IFL in 2013 but no longer in 2016. Those concessions were treated by the announcement of the IFL policy but not treated when the policy was implemented in 2017.[13] 206 treated and 3206 control concessions remain for the analysis.

Figure 2.2: Certification status by year and IFL

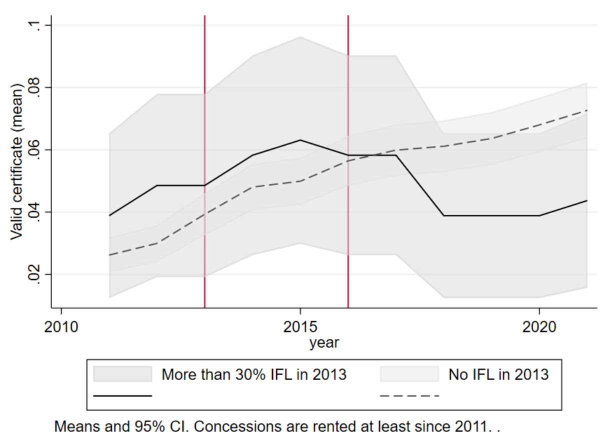
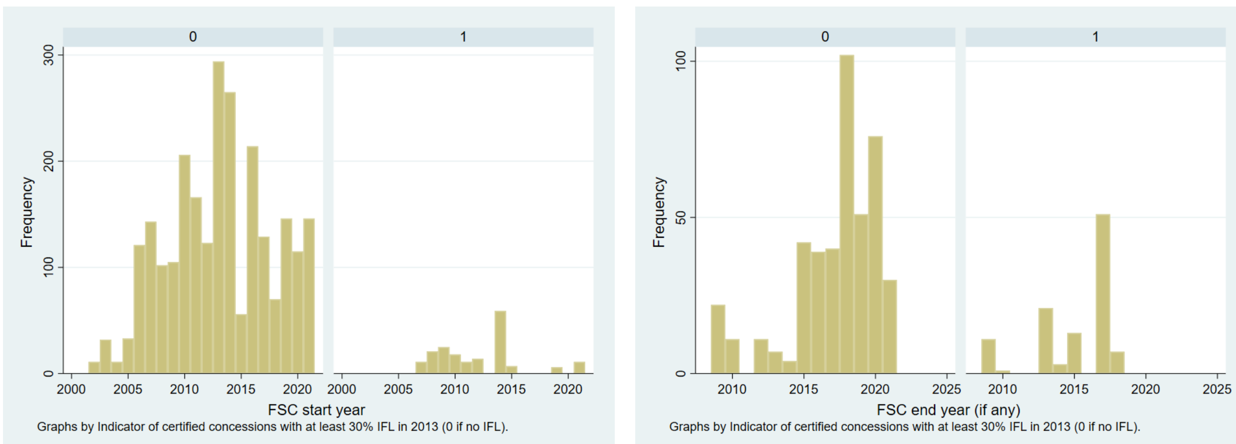


Figure 2.2 compares the certified fraction of concessions with and without IFL in a given year. From 2011 to 2015, certification was more likely among concessions with IFL. The increasing trends are approximately parallel for both groups in that period. After 2015, the certification rates among concessions with IFL decreased, while those among concessions

[13] All those concessions never got certified in 2011-2020.

without IFL continued to grow. Figures 2.3 show the distribution of start and end years of concessions' certification by treatment status. It shows that changes in the certification rates of the treatment and comparison groups are both driven by concessions getting certified and concessions dropping out of certification.

Figure 2.3: Distribution of years of initial certification and dropouts for certified concessions with and without IFL, excluding concessions rented after 2013 or concessions with less than 20% IFL.



(a) Years of initial certifcation

(b) Years of termination (if any)

2.4.2 IFL preservation in certified concessions

To study the effect of the FSC rule change on IFL preservation in certified concessions, we compare the development in IFL and tree cover loss rates in concessions with IFL that were certified to those uncertified in 2013. That is, we base the treatment definition on the certification status before the policy change. We, thereby, avoid confounding the effect of the policy with an effect of selection of certain types of companies into the treatment. It allows to estimate lower bound of the policy's effect on concessions that are certified, as we further outline in the next section.

Table 2.2: Concessions excluded from the analysis of IFL preservation

	Shares in full sample	
	(1)	(2)
	Treated	Control
% with rental agreement after 2013	24	45
Observations	76	1077

Table 2.2 shows the types of concessions excluded from the main analysis. We assess the robustness of our results to that exclusion in Section 2.7. We focus on concessions that were rented before 2013 since we want to avoid confounding changes due to changes in the tenant, as outlined in the previous section. We thus remove 24% of the 76 concessions with IFL that were certified in 2013 and 45% among the 1077 concessions with IFL that were uncertified in 2013. In the following, we present statistics for the remaining 589 concessions, of which 58 were certified in 2013 and thus treated.

Table 2.3: Summary statistics of time-invariant variables

	(1)		(2)	(3)		(4)	(5)		(6)
	Canada - Treated		Treated-Control	Congo Bassin - Treated		Treated-Control	Russia - Treated		Treated-Control
	Mean	SD	Mean diff.	Mean	SD	Mean diff.	Mean	SD	Mean diff.
Area of concession in K ha	1684.1	1610.1	-120.4	298.5	317.1	167.5	416.0	498.4	350.1**
IFL area in 2000 in K ha	302.6	531.6	-94.6	149.2	246.3	80.9	62.2	81.4	41.7*
Tree cover in 2000 in K ha	1042.6	906.1	241.8	279.4	310.2	162.3	224.7	236.5	185.4**
Year of the rental agreement	1998.6	14.8	-1.8	1999.1	6.7	-4.0*	2009.1	1.5	-0.8*
Observations	22		54	15		205	21		387

Notes: Concessions rented at least since 2013 and with more than 0 % IFL in t=0. Treated (control) are those (not) certified in 2013.
Mean differences are treated minus control means. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.3 shows baseline characteristics by world region. In the Congo Basin and Russia, treated concessions tend to consist of larger areas, more IFL and more tree cover in 2000. They also tend to be rented to the current tenant more recently than the control concessions. In Canada, the differences are instead the opposite, but not significant. There is also much variation among the treated concessions. All of that variation motivates including corresponding fixed effects, controls and weighting in the analysis.

Table 2.4: Summary statistics of outcome variables

	(1)		(2)	(3)		(4)	(5)		(6)
	Treated 2013		Mean diff.	Treated 2016		Mean diff.	Treated 2020		Mean diff.
	Mean	SD	2013	Mean	SD	2016	Mean	SD	2020
80% IFL remain from last IFL map in p.p.	34.5	47.9	-31.1***	84.5	36.5	5.0	74.1	44.2	-10.4
80% IFL minus 0.5% area/yr. remain from last IFL map in p.p.	56.9	50.0	-14.6*	91.4	28.3	9.7*	84.5	36.5	-2.4
Preserved share of 2000 IFL in p.p.	62.0	28.7	-17.1***	56.1	28.5	-12.6**	48.9	29.3	-11.9**
Avg. IFL loss rate since last map in p.p.	2.9	2.2	1.3***	2.0	3.8	-1.5*	1.8	2.7	-0.2
Tree cover loss rate in 2000 IFL in p.p.	0.6	1.5	0.3	1.4	4.0	0.5	0.7	1.8	0.3
Non-fire-caused tree cover loss rate in 2000 IFL in p.p.	0.4	0.6	0.2*	1.0	3.3	0.8	0.4	0.7	0.2
Tree cover loss rate in p.p.	0.6	1.4	0.2	1.2	3.1	0.3	0.7	1.4	0.2
Non-fire-caused tree cover loss rate in p.p.	0.3	0.3	0.1***	0.6	0.7	0.2*	0.4	0.5	0.2*
Observations	58		647	58		647	58		647

Notes: Concessions rented at least since 2013 and with more than 0 % IFL in t=0. Treated (control) are those (not) certified in 2013.
Mean differences are treated minus control means. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.4 presents means in outcome variables among the treated and differences in means between treated and control concessions in 2013, 2016 and 2020. The first row shows the fraction of concessions that had preserved 80% of the IFL area from the last IFL map. For the treated in 2020 in column (5), about every fourth of the 58 treated concessions did not comply with the initial 80% rule of the IFL policy.[14] The second row of column (5) implies that nine of those 15 concessions did not even comply with the relaxed 80% rule from 2018. The relaxed rule allowed companies to go beyond the 20% loss of IFL if this does not affect more than 0.5% of the whole forest management unit per year.[15]

The available audit reports of the concessions that did not comply with the relaxed 80% rule in 2017-2020 do not document major violations related to IFL degradation. For one of the concessions that did not comply with the initial strict rule, an IFL-related major violation is documented. The audit reports for that concession are an example of how many FSC audits work and what might impede the implementation of the IFL policy: Findings of clear-cuts in IFL areas were based on the review of documents and reports from NGOs rather than detection during field audits or analysis of satellite imagery.[16] For correcting the violations and maintaining the certificate, updated management plans and procedures were sufficient. Compensation for induced IFL degradation was not required.[17]

The last six rows of Table 2.4 show statistics on IFL and tree cover loss over time. The time series of these measures are presented graphically in Figures 2.4, 2.5 and 2.6, by treatment status.

Figure 2.4 plots the share of IFL that is preserved in 2013, 2016 and 2020 in panel (a) and the average annual loss rate in panel (b). The IFL loss data do not allow for pre-trend analysis, as the first observations of IFL loss are the aggregated losses from 2000-2013, the only observation before the announcement of the IFL policy. Treated concessions have a

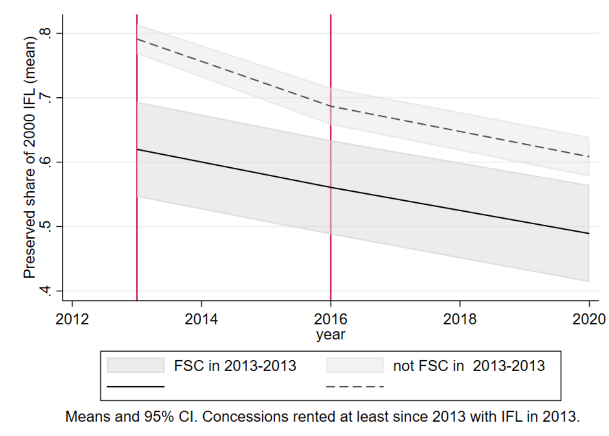
[14]That is an approximation since we do not have a good measure of whether and to what extent concessions are partially certified. That means some of the IFL losses we observe might have happened outside the certified area within a partially certified concession. However, in that case, the certificate holder arguably still violates the certification rules since the FSC's Policy for Association requires certified forest units not to destroy high conservation values even in non-certified areas under the same management. Certifiers will just detect that less often since they do not directly audit uncertified areas (FSC, 2011).

[15]The 85% share of compliant concessions is likely a lower bound since the IFL maps only allow to compute the average loss per year.

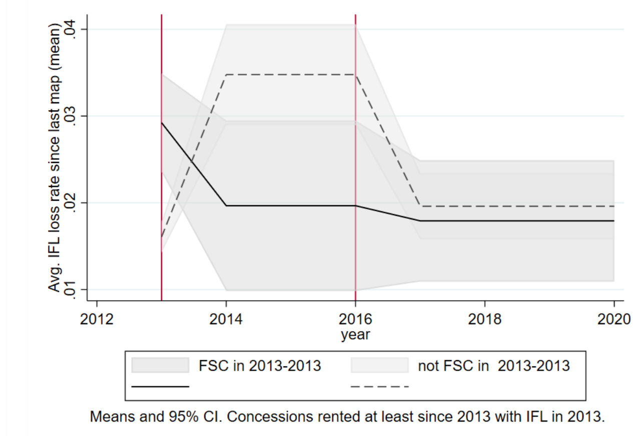
[16]Analysis of other audit reports with IFL related violations (but compliance with the 80% rule) confirm that.

[17]After logging in intact forests had been designated as a "minor" violation in 2017, a new case of clear-cuts in intact forests led the auditor to "upgrade" the violation to a "major". Major violations are costly for companies, as they lead to suspension if not corrected until an extra audit three months and lead to direct suspension if five or more are revealed in one audit (FSC, 2009). Documentation for the extra audit three months after the 2018 audit is missing. The audit of 2019 reveals harvesting in intact forests in other parts of the concession. In the extra audit three months later, this violation is considered as corrected based on updated procedures and related documents that are supposed to prevent such harvesting in the future. See https://search.fsc.org/en/certificate/a0240000005snqnAAA/ for most of those findings. The report from 2018 is no longer online, but was scraped in 2020 for (Isman, 2023).

Figure 2.4: Comparison of means in shares of preserved/lost IFL of concessions certified in 2013 (treated) to those in uncertified concessions (control)



(a) Preserved share of IFL area from 2000



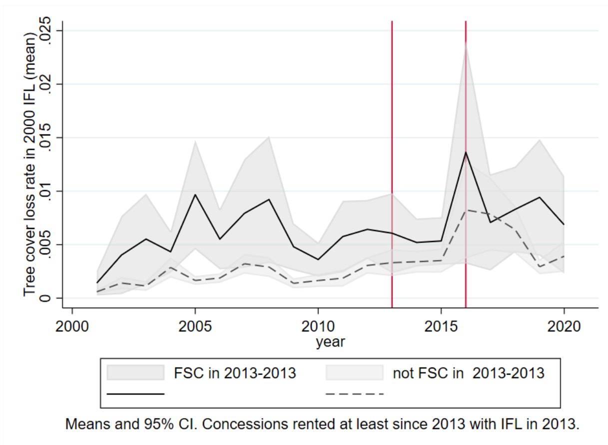
(b) Avg. annual IFL loss rate since last IFL map from IFL area from 2000

significantly lower level of preserved IFL than the control group. On average, about almost 38% of the IFL area in treated concessions were degraded in that pre-policy period, but only 21% in the control group. On average, that suggests annual IFL loss of 2.9 percentage points among the treated and 1.6 percentage points in untreated concessions. In the period of anticipation of the IFL policy, 2013-2016, the rate of IFL degradation slowed down by almost one percentage point in treated concessions, but increased by 1.9 percentage points in the control concessions. Loss rates almost only dropped 0.2 percentage points further after the policy was officially introduced in 2017. In parallel, it decreased to 2 percentage points per year in the control concessions, remaining above the new level among the treated.

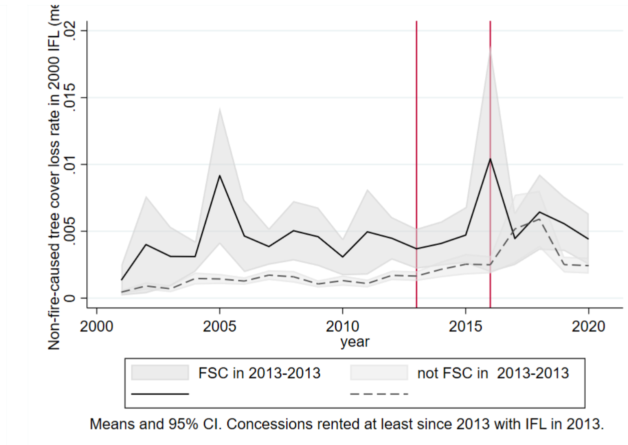
Figure 2.5 presents rates of tree cover loss within areas that were IFL in 2000. Panel (a) includes any tree cover loss in those areas, while panel (b) only shows rates of tree cover loss that are not due to fires. The latter is the more relevant measure for IFL preservation since tree cover loss due to wildfires does not count as IFL degradation and is less problematic from a forest conservation perspective. A caveat is that there is only yearly data on any fires, while wildfires are fires that happen sufficiently far from human infrastructure to be likely caused by humans.

Figure 2.6 plots tree cover loss as shares of the whole concession area. Panel (a) includes any tree cover loss, while panel (b) excludes tree cover loss caused by fires. Looking at the whole area might allows us to assess whether the policy might have caused a substitution of

Figure 2.5: Comparison of means in tree cover loss in areas that were IFL in 2000 in concessions certified in 2013 (treated) to those in uncertified concessions (control).

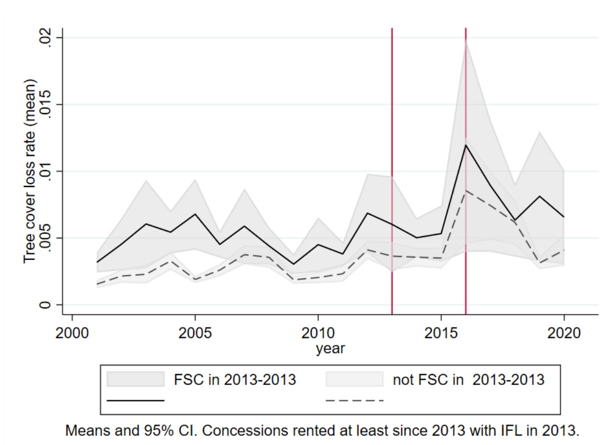


(a) Rate of any tree cover loss in IFL

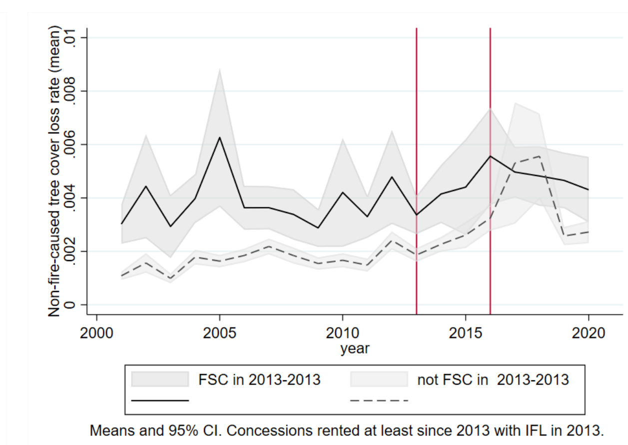


(b) Rate of tree cover loss that is not due to fires in IFL

Figure 2.6: Comparison of means in tree cover loss throught concessions certified in 2013 (treated) to those in uncertified concessions (control).



(a) Rate of any tree cover loss in and outside of IFL



(b) Rate of tree cover loss that is not due to fires in and outside of IFL logging in IFL with logging in non-IFL areas. All the figures of tree cover loss suggest more fluctuation among treated than untreated concessions, but movements in the same direction in most periods before the policy change. There were no extreme changes in the trends after the announcement and implementation of the policy, but there was an increase in tree cover loss in 2016, followed by a drop in certified concessions. In untreated concessions, fires seem to explain the increase in 2016; in treated, they don't not. However, those numbers are pretty noisy. The decrease in tree cover loss in certified concessions following 2016 suggests that reduced logging in IFLs was at least not fully compensated by logging outside of IFLs.

2.5 Empirical strategy

To study the effect of the IFL policy on certification choices and the IFL preservation, we use Difference-in-Differences (DiD) approaches. We compare the differences in the outcome between treated and untreated concessions before and after the treatment. Throughout the analysis, i denotes the concession, c its country and r its cohort, i.e. the year of its rental agreement. t denotes the period relative to the treatment: 0 in the period before the treatment, 1 in the first period of the treatment, and 2 in the second. Period 0 is 2013 since the IFL policy was announced in 2014 such that anticipation may already have affected certification decisions and IFL preservation, given that forest management is typically planned for multiple years. Standard errors are robust to heteroskedasticity and clustered at the company level or, where the company is not observed, at the concession level.

2.5.1 Certification choices

To estimate the effect of the IFL policy on concessions' likelihood to obtain or maintain FSC certification, we consider a dynamic two-way fixed effects regression specification:

$$FSC_{irct} = \alpha_i + \phi_{ct} + \phi_{rt} + \beta_t IFL_i + \epsilon_{irct} \tag{2.1}$$

The unit of the period t is a year. FSC_{irct} is an indicator that is 1 if the concession was certified in period t. IFL_i is an indicator that is 1 if the concession is treated, that is if an IFL covered at least 30% of the concession in 2013-2016. The comparison group are concessions that had no IFL in 2013-2016. α_i is a concession-level fixed effect. ϕ_{ct} and ϕ_{rt} are fixed effects for each country-year and each cohort-year pair. ϵ_{irct} is a mean independent shock. We weigh concessions by their area to account for the differential importance of certification in terms of the area.

Next to the Linear Probability Model in (2.1), we consider a Probit model:

$$Pr(FSC_{irct} = 1|IFL_i, r, c, t, \boldsymbol{x_i}) = \Phi\Big(\alpha_0 IFL_i + \phi_{ct} + \phi_{rt} + \beta_t IFL_i + \boldsymbol{x_i'}\boldsymbol{\beta_4}\Big) \qquad (2.2)$$

In the Probit model, we cannot control for concession fixed effects but only for an indicator of being treated. We, therefore, include the total area and the IFL area in 2000 as controls $\boldsymbol{x_i'}$.

The identifying assumption in both models is that the change in the difference between the relative likelihood of being certified before and after the announcement of the IFL policy can be entirely attributed to the policy, apart from changes at the national level or changes due to the time since the rental agreement, as the fixed effects capture these. That assumption requires that the trends of the two groups within the same country and cohort would have been parallel without the IFL policy. The dynamic TWFE specification allows us to assess whether the groups followed parallel trends before the announcement of the policy, which provides suggestive evidence for or against the parallel trends assumption. Even in the presence of parallel pre-trends, potential threats to the parallel assumption could come from other policies or market developments that change the difference between the net benefits of certification of concessions with and without IFL. We are unaware of any developments other than the IFL policy that would affect those two groups differently. The IFL concept was still relatively new in 2013 and received little attention among policymakers and businesses (Ptichnikov, 2021; Williams et al., 2021).

One concern is a potential spillover of the policy effects on certified concessions without IFL. There are two potential channels for that. First, gross price premia for certification might rise due to the increased stringency of certification. If they do, they should rise proportionally in and outside of IFL, as the FSC label does not distinguish between the two types of concessions. A proportional increase in price premia could improve the trends in certification rates but should change them for both in parallel. The DiD estimate would then underestimate the overall effect on certification rates, as it would consider an increase in certification rates in concessions without IFL as the counterfactual trend rather than an outcome of the policy. However, there is no evidence for an increase in price premia, not even anecdotally. It is unlikely that consumers or investors are sufficiently informed about increases in FSC's requirements to compensate for such changes over a few years.

The second potential channel for spillover effects is more relevant than the first: The IFL policy requires concessions to reduce their activity in IFLs. Compliance with the policy could thus reduce the supply of FSC-certified wood from those concessions. Moreover, reductions in certification rates would reduce the supply as well. That could increase pressure on certified concessions without IFL to log more intensely. However, the descriptives of tree cover loss shown in the last section do not show any indications of that. Certified concessions are typically already under a tight policy of sustainable forest management, limiting the scope for intensified logging, particularly in the short run.

2.5.2 IFL preservation in certified concessions

We assess the effect of the IFL policy on IFL preservation with two measures, denoted by y_{irct}. The first measure of IFL preservation is the lost share of IFL. The second measure of IFL preservation is the share of tree cover in IFL that is lost due to causes other than fire, one of the main drivers of IFL loss.[18] We consider a dynamic two-way fixed effects regression specification for the subset of concessions with IFL in $t = 0$:

$$y_{irct} = \alpha_i + \phi_{ct} + \phi_{rt} + \beta_t FSC_i + \epsilon_{irct} \tag{2.3}$$

where FSC_i is an indicator which is 1 if a concession had a valid certificate in $t = 0$. That is the treated group. The comparison group are concessions without valid certificates in period 0. We include the same fixed effects as in the analysis of certification decisions. We weigh concessions by their IFL area to account for the differential impact IFL loss in each concession in terms of shares has on aggregate IFL loss.

Under a few assumptions discussed below, β_t identifies a lower bound of the effect of the policy among the certified. Note that the treated group in our specification consist of (1) concessions that remained certified throughout 2014-2020, and (2) concessions that left FSC in that time. The control group, i.e. those uncertified in 2013, consist of (3) concessions that remained uncertified and (4) concessions that become certified in 2014-2020. We want to assess the policy's causal effect on group (1), compared to group (3).[19] However, directly

[18] While the main rule of the policy is the relaxed 80% rule discussed earlier, assessing the policy's effect on compliance with that rule is difficult, as it is not clear what an equivalent rule would have been in earlier periods. Moreover, the 80% rule is likely simply an operationalization of the broader aim of preserving substantially more IFL than in earlier periods.

[19] In the future, we also want to quantify the overall effect of the policy. Comparing concessions that were certified and uncertified in 2013 does not directly quantify overall effect, since the concessions in the treated (or control) group are not affected (or unaffected) by the policy in all years. To assess the overall effect of the

comparing group (1) and (3) would confound the causal effect of the policy with a selection effect, as concessions may take the policy change into account when deciding whether they remain certified or uncertified, or not. Comparing the effect on groups (1) and (2) to groups (3) and (4) avoids that. It establishes a lower bound of the causal effect on group (1), under two plausible assumptions: First, we expect that the IFL policy improves IFL preservation in group (1) at least as much as in groups (2) and (4), respectively. Second, we expect that it improves IFL preservation in group (4) as least as much as in group (3), where the effect on group (3) should even be zero.

In addition, identification of a causal effect requires that concessions that were certified or uncertified in 2013 would have followed parallel trends in IFL preservation without the IFL policy. For tree cover loss, the dynamic TWFE specification allows us to assess whether the groups followed parallel trends before the policy announcement.

Threats to the parallel trends assumption could come from other policies or market developments that change the difference between the net benefit of logging or construction of infrastructure in IFL areas between certified and uncertified concessions. As discussed above, we are unaware of other relevant changes related to IFL specifically. Neither have we found indications of other substantial changes in the FSC policy that could explain reduced forest degradation in certified concessions more generally. In Section 2.7, we also look for indications of such general changes in the FSC system by comparing the differences in differences in non-fire-caused tree cover loss in the non-IFL areas of concessions with IFL.

Another potential threat to the identifying assumption could come from general equilibrium effects of the policy. Reduced logging in certified IFL would reduce the supply of wood, which could affect the profitability of logging more in uncertified areas. However, substantial effects in this direction are unlikely since certified wood from IFL corresponds to a tiny share of global wood production. In our data, certified concessions with IFL correspond to only 1% of all concessions.

2.5.2.1 IFL loss

IFL loss is defined as the average annual share of IFL from 2000 that human intervention degraded since the last IFL map. We focus on the average annual share to account for the difference in the number of years between the IFL maps. We consider 2013 as the

policy, we will, thus, need to apply DiD estimators that account for those differential effects across the years.

pretreatment period 0, 2016 as the first treatment period 1, and 2020 as the second treatment period 2, since these are the years with data on IFL loss. We include the three periods in the dynamic TWFE since the anticipation effect from 2013 to 2016 may differ in magnitude from the effect of the policy implementation from 2017-2020. The main caveat is that we cannot test for the pre-trend since no data on IFL loss before 2013 are available.

2.5.2.2 Non-fire-caused tree cover loss in IFL

Non-fire-caused tree cover loss in IFL is the total share of tree cover in IFL measured in 2000 that is lost in year t, minus the share lost due to fires. That is usually one of the main drivers of IFL loss, a critical outcome in itself, and allows for yearly and geographically more granular analysis, including pre-trend analysis. We subtract loss due to fires since this includes much loss due to wildfires that do not count as IFL loss. We define the period since treatment t as the year minus the pretreatment year 2013. We consider the years from 2001 to 2021.

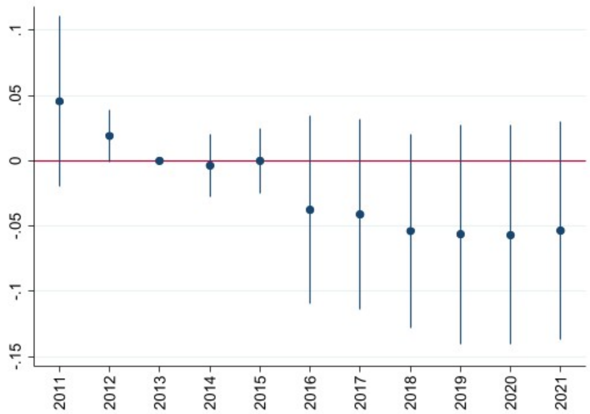
2.6 Results

2.6.1 Certification choices

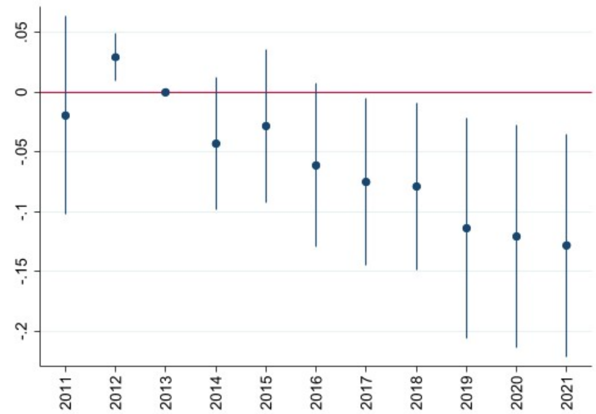
Figure 2.7(a) shows the effect sizes and 95% confidence intervals of the yearly DiD estimators from the estimation of model 2.1. Table 2.6 in the Appendix shows the coefficients compared to the results from the probit model in equation 2.2. The results are similar. There is a slight downward jump in the point estimates one year before the official implementation of the IFL policy. After that, the difference in certification rates remains stable until 2021, according to the point estimates. None of those coefficients is statistically significant, however. Confidence intervals suggest that the IFL policy could be associated with anything from a relative decrease of certification rates among concessions with IFL of almost 13% up to increases of almost 3%. There is also no significant pre-trend, even though the point estimates suggest that a decrease in certification rates among concessions with IFL might already have started before the announcement of the IFL policy, potentially with early discussions about the concept of IFL.

Figure 2.7(b) shows the DiD coefficient sizes from the same regression, but only including Russian concessions. The DiD coefficients suggest a reduction in certification rates of about 3 to 22% by 2021, with 95% confidence. However, we also find a significant difference in the differences in certification rates between 2012 and 2013. That difference hints at a potentially

Figure 2.7: DiD coefficients estimates according to equations 2.1, with 95% CI. Standard errors are clustered at the company level. The observations are weighted by area.



(a) Full sample.



(b) Russian concessions only.

earlier start of calls for stricter IFL requirements than we know. At least, the trend goes in the same direction. Robustness checks in that regard are needed to affirm the suggestive evidence of a negative effect on certification rates in Russia.

In FSC's audit reports, we find two cases in which certificate terminations followed upon reported violations of the IFL policy and certifier requests to correct them.[20] Otherwise, the cost of complying with the policy and anticipation of corresponding violation reports are probably the primary explanations for the reduction in certification rates.

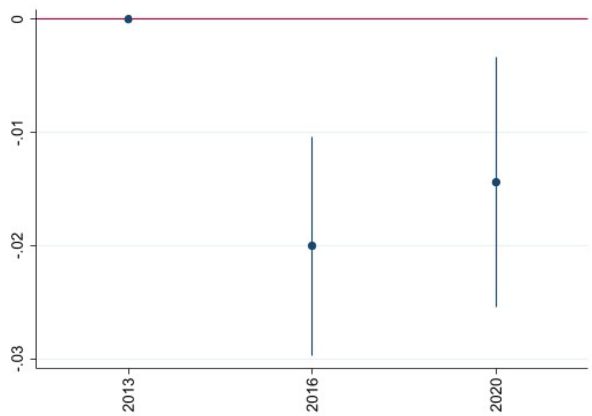
2.6.2 IFL preservation in certified concessions

2.6.2.1 IFL loss

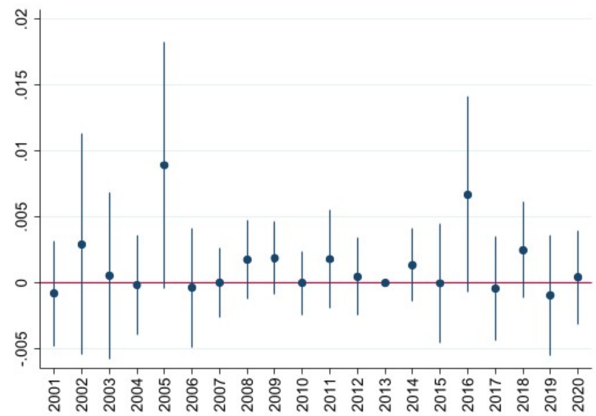
Figure 2.8(a) shows the effect sizes of the DiD estimators from the estimation of the dynamic TWFE model 2.3. Table 2.5 shows the plotted coefficients in column (2), our main specification. Column (1) does not include cohort-year fixed effects. Under the assumption discussed in the last section, the coefficients suggest that the IFL policy caused a significant reduction in the annual IFL loss rate from 2000-2013 to 2014-2016, the years in which the policy was

[20] In the 2018 audit summary reports of certificate FC-FM/COC-643024, the certifier reports that "Cuts are allowed in forests where HCVF2 of the regional level are identified as intact forests of Priangarye". In the 2018 report for certificate FC-FM/COC-643107, the certifier notes that "The certificate holder has allowed cases of fragmentation of intact forest territories in 2017, which doesn't allow to preserve them in accordance with the established regime of strict protection, envisaged by the Agreement between the enterprise and the non-commercial organisation." See the public audit summary reports on https://search.fsc.org/en/ for details.

Figure 2.8: DiD coefficients estimates as lower bound estimates of effects on IFL preservation outcomes, with 95% CI. Standard errors are clustered at the company level. The observations are weighted by IFL area.



(a) DiD in IFL loss rates, from Table 2.5, according to equation 2.3.



(b) DiD in non-fire-caused tree cover loss, from column (2) in Table 2.7, according to equation 2.3.

debated and announced. According to the 95% confidence interval, it is a reduction of at least 1.1 percentage points. We find an additional, albeit smaller reduction after the official policy implementation in 2017-2020. The estimate from column (1) is similar, but smaller and insignificant. Appendix-Figure 2.10 shows the DiD coefficient for a joint analysis of 2013-2020, including cohort-year fixed effects. It suggests an overall reduction in IFL loss rates by at least 0.7 percentage points.

Two factors might push most of the effect into the anticipation period 2013-2016. First, the common practice of long-term planning in forest management could explain a strong anticipation effect. Concessions might have reacted after the decision about the IFL policy by updating their harvesting plans from the year forward. Second, as described earlier, FSC slightly relaxed some of its rules soon after the policy implementation, including the 80% threshold of IFL preservation.

2.6.2.2 Non-fire-caused tree cover loss in IFL

Figure 2.8(b) shows the effect sizes of the yearly DiD estimators from the estimation of dynamic TWFE according to model 2.3, with non-fire-caused tree cover loss as the outcome variable. Table 2.7 in the Appendix shows the plotted DiD coefficients in column (2), while column (1) omits the cohort-year fixed effects. Most coefficients are positive, all are insignificant at a 5% level. Most coefficients are close to zero, except those in 2005 and

Table 2.5: Annual average IFL loss rate since last IFL map

	(1)	(2)
FSC x 2016	-0.0122	-0.0200***
	(-1.77)	(-4.07)
FSC x 2020	-0.0121*	-0.0144*
	(-2.24)	(-2.56)
Constant	0.0192***	0.0200***
	(21.67)	(26.93)
Observations	1908	1908
Year - country FE	Yes	Yes
Concession FE	Yes	Yes
Year - rental agreement FE	No	Yes
Weighted by area	Yes	Yes

Notes: The dependent variable is a the rate of average annual IFL loss since the last IFL map as a share of the IFL in 2000. Standard errors are clustered at the company level (or the concession in the concessions, where the company name is missing). Concessions rented at least since 2013 and with more than 0 % IFL in t=0. Treated (control) are those (not) certified in 2013. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

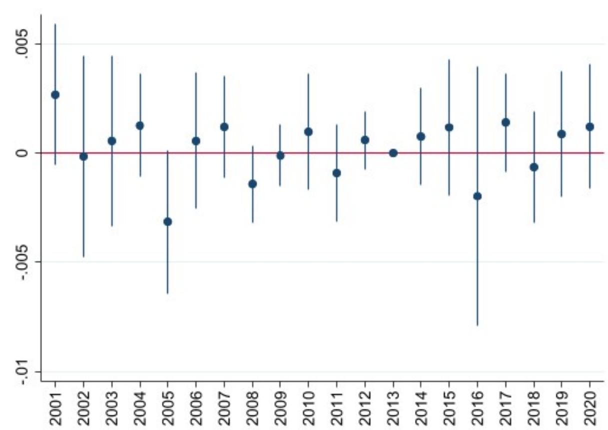
2016. The latter suggests an increase in non-fire-caused tree cover loss rates in certified IFL of a bit more than zero to 1.5 percentage points due to the policy. Yet, recall that those coefficients are only lower bounds of a potential reduction in tree cover loss. The true effect is closer to zero and might even be a reduction, as we would expect. However, the estimates do certainly not allow to conclude that the cause for the decrease in IFL loss is a reduction in non-fire-caused tree cover, even though this is usually the main driver for IFL loss.[21]

2.7 Robustness

For the whole analysis, including concessions with a rental agreement from 2013-2020 does not change the results much, as the coefficient plots 2.11 and 2.12 in the Appendix show. The signs of the coefficients of interest are the same. The only main difference is that the effect of the policy on IFL preservation in 2020 is insignificant in that version.

When including concessions with less than 30% IFL, estimated effects on certification rates are significantly negative in some years for the whole sample. We show the coefficient plot in the Appendix-Figure 2.13. In that case, the effects are not only driven by Russia. The DiD coefficients in a Russian-only two-way fixed regression are no longer significant.

Figure 2.9: DiD coefficients estimates from analysis of non-fire-caused tree cover loss in non-IFL areas according to equation 2.3, with 95% CI.



We check for the possibility that the policy had spillover effects on non-IFL areas in concessions with IFL or that other developments affected the difference between certified and uncertified

[21]Note that concessions in the control group may have been certified in the years before 2013, while treated concessions may not have been certified yet. In the future, we will test how accounting for these differences affects the results.

concessions in general around the timing of the IFL policy. To do so, we rerun the same dynamic TWFE specification as in the analysis of non-fire-caused tree cover loss in IFL, but with the outcome variable being non-fire-caused tree cover loss in non-IFL areas.

Figure 2.9 presents the effect sizes with 95% confidence intervals. The fluctuations in the differences between certified and uncertified concessions are not statistically significant, suggesting parallel trends. It is striking that the point estimates in most years have the inverse sign of those in the analysis of non-fire-caused tree cover loss in IFL. That might be related to the rotation of harvesting practices, but we need to investigate it further. One option to account for such differences is a triple differences-in-differences analysis.

2.8 Conclusion and outlook

This paper studies the attempt of the voluntary certification scheme FSC to protect intact forests. First, we analyse whether this attempt causes forestry concessions to leave the certification scheme altogether. In Russia, we find suggestive evidence that the likelihood of certification of areas affected by FSC's new policy decreased compared to unaffected areas. For other countries, we do not find clear evidence in that regard. Second, we assess the effectiveness of the policy in reducing the degradation of intact forests among concessions that remained certified despite the policy. We find substantial improvements, particularly in anticipation of the actual policy implementation. The channel for such improvements remains unclear. The policy did not reduce tree cover loss in the affected areas.

Additional data and analysis are necessary to assess the robustness of those results and the aggregate effects of the policy. The number of observations used for the above analysis is small. Soon, we will add additional observations from British Columbia and Indonesia. Matching treated and control concessions with similar pre-trends in the outcome variables or triple differences-in-differences with non-intact forest areas in the same concessions could mitigate concerns about potential violations of the parallel trends assumption. Power analysis will be helpful to assess how confidently we can interpret insignificant effects on certification choices globally and tree cover loss in most years as null effects. Assessing the aggregate effects requires the use of DiD estimators that account for differential treatment of observations over time.

2.9 Appendix

2.9.1 Results

Table 2.6: Yearly probability of being certified for concessions with and without IFL

	(1) LPM	(2) Probit
IFL x 2011	0.0505	0.0994
	(1.50)	(1.17)
IFL x 2012	0.0213	0.0703**
	(1.91)	(2.61)
IFL x 2014	-0.00376	-0.00994
	(-0.31)	(-0.16)
IFL x 2015	-0.000136	0.00730
	(-0.01)	(0.11)
IFL x 2016	-0.0379	-0.253
	(-1.03)	(-1.18)
IFL x 2017	-0.0413	-0.273
	(-1.11)	(-1.27)
IFL x 2018	-0.0537	-0.347
	(-1.42)	(-1.62)
IFL x 2019	-0.0548	-0.409
	(-1.28)	(-1.81)
IFL x 2020	-0.0558	-0.410
	(-1.30)	(-1.83)
IFL x 2021	-0.0519	-0.377
	(-1.21)	(-1.68)
IFL		0.430
		(1.20)
Constant	0.207***	-48.47
	(34.05)	(-1.24)
Observations	37422	37374
Year - country FE	Yes	Yes
Concession FE	Yes	No
Concession controls	No	Yes
Weighted by area	Yes	Yes

Notes: Standard errors are clustered at the company level (or the concession in the 153 concessions, where the company name is missing). Treated are concessions with more than 10 % IFL in 2013. Control are concessions without IFL in 2013. All concessions are rented at least since 2013. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 2.10: Joint DiD coefficient on IFL loss rates, with 95% CI.

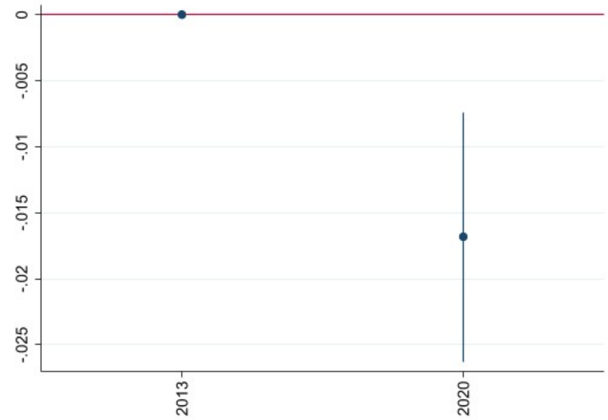


Table 2.7: Non-fire-caused tree cover loss in 2000 IFL in percentage points per year

	(1)	(2)
FSC x 2001	0.000241	-0.000795
	(0.11)	(-0.39)
FSC x 2002	0.00372	0.00293
	(0.88)	(0.69)
FSC x 2003	0.00203	0.000571
	(0.59)	(0.18)
FSC x 2004	0.000438	-0.000146
	(0.25)	(-0.08)
FSC x 2005	0.00750	0.00893
	(1.63)	(1.87)
FSC x 2006	0.0000260	-0.000361
	(0.01)	(-0.16)
FSC x 2007	0.000840	0.0000398
	(0.59)	(0.03)
FSC x 2008	0.00303	0.00177
	(1.47)	(1.18)
FSC x 2009	0.00286	0.00190
	(1.65)	(1.36)
FSC x 2010	0.000774	0.00000744
	(0.58)	(0.01)
FSC x 2011	0.00312	0.00182
	(1.45)	(0.95)
FSC x 2012	0.00166	0.000488
	(0.89)	(0.33)
FSC x 2014	0.00230	0.00137
	(1.33)	(0.97)
FSC x 2015	0.000931	-0.0000205
	(0.34)	(-0.01)
FSC x 2016	0.00841	0.00669
	(1.84)	(1.77)
FSC x 2017	0.000957	-0.000416
	(0.40)	(-0.21)
FSC x 2018	0.00334	0.00250
	(1.74)	(1.35)
FSC x 2019	-0.000430	-0.000916
	(-0.17)	(-0.39)
FSC x 2020	0.00131	0.000429
	(0.60)	(0.24)
Constant	0.00264***	0.00279***
	(9.30)	(11.67)
Observations	12680	12680
Year - country FE	Yes	Yes
Concession FE	Yes	Yes
Year - rental agreement FE	No	Yes
Weighted by IFL area	Yes	Yes

Notes: The dependent variable is non-fire-caused tree cover loss in 2000 IFL in p.p./year. Standard errors are clustered at the company level (or the concession where the company name is missing). Concessions rented at least since 2013 and with more than 0 % IFL in t=0. Treated (control) are those (not) certified in 2013. t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.9.2 Robustness

Figure 2.11: DiD coefficients according to equations 2.1, with 95% CI, but including concessions with rental agreement in 2013-2020.

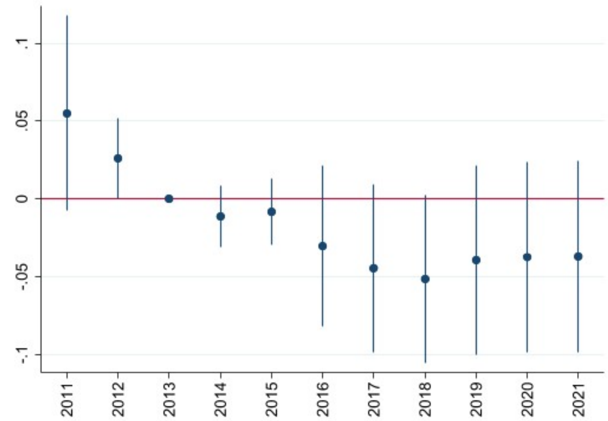
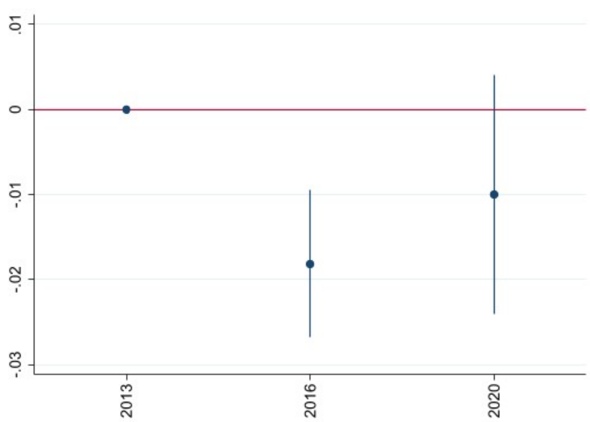
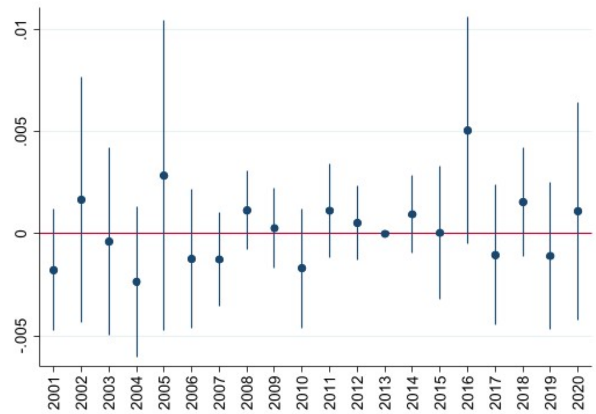


Figure 2.12: DiD coefficients estimates as lower bound estimates of effects on IFL preservation outcomes, with 95% CI, including concessions with rental agreement in 2013-2020. Standard errors are clustered at the company level. The observations are weighted by IFL area.

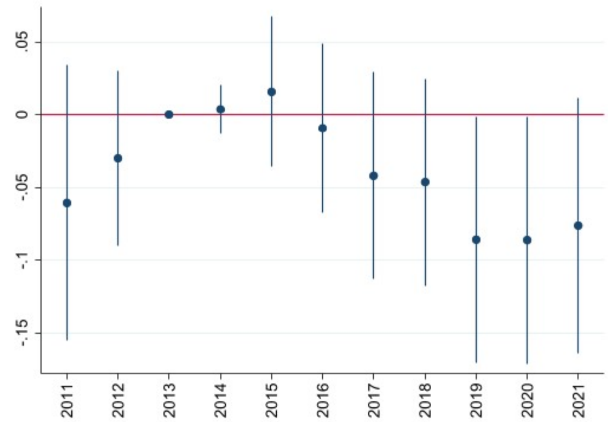


(a) DiD in non-fire-caused tree cover loss, according to equation 2.3.



(b) DiD in non-fire-caused tree cover loss, according to equation 2.3.

Figure 2.13: DiD coefficients according to equations 2.1, with 95% CI, but including concessions with less than 30% IFL.



Bibliography

Azevedo-Ramos, C., Macedo Silva, J. N., and Merry, F. (2015). The evolution of Brazilian forest concessions. *Elementa: Science of the Anthropocene*, 3.

Boudreau, L. (2021). Multinational enforcement of labor law: Experimental evidence from Bangladesh's apparel sector. Working paper.

Cerutti, P. O., Tacconi, L., Nasi, R., and Lescuyer, G. (2011). Legal vs. certified timber: Preliminary impacts of forest certification in Cameroon. *Forest Policy and Economics*, 13(3):184–190.

Dragusanu, R., Giovannucci, D., and Nunn, N. (2014). The Economics of Fair Trade. *Journal of Economic Perspectives*, 28(3):217–236.

Dragusanu, R., Montero, E., and Nunn, N. (2022). The Effects of Fair Trade Certification: Evidence From Coffee Producers in Costa Rica. Working paper.

FAO (2018). State of the world's forests 2018 - Forests pathways to sustainable development. Technical report, FAO, Rome.

FSC (2009). Forest management evaluations (FSC-STD-20-007 (V3-0) EN).

FSC (2011). Policy for the Association of Organizations with FSC (FSC-POL-01-004 V2-0 EN). https://connect.fsc.org/document-centre/documents/retrieve/53eab6fd-915c-4210-a56b-f6b57d67e153?mode=view#viewer.action=download.

FSC (2015). FSC principles and criteria for Forest Stewardship (FSC-STD-01-001 V5-2 EN).

FSC (2018). International Generic Indicators (FSC-STD-60-004 V2-0 EN). https://fsc.org/en/document-centre/documents/retrieve/2a0d35e7-2707-43ad-a493-3192b3daea6a?mode=view#viewer.action=download.

FSC (2020a). The FSC National Forest Stewardship Standard of Cameroon - FSC-STD-CMR-02-2020 EN.

FSC (2020b). The FSC National Forest Stewardship Standard of Russian Federation - FSC-STD-RUS-02.1-2020 EN.

FSC (2020c). The FSC National Forest Stewardship Standard of The Gabonese Republic - FSC-STD-GAB-02-2020 EN.

FSC (2020d). The FSC National Forest Stewardship Standard of The Republic of Congo - FSC-STD-COG-02-2020 EN.

FSC (2020e). Intact Forest Landscapes Guidance for Forest Managers (FSC-GUI-30-010 V1-0 EN). https://connect.fsc.org/document-centre/documents/resource/424.

FSC (2020f). Public Certificate Search. https://info.fsc.org/certificate.php#result.

FSC (2022). FSC Directive on FSC Forest Management Evaluations.

Greenpeace International (2021). Destruction: Certified. Technical report. https://www.greenpeace.org/static/planet4-international-stateless/2021/04/b1e486be-greenpeace-international-report-destruction-certified_finaloptimised.pdf.

Greenpeace International, World Resources Institute, Wildlife Conservation Society, Department of Geographic Sciences, University of Maryland, Transparent World, and WWF-Russia (2021). Intact Forest Landscapes. https://www.intactforests.org/index.html.

Gutierrez Garzon, A. R., Bettinger, P., Siry, J., Abrams, J., Cieszewski, C., Boston, K., Mei, B., Zengin, H., and Yeşil, A. (2020). A Comparative Analysis of Five Forest Certification Programs. *Forests*, 11(8):863.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., and Townshend, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 342(6160):850–853.

Haurez, B., Daïnou, K., Vermeulen, C., Kleinschroth, F., Mortier, F., Gourlet-Fleury, S., and Doucet, J.-L. (2017). A look at Intact Forest Landscapes (IFLs) and their relevance in Central African forest policy. *Forest Policy and Economics*, 80:192–199.

Isman, J. J. (2023). Auditor leniency and certification in voluntary forest certification. Unpublished.

Karsenty, A. (2019). Certification of tropical forests: A private instrument of public interest? A focus on the Congo Basin. *Forest Policy and Economics*, 106(101974).

Kleinschroth, F., Garcia, C., and Ghazoul, J. (2019). Reconciling certification and intact forest landscape conservation. *Ambio*, 48(2):153–159.

Lee, P. G. (2009). Caution Against Using Intact Forest-Landscapes Data at Regional Scales. *Ecology and Society*, 14(1):resp1.

Moog, S., Spicer, A., and Böhm, S. (2015). The politics of multi-stakeholder initiatives: The crisis of the forest stewardship council. *Journal of Business Ethics*, 128:469–493.

Potapov, P., Hansen, M. C., Laestadius, L., Turubanova, S., Yaroshenko, A., Thies, C., Smith, W., Zhuravleva, I., Komarova, A., Minnemeyer, S., and Esipova, E. (2017). The last frontiers of wilderness: Tracking loss of intact forest landscapes from 2000 to 2013. *Science Advances*, 3(1).

Potapov, P., Yaroshenko, A., Turubanova, S., Dubinin, M., Laestadius, L., Thies, C., Aksenov, D., Egorov, A., Yesipova, Y., Glushkov, I., Karpachevskiy, M., Kostikova, A., Manisha, A., Tsybikova, E., and Zhuravleva, I. (2008). Mapping the World's Intact Forest Landscapes by Remote Sensing. *Ecology and Society*, 13(2).

PPECF (2014). Motions for the 2014 FSC General Assembly. http://www.ppecf-comifac.com/files/motion_n%C2%B0_65.pdf.

Ptichnikov, A. (2021). Impact assessment research of Motion 65 implementation for boreal forests of the Russian Federation. https://fsc.org/sites/default/files/2021-10/Motion%2034%20Russia%20report%20TC%20FSC%20Russia.pdf.

Rotherham, T. (2016). Forest Management Certification in Canada and Around the World. *The Forestry Chronicle*, 92(02):142–146.

Sist, P., Piponiot, C., Kanashiro, M., Pena-Claros, M., Putz, F. E., Schulze, M., Verissimo, A., and Vidal, E. (2021). Sustainability of Brazilian forest concessions. *Forest Ecology and Management*, 496(119440).

The IFL Mapping Team (2021). The Intact Forest Landscapes 2000/2013/2016/2020. https://intactforests.org/shp/IFL_readme_2021.pdf.

Tritsch, I., Le Velly, G., Mertens, B., Meyfroidt, P., Sannier, C., Makak, J.-S., and Houngbedji, K. (2020). Do forest-management plans and FSC certification help avoid deforestation in the Congo Basin? *Ecological Economics*, 175(106660).

Tyukavina, A., Potapov, P., Hansen, M. C., Pickens, A. H., Stehman, S. V., Turubanova, S., Parker, D., Zalles, V., Lima, A., Kommareddy, I., Song, X.-P., Wang, L., and Harris, N. (2022). Global Trends of Forest Loss Due to Fire From 2001 to 2019. *Frontiers in Remote Sensing*, 3:825190.

Ulybina, O. and Fennell, S. (2013). Forest certification in Russia: Challenges of institutional development. *Ecological Economics*, 95:178–187.

Venier, L., Walton, R., Thompson, I., Arsenault, A., and Titus, B. (2018). A review of the intact forest landscape concept in the Canadian boreal forest: Its history, value, and measurement. *Environmental Reviews*, 26(4):369–377.

Watson, J. E. M., Evans, T., Venter, O., Williams, B., Tulloch, A., Stewart, C., Thompson, I., Ray, J. C., Murray, K., Salazar, A., McAlpine, C., Potapov, P., Walston, J., Robinson, J. G., Painter, M., Wilkie, D., Filardi, C., Laurance, W. F., Houghton, R. A., Maxwell, S., Grantham, H., Samper, C., Wang, S., Laestadius, L., Runting, R. K., Silva-Chávez, G. A., Ervin, J., and Lindenmayer, D. (2018). The exceptional value of intact forest ecosystems. *Nature Ecology & Evolution*, 2(4):599–610.

Williams, J., Lapointe, U., Paixao, C., Zanotta, D., and Clark, T. (2021). Assessing the Impact of Implementing FSC's Protection Measures for Intact Forest Landscapes in Canada. Technical report. https://fsc.org/sites/default/files/2021-10/FSC%20CA%20_%20FINAL_M34Report.IFL%20impact%20assessment.pdf.

Zhuravleva, I., Turubanova, S., Potapov, P., Hansen, M., Tyukavina, A., Minnemeyer, S., Laporte, N., Goetz, S., Verbelen, F., and Thies, C. (2013). Satellite-based primary forest degradation assessment in the Democratic Republic of the Congo, 2000–2010. *Environmental Research Letters*, 8(024034).

Chapter 3

The effect of accreditation regulation on the credibility of voluntary certification

Johanna Isman[1]

[1]Toulouse School of Economics

3.1 Introduction

Due to adverse selection and moral hazard, asymmetric information about quality attributes causes inefficiencies in many markets (Akerlof, 1970; Arrow, 1963). Credible quality signaling is particularly challenging in the case of *credence attributes*, which can only be revealed at a high cost through inspection of the production process or testing technologies (Darby and Karni, 1973). Examples are pesticide residues in food or ethical treatment of workers abroad. Buyers learn about credence attributes only with a very low probability, based on some level of public monitoring, for instance, through governmental or non-governmental organizations, investigative journalists, or researchers.

This paper theoretically studies the potential of *accreditation* to increase the credibility of signaling credence qualities through third-party certification. While third-party certification is the external assessment and disclosure of a firm's compliance with a quality standard through inspection of production processes or tests of products,[2] accreditation is the external assessment and disclosure of the third party's competence to certify. Accreditation can improve a certifier's incentives to assess and disclose firms' compliance rigorously and truthfully. It typically involves inspections, is based on international accreditation standards, and is often conducted by public authorities (IAF, 2018a).

This paper's core is a model that explains the effect of the establishment and competition regulation of accreditation bodies on the credibility of the certification of voluntary credence standards. I provide motivating evidence from Uruguay to illustrate the potential role of establishing an internationally recognized national accreditation body in spreading certification.

In the model, atomistic buyers, a certifier, and zero to two accreditation bodies play a repeated game with an infinite horizon. The supplier is a new entrant in each stage game to focus the analysis on cases where truthful signaling without certification is not incentive compatible. I consider a monopolist certifier as I want to concentrate on the market structure in which incentive compatibility of truthful certification is most likely. The baseline model without accreditation then reproduces a result from the literature: Even a monopolist certifier may not have sufficient incentives to assess and disclose the supplier's compliance truthfully

[2] Certifiers are independent of the owners and typically also from the developers of the standard. Standards can be set at many different levels. See Figure 3.3 in Appendix 3.8.1 for an overview. The organizations that set a standard usually own it. That means they charge license fees when a certifier certifies according to the standard.

if they do not care enough about the future or if the probability that fraud becomes known to buyers is too low (Strausz, 2005).

My contribution is the introduction of accreditation, a second layer of monitoring. Obtaining accreditation has two opposing effects for the certifier's incentives to cheat: On the one hand, it requires them to spend a certain fraction of the truthful monitoring cost to demonstrate their competence. The fraction they need to spend is a proxy of the accreditation quality since it reduces their incentives to cheat. On the other hand, accreditation requires the payment of a fee to the accreditation body in each period, which reduces future profits and thus reduces potential loss in future profits from cheating. The fee is increasing in the quality of accreditation. The optimal level of accreditation quality depends on the strengths of the two opposing effects for the certifier's incentives.

Based on this model, I derive the following conclusions. First, accreditation can only prevent certifiers from cheating if buyers are sophisticated enough to demand accreditation from certifiers or if accreditation is compulsory. Second, assuming sophisticated buyers, I show that a profit-maximizing accreditation body reduces fraud in the certification market whenever a welfare-maximizing accreditation body can do so. This result does not hold if buyers naively trust in any accreditation. Third, under a few plausible assumptions, the model suggests that the optimal accreditation quality is, on average, higher in developing than advanced economies. Fourth, it shows that developing and small economies should only establish their own national accreditation body if they cannot find a foreign accreditation body of sufficiently high quality and sufficiently low transportation cost.

The theoretical analysis is motivated by an empirical analysis of the association of the availability of an internationally recognized national accreditation body with the share of companies that seek certification. Using data from the World Bank Enterprise Surveys, I compare changes in certification rates in Uruguay and Nicaragua before and after the official international recognition of Uruguay's national accreditation body. The results suggest that the availability of globally recognized national accreditation is associated with increased certification rates. Increased certification rates might be driven by increased credibility of certification through accreditation, such that this case illustrates the potential of accreditation highlighted in the model. However, the setting does not allow to establish causality with certainty.

The questions analyzed in this paper are policy-relevant for the following reasons. Demand for accreditation is increasing globally, partly driven by downstream buyers preferring to buy products certified by an accredited certifier (Heras-Saizarbitoria and Boiral, 2019). Many developing countries still have to decide whether to establish their national accreditation body. International organizations such as UNIDO have supported some such endeavors (UNIDO et al., 2017). Donors have to choose in which cases to contribute. Finally, the regulation of accreditation bodies started only about fifteen years ago, and it may still need refinements, particularly concerning the degree to which private accreditation bodies are allowed. Approaches differ by country, from little regulation in the United States to a restriction to a single accreditation body appointed by the state in the European Union and China.

The paper is organized as follows. I first give an overview of the literature on certification and outline in what ways I contribute to it. Section 3.2 provides an overview of the organization of accreditation worldwide and motivating evidence about its relevance. Section 3.3 presents the baseline model in which no accreditation is available. In section 3.4, the model is extended to account for certification. I analyze first the case of a monopoly accreditation body and second, Bertrand competition among different accreditation bodies. Section 3.5 shows how the results of section 3.4 change if buyers do not understand the value of accreditation. In section 3.6, I adapt the results to account for the case where buyers only trust accredited certification. Section 3.7 summarises and gives an outlook on potential extensions to this paper.

Literature and contribution

This paper contributes to the literature on the economics of certification. The syntheses of Dranove and Jin (2010) and Bonroy and Constantatos (2014) provide helpful overviews. Many papers cover companies' decisions to seek certification between cost and benefits. However, fewer models endogenize the choices of the certifying company (henceforth the *certifier*) (Farhi et al., 2013), as I do in this paper.

This paper focuses on the certifiers' choice between honest and fraudulent certification. I concentrate on for-profit certifiers.[3] I follow former papers regarding the certifier as a black box (Bolton et al., 2012; Giannakas, 2002; Strausz, 2005).[4] I do not model certifiers' public

[3]For many standards, such as ISO management system standards, certifiers are typically for-profit companies. Two of the biggest players are Bureau Veritas and SGS. For ethical standards, such as Fairtrade, certifiers are typically non-profit companies.

[4]I do not analyze potential principal-agent issues between the certifier as a company and its agents. The

choices, such as the limited disclosure of inspection results, which Lizzeri (1999), Farhi et al. (2013), Stahl and Strausz (2017) and Kastl et al. (2018) have analyzed. I instead adopt a result of Lizzeri (1999) as an assumption. Lizzeri (1999) shows that, under monopoly and oligopoly certifiers, there exists an equilibrium in which the certifiers only publish the names of successful applicants. This equilibrium is in line with the practice of most certifiers: Certifiers typically do not disclose unsuccessful applicants and grant applicants sufficient time to correct undetected nonconformities.[5] Hence, any certification applicant in my model obtains certification if they pay the necessary cost. Thus, the model focuses on the firm's moral hazard rather than adverse selection.

Many certification models, including the papers cited above, assume full commitment of the certifier to the disclosure rule. That implies that the information the consumer obtains is limited but correct (Giannakas, 2002). The papers that investigate the credibility of certifiers are almost entirely based on reputation building.[6] Ratings of financial products are the most-studied examples of certification (Bar-Isaac and Shapiro, 2011; Bolton et al., 2012; Faure-Grimaud et al., 2009; Mathis et al., 2009; Opp et al., 2013; Skreta and Veldkamp, 2009). Reputation building plays a crucial role in this industry since investors can systematically observe the overall accuracy of credit rating agencies ex-post. Nevertheless, various scandals have brought attention to severe conflicts of interest in this industry (Bolton et al., 2012). The literature on certification in financial markets finds that increased competition exacerbates asymmetric information (Faure-Grimaud et al., 2009; Skreta and Veldkamp, 2009). Without a particular focus on the financial market, Strausz (2005) obtains an even more striking result: He shows that, under low discount rates, collusion can only be deterred if certification fees are above the monopoly price. He obtains this result even though he assumes that

certifiers' agents are (1) employed or contracted auditors and (2) the employees that evaluate the auditors' inspection reports and make the certification decisions. These agents can be subject to moral hazard, adverse selection, or corruption. However, a sufficiently committed certifier can implement mechanisms that minimize such issues. For example, they could send multiple auditors to each client or double-check auditors' reports. In an RCT on environmental audits in India, Duflo et al. (2013) also show that introducing (1) random auditor assignment, (2) flat payment by companies to the auditors, (3) random backchecks of reports through an upstream agency, and (4) linking salaries of auditors to backcheck results, altogether improve auditing accuracy. Moreover, due to the lack of systematic asymmetric information between top management and the employees evaluating inspection reports, I do not suspect many difficulties for the company to implement their policies on report evaluations and certification decisions.

[5]There are, however, standards for which certifiers have to publish inspection reports, as MSC Fisheries (Auld and Gulbrandsen, 2010). One explanation for this might be the separation of standard setting and certification. The standard setters, who often also prescribe disclosure rules, have different incentives than the certifiers. Ways of modeling those can be found in Swinnen et al. (2015).

[6]This literature took off most significantly with the contributions of Biglaiser (1993); Biglaiser and Friedman (1994).

customers perfectly observe quality after their purchase. The finding implies that there can be fraudulent certification even under a monopoly certifier. My model builds on that idea and analyses whether accreditation can prevent fraudulent certification.

There is significant evidence of incorrect certification, especially for credence goods. For example, the exploratory study of Heras-Saizarbitoria and Boiral (2019) shows large-scale fake certifications of the ISO 9001 management system standards in China. The study finds that fraudulent certification exists even for accredited certifiers. Heras-Saizarbitoria and Boiral (2019) finds this problem mainly for certifiers accredited by private accreditation bodies which are not internationally recognized.[7] My model, therefore, distinguishes between welfare-maximizing and profit-maximizing accreditation bodies. This distinction allows me to give some suggestive explanations on the findings of Heras-Saizarbitoria and Boiral (2019).[8]

3.2 Context and motivating evidence

3.2.1 The regulation of accreditation

This section explains the functioning and regulation of accreditation to motivate some critical assumptions made in the model.

Accreditation bodies are only widely recognized if they work according to ISO standard 17011.[9] This standard defines an *accreditation body* as an "authoritative body that performs accreditation" (ISO, 2017). This authority is typically derived from the government (ISO, 2017). It can also be derived from designation by private standard owners.[10] Accreditation is

[7]I will explain the role and system of international recognition of accreditation bodies in section 3.2.2.

[8]Cases of fraud in certification have been observed even in advanced economies. Examples are the Anderson-Enron accounting scandal and the case in which Sony Pictures was fined for having certified their films based on invented reviews (Strausz, 2005). Strausz (2005) claims that these cases are out-of-equilibrium events that were sufficiently punished to deter systematic fraud in the market. However, other cases cast doubt on that and show the difficulties of monitoring certification. TÜV Rheinland, for example, was accused of misreporting construction safety in the garment factories of Rana Plaza whose collapse in 2013 cost the lives of over 1100 workers (OECD NCP, 2018). In 2013, a French court held the company liable for falsely certifying ruptured breast implants (Van Leeuwen, 2017). Despite such incidences, TÜV Rheinland's revenues have been steadily growing in the subsequent years, what implies that their reputation was not strongly affected by the accusations (TÜV Rheinland AG, 2015, 2016, 2017).

[9]See Heras-Saizarbitoria and Boiral (2019); ISO (2017). ISO 17011 has been internationally harmonized and developed by the International Organization for Standardization (ISO). It describes how accreditation bodies should be organized and work.

[10]This is, for example, the case for Assurance Services International, which sustainability standards such as Marine Stewardship Council designated as their sole accreditation body (Auld and Gulbrandsen, 2010; MSC, 2019).

the attestation that a certifier is competent to assess firms' conformity with specific standards, typically based on ISO standards for certifiers. (ISO, 2017).[11] These standards are not very detailed, so accreditation bodies still have some discretion in their assessments of certifiers.[12] This is why this paper allows the quality of accreditation to vary between different bodies.

The state often regulates accreditation. In the European Union, each member state can only have one accreditation body appointed by the state.[13] For voluntary standards, private certification bodies are not prevented from seeking accreditation from non-European bodies.[14] Other advanced economies such as the USA allow activities of accreditation bodies that the state did not appoint even with-in the USA (NIST 2018).[15] Also, many emerging and developing countries, such as Argentina, India, and Kenya, do not regulate their accreditation markets a lot.[16] Stricter regulations, similar to Europe, apply, for example, in China and Colombia.[17][18] These varying regulations motivate analyzing both the case of monopoly accreditation and competition among foreign and national accreditation bodies.

Figure 3.1(a) serves as an illustration of the accreditation and certification system. It shows a typical case in advanced economies: The certified company (Airbus), the certifier (AFNOR),

[11]There are ISO standards for various types of certifiers, for example, ISO 17065 for certifiers of products, processes, and services, and ISO 17021 for certifiers of management systems. Accreditation attests to their competence by checking compliance with these standards. They include organizational requirements favoring impartiality and thoroughness in the certification process. For example, accredited certifiers and their subsidiaries must be independent of all companies involved in the value chain of certified products or processes (ISO, 2012).

[12]Standard ISO 17065:2012 has about 25 pages, ISO 17021 about 35.

[13]If certifiers in Europe want to get accredited by a European accreditation body, they must refer to the accreditation body of their headquarters' state (Reg. (EC) No 765/2008). If that accreditation body did not offer the required service, they could refer the certifier to the accreditation body appointed by another EU member state. The EU has taken measures against continued activities of other private accreditation bodies (European Commission, 2017; Ref. Ares, 2014).

[14]The British BSI Group, for example, is accredited by the US accreditation body ANAB (ANAB 2018). These accreditations will not be recognized by the European states (European Commission, 2014).

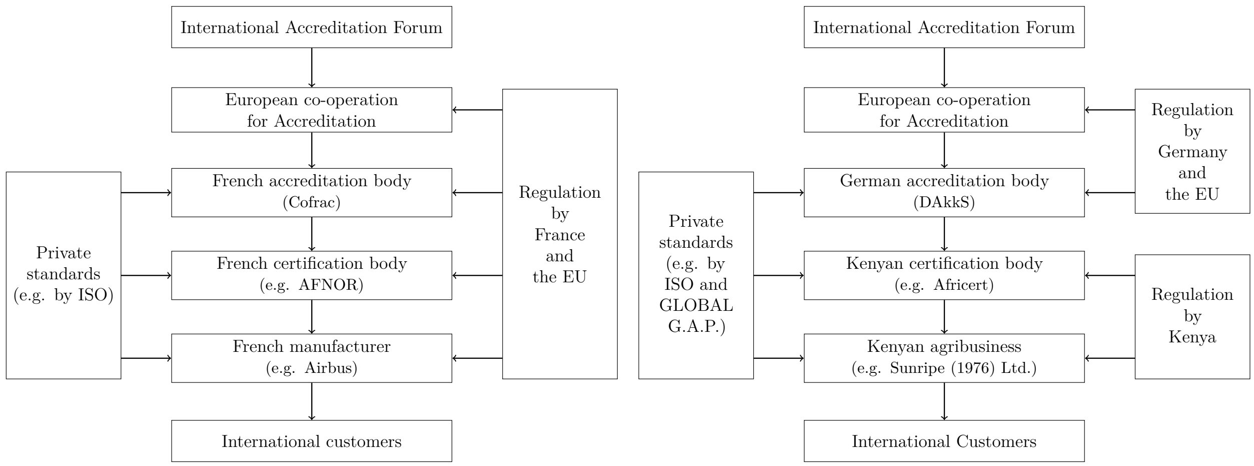
[15]Also, some private accreditation bodies specialized in particular standards and transnationally active are based in the USA, such as the International Organic Accreditation Service (IOAS). Such bodies build their credibility on their reputation and creation through NGOs but are also increasingly integrated with the international peer-evaluation system (IOAS, 2018).

[16]See AfriCert (2017); Competition Appellate Tribunal (2014); Eguino (2018).

[17]See CNAS (2012); UNIDO et al. (2017).

[18]What the accreditation bodies in most economies have in common is that they are created with the support of the state (Quality Council of India, 2015; UNIDO et al., 2017). An important reason for such cooperation is that accreditation activities are often not profitable in their first years of business but have positive externalities for the credibility of quality assurance and thus for the development of an economy (Eguino, 2018; UNIDO et al., 2017). Many states use accreditation bodies also to support the controls of public regulations (UNIDO et al., 2017). In the long term, many accreditation bodies try to become financially independent, as dependence on public resources can also affect their credibility negatively (UNIDO et al., 2017).

Figure 3.1: Value chains of recognition, accreditation, certification and product sale



(a) In manufacturing in France with national accreditation **(b)** In agriculture in Kenya with foreign accreditation

and the accreditation body (Cofrac) are all based in the same country.[19] In developing countries, certifiers are often accredited by foreign accreditation bodies as shown in Figure 3.1(b): The German accreditation body DAkkS accredited the Kenyan certifier AfriCert to certify farms according to the GLOBALG.A.P. standard.

3.2.2 Motivating evidence: The role of internally recognized national accreditation bodies

In this section, I provide motivating evidence of the role of national accreditation bodies in increasing demand for certification, which might be related to increased credibility. I expect to see increased demand for *internationally recognized* certification not directly after establishing the accreditation body but after the international recognition of that body through peer review. In a peer review, accreditation bodies from other countries check compliance with the ISO standard 17011 through regional or international associations, such as the European co-operation for Accreditation and the International Accreditation Forum in Figure 3.1.[20] In international trade, this likely increases the recognition of certifiers accredited by the

[19] See AFNOR Group (2018); Cofrac (2018).

[20] Figure 3.4 in Appendix 3.8.2 shows the countries which have officially gained international or regional recognition in that way. Membership in a regional accreditation group, such as the African Accreditation Cooperation (AFRAC) or the European co-operation for Accreditation, facilitates the recognition of an accreditation body in the limited geographical scope of that organization. The International Accreditation Forum, in turn, reviews the regional groups. That review gives their members certain credibility, even internationally (IAF, 2018a; 2018b).

peer-reviewed body. In the following, I test whether the international recognition of the Uruguayan accreditation body through peer review is associated with increased certification demand. Here, *certification demand* is the decision of companies to obtain internationally recognized certification, not the demand for products from certified companies.

3.2.2.1 Empirical specification

I use a difference-in-differences (DD) estimator, which compares changes in the certification status y_{ict} of companies affected by that recognition to changes in the certification status of companies unaffected by this recognition. I consider a Linear Probability Model with two-way fixed effects:

$$y_{ict} = \alpha_i + \beta Post_t + \delta\, Treated_i \times Post_t + \varepsilon_{it} \tag{3.1}$$

i indicates the company and t the period. $Treated$ is a dummy variable that is one if the company is located in Uruguay, the country whose accreditation body became internationally recognized through peer review in 2011 (International Accreditation Forum, 2018b). Since the international recognition affected all certifiers in Uruguay, it likely also affected all companies in Uruguay. The control group consists of companies located in Nicaragua, whose trend in certification is approximately parallel to that of Uruguay prior to 2011, as I show in the results from a placebo test comparing 2006 and 2010 in Table 3.3 in the Appendix. Both Nicaragua and Uruguay have had national accreditation bodies since the 1990s, but the international community has not peer-reviewed and recognized Nicaragua's accreditation body until today. $Post$ is a dummy that takes on the value one if the observation is from the period after 2011 and zero otherwise.

3.2.2.2 Data

I use panel data from the World Bank Enterprise Surveys, which are representative samples of private firms (World Bank, 2017). The survey covers Nicaraguan firms in 2006, 2010, and 2016 and Uruguayan firms in 2006, 2010 and 2017. I use 2016/2017 as the period after the treatment, 2010 as the period prior to the treatment, and 2006 to test for a parallel trend from 2006 to 2010 in Table 3.3. Table 3.2 in the Appendix shows descriptive statistics of the data. The table illustrates that the difference in the share of firms with internationally recognized certification between Nicaragua and Uruguay changed after the treatment, but the differences in other observables did not change much.

3.2.2.3 Results and discussion

Table 3.1 presents the main results. Column (1) follows specification 2.3, and column (2) additionally controls for two company-level variables: the percentage of the firm owned by private domestic individuals, companies, or organizations; and the years of experience in the company's sector the top manager has. The reported standard errors are robust to heteroskedasticity. Both specifications result in a positive DD estimate that is significant at a 1% significance level. The international recognition of the Uruguayan accreditation body is associated with a 21 percentage point increase in the likelihood that a company in Uruguay has internationally recognized certification, compared to Nicaragua.

These results suggest that the international recognition of the Uruguayan accreditation body positively affected the likelihood of Uruguayan companies being certified. If companies choose certification rationally and under the common trend assumption, this points to an increased net benefit of certification for companies due to the availability of internationally recognized accreditation at a reduced cost through a national accreditation body. This increase might be due to risen credibility of certification by certifiers previously not accredited by an internationally recognized accreditation body or due to the reduced cost of certification by certifiers previously accredited by a foreign, more expensive accreditation body.

However, the result is only suggestive. Firstly, the small sample reduces reliability in the regression output. Secondly, the data for the post-treatment period come from two different years: 2016 for Nicaragua and 2017 for Uruguay. If many companies obtained certification in 2017 for some reason, I overestimate the result. Thirdly, more favorable development in Uruguay compared to Nicaragua from 2010 to 2016/2017 in dimensions not caused by the treatment might explain the rise in certification demand. However, GDP per capita growth between 2010 and 2016 has been, on average, higher in Nicaragua than in Uruguay (World Bank, 2019).

3.3 The baseline model

The extensive-form game of my model is shown in Figure 3.2. Time is discrete. The horizon is infinite. There are three players.

The players are a buyer or a mass 1 of identical buyers, a supplier, and a certifier. The

Table 3.1: Main results

	(1)	(2)
Treated x Post	0.210***	0.211***
	(3.89)	(3.89)
Post	-0.183***	-0.181***
	(-4.88)	(-4.78)
Constant	0.241***	0.295**
	(12.65)	(3.20)
Observations	464	464
Controls	No	Yes
Company FE	Yes	Yes
F-stat.	12.17***	6.18***

Notes: Estimates from regression of specification 2.3 in column (1), adding company controls in column (2) (percentage of domestic shareholders and the years of sector experience of the top manager). t statistics in parentheses, robust to heteroskedasticity. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

certifier is a monopolist.[21] Buyers are atomistic. The buyers and the certifier are long-lived players. The supplier is short-lived. All players are risk-neutral.

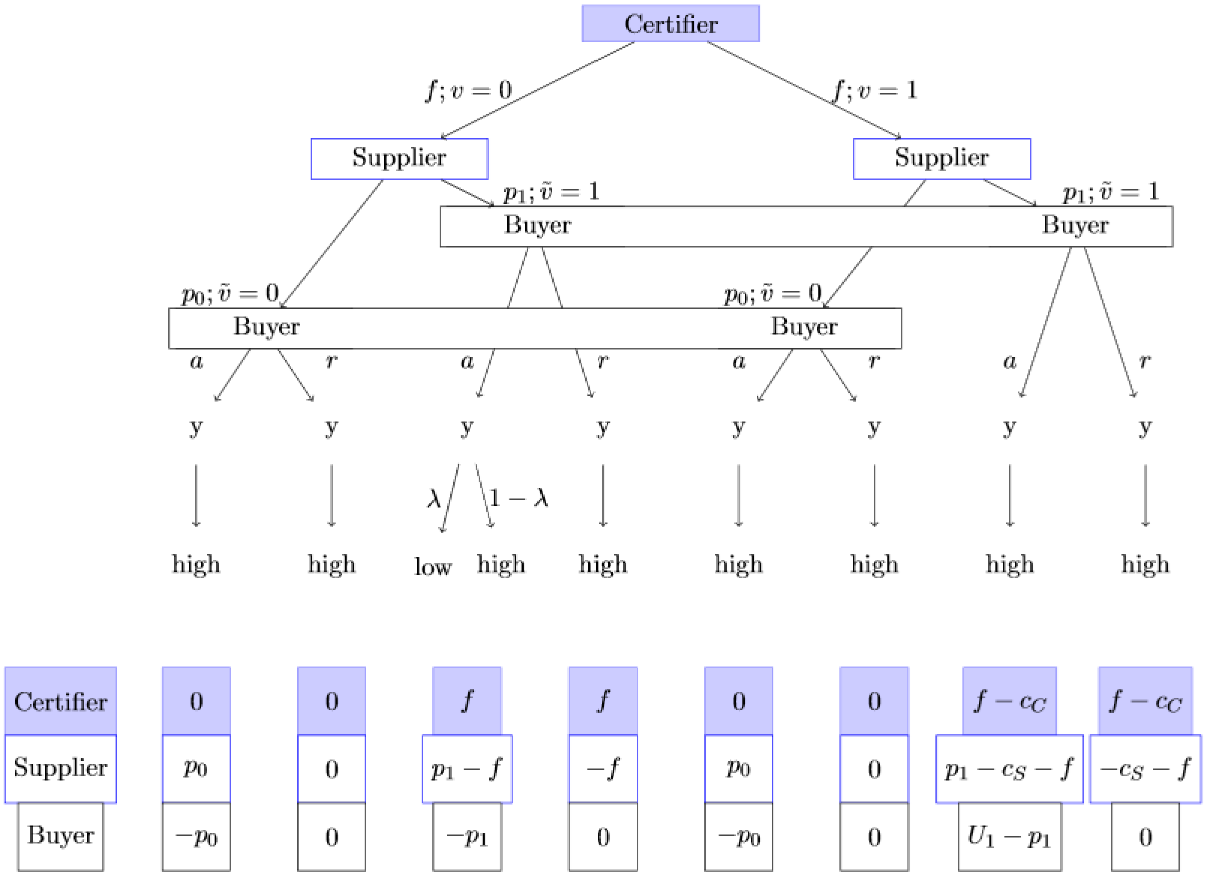
Remark. By assuming that the supplier is short-lived, I abstract from the case where the supplier can build a reputation. This assumption allows me to focus on the incentives of the certifier since it implies that the supplier cannot credibly signal quality without credible certification.

The action sets are the following: The certifier offers a certification contract (v, f). $f \in \mathbb{R}_+$ is the fee for which they certify the supplier. $v \in \{0, 1\}$ is the quality level that the certifier verifies. The certifier tells the public that the supplier has high quality 1 whenever their quality exceeds v.[22] In other words, the certifier grants a high-quality label $\tilde{v} = 1$. If $v = 1$, the certifier verifies that the supplier produces high quality. If $v = 0$, the certifier does not verify high quality and provides a fake certificate. The certifier's cost of monitoring a true

[21]This is a simplification of the theoretical insight that reputation mechanisms can only work if prices are above marginal cost (Darby and Karni, 1973) and the empirical observation that certification markets are oligopolistic: A few big players such as SGS and Bureau Veritas are dominant (Galland, 2017)

[22]Third-party certification is typically not a one-shot check if a producer complies with a standard. Rather, the certifier notes non-conformities and sets a time frame within which the producer has to correct those non-conformities (ISO, 2012, 2015a). Thus, compliance is rather a matter of moral hazard than adverse selection. In other words, I apply the empirical observation, and theoretical result of Lizzeri (1999) that certifiers usually do not publish rejected applications.

Figure 3.2: The baseline model



quality level v is (vc_C), which is 0 if the certifier cheats and $c_C \in \mathbb{R}_{++}$ otherwise.[23] The supplier chooses if they accept or reject the certification contract. In other words, they choose a quality label $\tilde{v}$:

$$\tilde{v} = \begin{cases} 1 & \text{if the supplier accepts the certification contract} \\ 0 & \text{otherwise} \end{cases}$$

Accepting a certification contract (v, f) to obtain a high-quality label $\tilde{v} = 1$ requires investing in product quality v at cost (vc_S). This cost is 0 if the certifier cheats and $c_S \in \mathbb{R}_{++}$ otherwise. The supplier further chooses a price $p_{\tilde{v}} \in \mathbb{R}_+$ at which they offer a product of quality label $\tilde{v}$ to the buyers. The buyers accept or reject.

[23]These costs do not only include the cost of actual auditing but also the cost of acquiring and maintaining the competence and infrastructure which are needed to certify correctly.

Remark. I assume that possible quality levels are exogenous and discrete not only for simplicity. It also allows focusing on the certification of quality standards which an independent standard-setting organization fixed and an individual company cannot influence.[24] Buyers may view the quality of complying with such a standard as binary: A company truly follows ISO 9001 or not, a product is Fairtrade or not.[25] This is also why I assume that the certifier only certifies high quality $\tilde{v} = 1$. The definition of the high-quality level as 1 is without loss of generality. Certification is always a positive signal; setting it at some level $\bar{v} > 0$ would not affect the qualitative results.

The players' information is as follows. At the end of each period $t \geq 1$, there is a public signal about honesty y: The signal can be high or low. The signal is always high if nobody cheats, i.e., if the supplier did not get certified or if the certifier's effort level was one:

$$\text{If } \tilde{v} = 0 \text{ or } v = 1, y = \text{"}high\text{"}$$

If the certifier cheats, the signal is sometimes high and sometimes low:

$$\text{If } v = 0 \text{ and } \tilde{v} = 1, y = \begin{cases} \text{"high" with probability } 1 - \lambda \\ \text{"low" with probability } \lambda \end{cases}$$

Thus, λ is the probability of detecting fraudulent certification or the level of public monitoring.

The game is common knowledge. Thus, all players know the history of public signals, the players, the action sets, the information sets, the timing, and the payoffs (see below). The certifier and the supplier observe all actions chosen. The buyers observe the quality label $\tilde{v}$, not the certification contract (v, f). Thus, the buyers do not observe whether the certifier cheats or not. They also do not observe their stage payoff.

The Timing of the stage game is:

Stage 1: The certifier offers a certification contract (v, f).

Stage 2: The supplier accepts or rejects the certification contract (choose quality label $\tilde{v}$) and offers the product at a price $p_{\tilde{v}}$ to the buyers.

[24] See Appendix 3.8.1, Figure 3.3 for a complete overview of standard setters. I focus here on large standard-setting organizations where an individual company has little influence on the standard.

[25] Unless due to intrinsic motivation, the supplier has no incentive to comply only with some Fairtrade criteria since this does not lead to certification and price premia.

Stage 3: The buyers purchase the product or not.

The payoffs of each stage game are as follows.

The valuation of the buyers is

$$U_v = \begin{cases} U_1 > c_S + c_C & \text{if } v = 1 \\ U_0 = 0 & \text{if } v = 0 \end{cases}$$

Thus, the buyers obtain a surplus S_B from buying a product of quality v at price $p_{\tilde{v}}$:

$$S_B = U_v - p_{\tilde{v}}$$

Remark. The most straightforward way is to see the buyers as final consumers who draw utility U_v from the true quality v, not the quality label $\tilde{v}$. The idea is that the true quality affects the buyers' well-being but that the buyers cannot attribute this effect to the product. One can think of long-term health or environmental impacts that affect well-being through the aggregate level of environmental protection. Alternatively, the buyers could also be firms that obtain an unaccounted benefit U_v if their supplier has implemented a meta-standard. The most widespread meta-standards are the ISO 9000 management system standards. One objective of these standards is to increase the effectiveness of firms, both in product and service quality (ISO, 2015b). Implementing such a management system can facilitate interaction with buyers and increase the profits from that interaction in the long term. The buyers cannot always directly observe such benefits in a given period and may later be unable to attribute it to the certification. [26]

The supplier obtains a surplus S_S from selling a product with quality label $\tilde{v}$ at price $p_{\tilde{v}}$:

$$S_S = \begin{cases} p_{\tilde{v}} & \text{if } \tilde{v} = 0 \\ p_{\tilde{v}} - vc_S - f & \text{if } \tilde{v} = 1 \text{ under certification contract } (v, f) \end{cases}$$

[26] It is also possible to extend the model to account for a retailer with a long-term relationship with final consumers. However, this would require the model to be adapted so that the retailer gains some profit which they would lose if they lost their customers. Alternatively, I would need to assume that they are honest and, to some extent, altruistic: Then they would disvalue a low-quality product through observing negative effects on their customers in the long run (see Kitzmueller and Shimshack (2012) on the role of intrinsically motivated agents in businesses).

The certifier obtains a surplus S_C from a contract (v, f):

$$S_C = f - (vc_C)$$

The long-time payoffs of the certifier and the buyers are the discounted sums of the expectation of stage game payoffs for infinitely many periods. Their common discount factor is $\delta \in [0, 1]$.

As a tiebreaking rule, I assume that all players prefer high quality: In case two actions or strategies lead to the same payoff, each player favors the action or strategy that leads to an equilibrium that ensures high quality.

3.3.0.0.1 Social optimum Assume a utilitarian social planner can determine all actions. The total welfare W in each stage game is

$$W = S_B + S_S + S_C = U_v - v(c_S + c_C) = \begin{cases} U_1 - (c_S + c_C) > 0 & \text{if } v = 1 \\ 0 & \text{if } v = 0 \end{cases} \tag{3.2}$$

Recall that the action sets do not contain the possibility that the supplier provides high quality without certification. In other words, it is not possible that $v = 1$ and $\tilde{v} = 0$ simultaneously. I assume that as a shortcut to account ex-ante for the fact that the short-lived supplier would never have the incentive to play these actions. If they would, this would have been the first best since it would give welfare $W = U_1 - c_S$.

Assuming this is impossible, the social planner's solution to the remaining problem is to induce certification and a truly high-quality level ($v = 1$) in each period $t \geq 1$. Thus, without accreditation, the social planner's (second best) solution is achieved by any "High-Quality Equilibrium" defined as follows:

Definition 1. *Any equilibrium in which the true quality level is $v = 1$ in each period $t \geq 1$ is a "High-Quality Equilibrium."*

Throughout the paper, I illustrate the model and results with the same numerical specification, separate from the general results:

Illustration. Assume that the buyers' valuation equals the true quality of the product $U_v = v$

and that $c_S = c_C = \frac{1}{4}v^2$. Then, total welfare is

$$W = v - \frac{1}{4}v^2 - \frac{1}{4}v^2 = \begin{cases} \frac{1}{2} & \text{if } v = 1 \\ 0 & \text{if } v = 0 \end{cases}$$

□

3.3.0.0.2 Equilibrium I only look for public-perfect equilibria and restrict myself to subgame perfection. The focus on public-perfect equilibria means I do not consider any private strategies, i.e., strategies in which actions are taken conditional on private information (Mailath and Samuelson, 2006). This focus is justified in Appendix 3.8.3 in the proof of Lemma 2.

Without repeated interaction, the certifier has a dominant incentive to cheat. This incentive comes from two sources: First, verifying the true quality $v = 1$ would incur monitoring cost c_C. By cheating, the certifier saves these costs, a moral hazard problem. Second, cheating allows the certifier to extract a higher surplus from the supplier since the supplier saves the cost c_S, a form of corruption. Thus, in a one-shot game, a certifier can only lose from truthful certification:

Lemma 1. *No High-Quality Equilibrium exists in the one-period game.*

Proof. See Appendix 3.8.3.1. □

This result easily extends to any finite horizon. At the infinite horizon, the buyers can play the following grim trigger strategy: If there were only high signals in the past, they buy the product at any price $p_{\tilde{v}} \leq U_{\tilde{v}}$ for a product of quality label $\tilde{v}$. Otherwise, their willingness to pay is zero. This grim trigger strategy allows the buyers to enforce a High-Quality Equilibrium if the certifier's discount factor is significant enough:

Lemma 2. *If no accreditation is available, the buyers enforce a High-Quality Equilibrium if and only if*

$$\delta \geq \underline{\delta}_{NA} \equiv \frac{1}{\lambda(\frac{U_1}{\bar{c}} - 1) + 1} \tag{3.3}$$

where $\bar{c} = c_S + c_C$.

The discount factor threshold $\underline{\delta}_{NA}$ from Lemma 2 is the incentive compatibility threshold of the certifier without accreditation.

Proof. See Appendix 3.8.3.2. □

The interpretation of Lemma 2 is as follows: If the discount factor δ is lower than the incentive compatibility threshold $\underline{\delta}_{NA}$, the certifier is too "impatient" or uncertain about the future and prefers to obtain a higher profit today by cheating, even though this might lower future profit if they are detected. The incentive compatibility threshold decreases in λ, the probability of a low signal if the certifier cheats. In other words, if the signal quality increases, a lower level of patience or certainty of future profits is sufficient to prevent cheating. Thus, increasing public monitoring could facilitate the achievement of the social optimum.

$\underline{\delta}_{NA}$ also decreases the buyers' willingness to pay for high quality U_1. However, it is increasing in the certifier's and the supplier's cost $\bar{c}$ of providing true quality $v = 1$. This result implies that the higher the net value of true quality certification, the easier it can be achieved.

Illustration. In the numerical example, ($U_v = v$, $c_S = c_C = \frac{1}{4}v^2$), Lemma 2 implies that the buyers enforce a High-Quality Equilibrium if and only if $\delta \geq \frac{1}{\lambda+1}$. □

The following sections extend the model to see whether the buyers can enforce a High-Quality Equilibrium more often if the certifier can get accredited. I also analyze under which type of accreditation body welfare is maximized and whether the state should favor monopoly or competition.

3.4 Model: Accreditation and sophisticated buyers

This section extends the baseline model to account for the possibility of accreditation, assuming that buyers are sophisticated. *Sophisticated* means that they can judge the accreditation quality. Accreditation forces the certifier to have a certain minimum level of rigor, i.e., to invest at least a fraction of their monitoring cost. This investment lowers the profit the certifier gains from cheating, such that there are cases in which the buyers can only enforce the High-Quality Equilibrium under accreditation.

3.4.1 Monopoly accreditation

As in the baseline model, time is discrete, and the horizon is infinite. There are four players now.

An additional player enters the game: the accreditation body. The accreditation body faces no competition, is long-lived, and is risk neutral. I will first derive a general result and then distinguish three cases: The accreditation body can be (1) a national accreditation body that maximizes total national welfare (or is regulated to do so), (2) a foreign accreditation body that maximizes total welfare of their foreign country, or (3) a profit-maximizing accreditation body. The other players are the same as in the baseline model.

The action sets are extended in the following way: The accreditation body offers an accreditation contract (β, g) where $g \in \mathbb{R}_+$ is the accreditation fee and $\beta \in [0,1]$ is the fraction of the certifier's monitoring cost c_C whose investment the accreditation body ensures. This specification reflects that accreditation bodies only inspect a sample of certifiers' activities and can thus only ensure competent certification for a fraction of the certifier's activity. I call β the accreditation quality since the effect of accreditation is increasing in β, as we will see. The certifier chooses whether to get accredited or not. The rest of the certifier's action set stays the same. The action sets of the supplier and the buyers stay the same.

The players' information stays the same, except that they have some additional information: All players observe the type of the accreditation body, the accreditation contract (β, g), and whether the certifier gets accredited or not. Throughout this section, this also holds for the buyers. I call the buyers *sophisticated buyers* whenever this information is relevant.

Definition 2. *A "sophisticated buyer" knows the game, the type of the accreditation body, the accreditation contract* (β, g)*, and whether the certifier gets accredited or not.*

The accreditation body has complete information.

The timing is as follows: Before the repeated interaction starts, the accreditation body chooses the accreditation fee g and the quality of accreditation β. Then, each stage game only changes in stage 1:

Stage 1: The certifier chooses to get accredited or not and offers a certification contract (v, f).

Stage 2: The supplier accepts or rejects the certification contract and offers to sell the product at price $p_{\tilde{v}}$.

Stage 3: The buyers accept or reject.

The payoffs of each stage game are as follows. The surplus of the buyers S_B and the surplus of the supplier S_S from trading a product of quality v at a price $p_{\tilde{v}}$ stay the same as in the baseline model. The certifier's surplus is:

$$S_C = \begin{cases} f - (vc_C) & \text{if they are not accredited} \\ f - \beta c_C - g - v(1-\beta)c_C & \text{if they are accredited} \end{cases}$$

The interpretation follows: Once the certifier gets accredited, they must invest βc_C. The decision to monitor high quality $v = 1$ only determines whether they incur the additional monitoring cost $(1-\beta)c_C$.

The accreditation body incurs a cost $\left(\frac{(\beta c_C)^2}{2} + K + t\right)$, where $t \in \mathbb{R}_+$ are the transportation cost and $K \in \mathbb{R}_+$ are the fixed establishment of the cost for each accreditation body.[27] $\frac{(\beta c_C)^2}{2}$ are the actual monitoring cost of accreditation. They are an increasing, convex function of the quality of accreditation β.[28] The surplus of the accreditation body, S_A, is then:

$$S_A = g - \left(\frac{(\beta c_C)^2}{2} + K + t\right)$$

The long-time payoffs of the accreditation body, the certifier, and the buyers are the discounted sums of the expectation of stage game payoffs for infinitely many stage games. Their common discount factor is δ. I assume the same tiebreaking rule as in the baseline model.

[27]The role of these variable will be clarified in sections 3.4.1.1 and 3.4.1.2. Since I have only one certifier per country in the model, the fact that K is a fixed cost does not play a role. It will, however, play a role when I introduce international trade of accreditation in section 3.4.1.2.

[28]The convexity of the monitoring cost of accreditation is justified by the fact that βc_C captures the accreditation quality. While the cost of checking a small fraction of the certifier's cost c_C, such as basic infrastructure and qualifications, the cost increase with a rising slope when it comes to checking not easily observable effort cost such as rigorousness in audits.

3.4.1.0.1 Social optimum Assume a utilitarian social planner determines all actions played. In each stage game, total welfare $W = S_B + S_S + S_C + S_A$ and thus

$$W = \begin{cases} U_1 - (\bar{c} + K) > 0 & \text{if v=1 \& the certifier is not accredited} \\ -K & \text{if v=0 \& the certifier is not accredited} \\ U_1 - (\bar{c} + \frac{(\beta c_C)^2}{2} + K + t) & \text{if v=1 \& the certifier is accredited} \\ -(\frac{(\beta c_C)^2}{2} + K + t) & \text{if v=0 \& the certifier is accredited} \end{cases} \tag{3.4}$$

where $\bar{c} = c_S + c_C$. The entrance of an accreditation body creates an establishment cost of K even if the certifier does not demand the body's services. I make the following assumption:

Assumption 1. *Any High-Quality Equilibrium yields strictly positive social welfare:*

$$U_1 - \left(\bar{c} + \frac{(\beta c_C)^2}{2} + K + t\right) > 0 \quad \forall \beta \in [0, 1]$$

That suggests that the entrance of the accreditation body improves welfare if and only if buyers cannot achieve a High-Quality Equilibrium without accreditation but can do so with accreditation. I will specify this result further below.

Illustration. In my numerical example, the total welfare under the presence of a single accreditation body is

$$W = \begin{cases} \frac{1}{2} - K & \text{if } v = 1 \text{ and the certifier is not accredited} \\ -K & \text{if } v = 0 \text{ and the certifier is not accredited} \\ \frac{1}{2} - \left(\frac{\beta^2}{32} + K + t\right) & \text{if } v = 1 \text{ and the certifier is accredited} \\ -\left(\frac{\beta^2}{32} + K + t\right) & \text{if } v = 0 \text{ and the certifier is accredited} \end{cases} \tag{3.5}$$

Assumption 1 implies positive welfare in a High-Quality Equilibrium even for $\beta = 1$. In the numerical example, Assumption 1 can thus only hold if:

$$(K + t) < \frac{15}{32} \tag{3.6}$$

□

3.4.1.0.2 Equilibrium Throughout the analysis, I focus on the conditions under which the buyers can enforce a High-Quality Equilibrium by playing a grim trigger strategy since this is the best possible strategy for the buyers, as shown in the baseline model. If the certifier never obtains accreditation, it is as if it is unavailable. Then, the buyers can enforce a High-Quality Equilibrium if and only if $\delta \geq \underline{\delta}_{NA}$, as shown in Lemma 2. If the certifier obtains accreditation, the incentive compatibility threshold changes:

Lemma 3. *If the certifier obtains accreditation in each period, the buyers can enforce a High-Quality Equilibrium if and only if*

$$\delta \geq \underline{\delta}_A \equiv \frac{1}{\lambda\left(\frac{U_1-\bar{c}-g}{\bar{c}-\beta c_C}\right)+1} \tag{3.7}$$

$\underline{\delta}_A$ is the incentive compatibility constraint when the certifier gets accredited in each period.

Proof. See Appendix 3.8.4.1. □

Recall that the buyers are sophisticated: they know the game, observe the accreditation contract (β, g), and whether the certifier obtains accreditation. This assumption implies that they know when truthful monitoring is incentive compatible. They can play a grim trigger strategy that forces the certifier to take the accreditation decision, which leads to a High-Quality Equilibrium if it exists:

Proposition 1. *If accreditation is available, sophisticated buyers enforce a High-Quality Equilibrium if and only if*

$$\delta \geq \min\{\underline{\delta}_A, \underline{\delta}_{NA}\} \tag{3.8}$$

Proof. See Appendix 3.8.4.2. □

Since accreditation is costly, it is only welfare improving if it enables a High-Quality Equilibrium that was not possible before:

Corollary 1. *Given sophisticated buyers, the entrance of an accreditation body improves welfare if and only if* $\underline{\delta}_{NA} > \delta \geq \underline{\delta}_A$.

A necessary condition for that is $\underline{\delta}_{NA} > \underline{\delta}_A$. This is equivalent to

$$g < \frac{\beta c_C(U_1 - \bar{c})}{\bar{c}} \tag{3.9}$$

This corollary shows that accreditation can only be welfare-improving if two conditions hold: The quality of accreditation β must be high enough. The accreditation fee g must be low enough. Thus, the choice of these variables is crucial. It is likely to vary with the objective of the accreditation body. In the following, I analyze the choices of different accreditation bodies and compare the effects.

Illustration. In my numerical example, $\underline{\delta}_A$, the threshold for enforceability of a High-Quality Equilibrium from Lemma 3 becomes the following:

$$\underline{\delta}_A = \frac{1}{\lambda\left(\frac{1-2g}{1-\beta/2}\right) + 1} \tag{3.10}$$

Then, $\underline{\delta}_{NA} > \underline{\delta}_A$ requires $g < \frac{\beta}{4}$ □

3.4.1.1 National welfare maximizing accreditation body

In this section, I assume that the only available accreditation body is a national accreditation body (denoted by subscript n). I assume that a national accreditation body offers an accreditation contract (β_n, g_n), which maximizes national welfare, where g_n is the accreditation fee and β_n is the accreditation quality. Accreditation under a national accreditation body does not incur transportation costs, so $t = 0$. However, the economy bears the cost $K \in \mathbb{R}_+$ of establishing the national accreditation body. Consequently, the accreditation contract (β_n, g_n) is the solution to the optimization problem defined in the following.

Definition 3. *The optimization problem of a welfare-maximizing accreditation body is:*

$$\max_{\beta,g} U - \bar{c} - \frac{(\beta c_C)^2}{2} - K$$

such that the Break Even constraint of the accreditation body is satisfied

$$g - \frac{(\beta c_C)^2}{2} - K \geq 0$$

and such that the Incentive Constraint holds:

$$\delta \geq \frac{\bar{c} - \beta c_C}{\lambda(U - \bar{c} - g) + (\bar{c} - \beta c_C)} \iff g \leq U - \bar{c} - \frac{1-\delta}{\delta\lambda}(\bar{c} - \beta c_C)$$

Remark. For simplicity, I assume that all players reside in the same economy, i.e., contribute to national welfare. I will discuss this assumption and possible extensions in the conclusion.

The solution to the problem in Definition 3 is $(\beta_n, g_n) = (K, 0)$, whenever the accreditation body cannot achieve a welfare-improving contract (see Appendix 3.8.4.3). The condition and solution for the case where it can improve welfare are as follows:

Proposition 2. *Under sophisticated buyers, the entrance of the national accreditation body improves welfare, if and only if (A1)*

$$\delta < \underline{\delta}_{NA} = \frac{1}{\lambda\left(\frac{U_1}{\bar{c}} - 1\right) + 1} \tag{3.11}$$

and (A2)

$$\delta \geq min\left\{\frac{1}{\lambda c_C + 1}; \frac{1}{\lambda \frac{1}{c_S}\left(U_1 - \bar{c} - \frac{c_C^2}{2} - K\right) + 1}\right\} \tag{3.12}$$

The accreditation contract is then (β_n, g_n) *where*

$$\beta_n = \frac{1-\delta}{\delta\lambda c_C} - \frac{1}{c_C}\sqrt{2\left(U_1 - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c}\right) + \left(\frac{1-\delta}{\delta\lambda}\right)^2} \tag{3.13}$$

and

$$g_n = \frac{(\beta_n c_C)^2}{2} + K \tag{3.14}$$

Proof. See Appendix 3.8.4.3. □

This proposition covers a particular case of Corollary 1. It is the best case of the corollary since it shows under which conditions a welfare-maximizing accreditation body can provide the conditions for a High-Quality Equilibrium when buyers cannot achieve it without accreditation. Assumption (A1) is simply the condition that a High-Quality Equilibrium cannot be achieved without accreditation. Assumption (A2) is needed so that an accreditation body can provide the conditions for a High-Quality Equilibrium. Note that a necessary condition for the combination of (A1) and (A2) is:

$$\frac{U_1}{\bar{c}} - 1 < \max\{c_C; \frac{1}{c_S}(U_1 - \bar{c} - \frac{c_C^2}{2} - K)\} \tag{3.15}$$

The interpretation is as follows: The establishment of the accreditation body can only be welfare improving if either the certifier's monitoring cost c_C are high enough, or if otherwise the buyers' valuation U_1 is high enough, but if the cost of the supplier c_S and the establishment cost of the accreditation body K are low enough. If the monitoring cost c_C is high, the cost saved by cheating would be high without accreditation but are much smaller under accreditation due to the sunk fraction $\beta_n c_C$. If the certifier's monitoring cost c_C is low, the effect of accreditation can still be sufficiently high if its establishment is not too costly and its net benefit high enough.

Illustration. In my numerical example, the necessary condition (3.15) translates into: $1 < \max\{\frac{1}{4}; \frac{15}{8} - 4K)\} \iff K < \frac{7}{32}$. Note that this condition also implies that welfare under a High-Quality Equilibrium and a national accreditation body is strictly positive (compare condition (3.6) in the Illustration of Assumption 1). To make the illustration more concrete, I assume that $\lambda = \frac{1}{10}$. That means I assume one out of ten fraudulent certificates is detected. Then, assumption (A1) in Proposition 2 ($\delta < \underline{\delta}_{NA}$) implies that $\delta < \frac{10}{11}$. Assumption (A2) implies that $\delta \geq \frac{80}{95-32K}$. Assume $K = \frac{1}{8}$ (such that condition (3.15) is satisfied, then (A1) and (A2) require that $\delta \in [\frac{80}{91}, \frac{10}{11})]$. $\delta = \frac{19}{22}$ would satisfy this requirement, for example. The lower K, the larger the interval of discount factors under which accreditation is welfare improving. Given the above assumptions, the optimal quality of accreditation $\beta_n \approx 0.781$, according to Proposition 2. The welfare gain through accreditation is then $\frac{1}{2} - \frac{\beta_n^2}{2} - \frac{1}{8} \approx 0.356$, since welfare would be zero without accreditation, since $\delta < \underline{\delta}_{NA}$. □

The illustration shows that the margin at which accreditation improves welfare might be smaller than expected. However, rigorous calibration and empirical testing of the model would be necessary to obtain a precise idea. Moreover, the model so far does not take into account scale effects. Scale effects will lower the participation of each accreditation fee to the overall establishment cost of K, and, as noted above, lower accreditation fees will make the case of welfare-improving accreditation more frequent.

Note that the optimal quality of accreditation β_n is decreasing in the buyers' valuation of high quality U_1, increasing in the establishment cost K as well as the supplier's and the certifier's cost of high quality c_S and c_C (and thus also $\bar{c}$). The larger the buyers' valuation of high quality and the lower the cost of high quality, the closer we are to the case where the High-Quality Equilibrium is enforced even without accreditation. Proposition 2 further implies:

Corollary 2. *If the entrance of a national accreditation body improves welfare, the optimal quality of accreditation β_n decreases in the level of public monitoring λ and the discount factor δ.*

Proof. See Appendix 3.8.4.4 □

The interpretation is as follows: The level of public monitoring and accreditation are imperfect substitutes. If λ goes up, i.e., if a higher fraction of cheating certifiers are detected, the threat of punishment for cheating increases. Hence, it relaxes the Incentive Constraint and, thus, the required accreditation quality to deter cheating. The same holds for the discount factor: A very long-lived or more patient certifier will disvalue the long-term threat of punishment more than a relatively short-lived or impatient certifier. A welfare-maximizing accreditation body can anticipate that and decrease the quality of accreditation β_n to save cost while still achieving the High-Quality Equilibrium.

3.4.1.2 Foreign welfare maximizing accreditation body

In this section, I assume that the entrance of a national accreditation body would be welfare-improving according to Proposition 2, but that the only available accreditation body is a foreign accreditation body. A foreign accreditation body is the national accreditation body of another country. The case of a foreign accreditation body is important since many small and developing countries do not have accreditation bodies but instead rely on foreign bodies.[29] This section will show under which conditions this is well-advised.

I make the following assumption:

Assumption 2. *If a foreign accreditation body is the only available accreditation body in a given country, they offer an accreditation contract*

$$(\beta_f, g_f) = \left(\beta_f, \frac{(\beta_f c_C)^2}{2} + t\right)$$

[29]The Kenyan certifier AfriCert, for example, is accredited by the German accreditation body DAkkS to certify farms according to the GLOBALG.A.P. standard, as Figure 3.0b illustrates. AfriCert has become one of the dominant players in agricultural certification in Sub-Sahara Africa (AfriCert, 2017; Fletcher, 2011). The German development agency GIZ supported their establishment, arguing that local certification would be more affordable and accessible for small-scale farmers than certification by international companies (Fletcher, 2011; Pannhausen and Untied, 2010). In the first years of its establishment and accreditation, AfriCert was not financially viable, but in 2012, it provided its investors with an internal rate of return of more than 30% (Fletcher, 2011).

where β_f is the solution to the welfare maximization problem from Definition 3 under the parameters $U_v = U_{vf} \in \mathbb{R}_+$, $c_C = c_{Cf} \in \mathbb{R}_+$, $c_S = c_{Sf} \in \mathbb{R}_+$, $\delta = \delta_f \in [0,1]$, $\lambda = \lambda_f \in [0,1]$ and $K = K_f$. $g_f \in \mathbb{R}_+$ is the accreditation fee and $t \in \mathbb{R}_+$ are the transportation cost.

This assumption implies that the foreign accreditation body sets the quality of accreditation β_f by solving the same optimization problem as the national accreditation body but based on the data of their country of origin. It further means that the foreign accreditation body keeps the accreditation quality β_f constant when providing accreditation abroad.[30] I make this assumption due to empirical indications that national accreditation bodies typically follow the same regulations when accrediting abroad.[31]

Assumption 2 also implies that the foreign accreditation body still prices at cost when providing services cross-country by adapting the accreditation fee by accounting for c_C. Pricing at a cost even abroad is common practice among national accreditation bodies.[32]. However, I assume that the economy in which the foreign accreditation body resides already paid the fixed establishment cost of K. I do so because national accreditation bodies are usually established to break even without providing service abroad and typically do so only after years of experience at home. Providing services abroad incurs transportation costs of t instead.

I assume that the existence of the foreign accreditation body in its home country is welfare improving.[33] Then, Proposition 2 and Assumption 2 imply that

$$\beta_f = \frac{1-\delta_f}{\delta_f \lambda_f c_{Cf}} - \frac{1}{c_C f}\sqrt{2\left(U_{1f} - \bar{c}_f - \frac{1-\delta_f}{\delta_f \lambda_f}\bar{c}_f - K_f\right) + \left(\frac{1-\delta_f}{\delta_f \lambda_f}\right)^2} \tag{3.16}$$

The home country of the foreign accreditation body can be a more advanced economy (denoted by subscript a) or a more developing country (denoted by subscript d).[34] I make the following assumptions:

Assumption 3. .

[30] A corresponding case is national accreditation bodies created for the accreditation of certification of ISO standards, but later competing with the international accreditation body IOAS in the accreditation of certification of organic products.

[31] See, for example, the regulations of the Dutch Accreditation Council (RvA) (2013).

[32] This is, for example, the case for the German and the Dutch accreditation body Dutch Accreditation Council (RvA) (2013); Kirmes (2019)

[33] Otherwise, the existence of that body would not be socially desirable

[34] A proxy for being more advanced or more developing can be income per capita.

(1) $c_{Cd} > c_C > c_{Ca}$ *(2)* $c_{Sd} \geq cS \geq c_{Sa}$ *(3)* $U_{vd} \leq U_v \leq U_{va}$

(4) $\lambda_d \leq \lambda \leq \lambda_a$ *(5)* $\delta_d \leq \delta \leq \delta_a$ *(6)* $K_d \geq K \geq K_a$

This assumption should be understood as follows. Assumption 3 (1) means that the certifier's cost of effectively monitoring the supplier (c_C) is lower in more advanced economies and higher in more developing countries, for example, due to the use of technology such as satellite imagery, but also due to institutional embeddedness and infrastructure. Assumption 3 (2) implies that truthful implementation of high quality is less costly in a more advanced economy and costlier in more developing economies. I expect this to be the case since products in more advanced economies are already of higher quality on average. Also, many compulsory quality standards in advanced economies, such as CE Marking in the EU, are still voluntary in developing countries. Assumption 3 (3) is made since most international quality standards are more known in advanced economies than in developing economies (UNIDO et al., 2017). Hence, buyers' valuation for them is likely to be higher in advanced economies than in developing economies. Moreover, higher wealth in advanced economies likely leads to higher ability and, thus, higher willingness to pay for high quality. Assumption 3 (4) means that the level of public monitoring in developing countries is likely lower than in advanced economies. A lower level of monitoring implies that the probability of detecting fraudulent certification is lower. There are many indications for that: developing countries are, on average, faced with lower freedom of the press, lower internet connectivity, lower resources for NGOs, state investigations, media, and research, higher levels of corruption, and lower levels of education than advanced economies.[35] Assumption 3 (5) says that certifiers and buyers are less patient or long-lived in developing countries than in more advanced economies. I assume that for two reasons: First, financial constraints typically reduce companies' and individuals' patience. Second, developing countries are, on average, faced with more political instability and a lower level of the rule of law than advanced economies (Institute for Economics and Peace, 2019; World Justice Project, 2019). Assumption 3 (6) means that establishing an accreditation body is more costly for developing than for more advanced economies. I assume that since establishing an accreditation body requires specialized expertise, more developing countries often have to fly in from abroad (UNIDO et al., 2017).

The following result is immediate:

[35] See Auriol and Picard (2006); Freedom House (2017a,b); Hanushek and Wößmann (2007); Olken and Pande (2012); Transparency International (2018)

Proposition 3. *Under the assumptions of Proposition 2, Assumption 2 and 3, the optimal quality of accreditation decreases with an economy's development:*

$$\beta_d \geq \beta \geq \beta_a \tag{3.17}$$

Where β is the optimal quality of accreditation in the reference country, β_d is the optimal quality of accreditation in a more developing economy, and β_a is the optimal quality of accreditation in a more advanced economy.

This proposition does not align with the widespread impression that accreditation bodies from developing countries are less rigorous. Indeed, the actual accreditation quality is likely weaker than the optimal level due to capacity constraints. Political economy and industrial policies also play an essential role and can lead to non-welfare maximizing accreditation regulation. Therefore, I will not continue the interpretation by distinguishing between a foreign accreditation body from a more developing or advanced economy. Instead, I will distinguish between a foreign accreditation body with a higher or a lower accreditation quality β_f. Nevertheless, the above analysis has illustrated that it is far from certain that accreditation bodies from more advanced economies have a higher quality of accreditation. Even though the precise determination of an optimal accreditation quality β seems far from reality, I expect that a benevolent regulator or accreditation body manager is aware of some of the issues I mentioned in the context of Assumption 3. Hence, it is not unreasonable to expect them to audit certifiers more carefully than accreditation bodies from advanced economies would do.

Replacing a national accreditation body with a foreign accreditation body affects the outcome. Recall that the quality of accreditation β_n is the minimum β for which the Incentive Constraint of the certifier is satisfied. Thus, the Incentive Constraint will no longer be satisfied under a foreign accreditation body with a lower accreditation quality. Suppose the quality of accreditation β_f of the foreign accreditation body is higher than that of the national body. In that case, the Incentive Constraint is still satisfied if the difference between β_n is small enough. If β_f is too high, the cost will exceed the benefit, and the Incentive Constraint will no longer be satisfied.

Proposition 4. *Under Assumption 2 and sophisticated buyers, it is welfare-improving to let certifiers obtain accreditation with a foreign accreditation body rather than establishing a national accreditation body if and only if*

(1) the quality of accreditation of the foreign accreditation body β_f is at least as high as β_n, the optimal quality of the national accreditation body,

(2) $\beta_f \leq \frac{1-\delta}{\delta\lambda c_C} + \frac{1}{c_C}\sqrt{2(U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} - K) + (\frac{1-\delta}{\delta\lambda})^2}$

(3) the establishment cost of a national accreditation body K is higher than the sum of the transport cost t and the additional variable accreditation cost $\frac{c_C^2}{2}(\beta_f^2 - \beta_n^2)$:

$$K \geq \frac{c_C^2}{2}(\beta_f^2 - \beta_n^2) + t \tag{3.18}$$

Proof. See Appendix 3.8.4.5 □

Thus, relying on a welfare-maximizing accreditation body of a similar economy (with a similar quality of accreditation) can be welfare-improving if the establishment cost of a national accreditation body K is high enough, transportation cost t low enough, and the quality of accreditation of the foreign body is not lower but not too high either. If the economy cannot find a foreign accreditation body satisfying these conditions, it should establish its own national accreditation body.

Proposition 4 thus explains why some small countries have cooperation agreements with other countries to use their accreditation services UNIDO et al. (2017). Other countries have formed regional accreditation groups, such as the Southern African Development Community Accreditation Service (SADCAS) and the CARICOM Regional Organisation for Standards and Quality in the Caribbean (CROSQ, 2014; UNIDO et al., 2017).

3.4.1.3 Profit maximizing accreditation body

I now assume that the only available accreditation body is profit maximizing. Such accreditation bodies exist but are usually not internationally recognized (International Accreditation Forum, 2018a).[36]

I assume the profit-maximizing accreditation body faces the same cost function as the national accreditation body but chooses the accreditation quality β_p to maximize profit. Given sophisticated buyers, the profit-maximizing accreditation body must still satisfy the Incentive Constraint. Otherwise, the buyers would not accept to buy a certified product

[36] Heras-Saizarbitoria and Boiral (2019) describe the importance of such accreditation bodies in China.

($\tilde{v} = 1$) at a positive price $p_{\tilde{v}} > 0$ if the certifier is accredited, such that the certifier would not demand accreditation.[37]

Thus, the profit-maximizing accreditation body's problem is:

$$\max_{\beta,g} g - \frac{(\beta c_C)^2}{2} - K \tag{3.19}$$

s.t. the Individual Rationality constraint of the certifier

$$U - \bar{c} - g \geq 0 \tag{3.20}$$

and s.t. the Incentive Constraint holds:

$$g \leq U - \bar{c} - \frac{1-\delta}{\delta\lambda}(\bar{c} - \beta_p c_C) \tag{3.21}$$

Note that under the Incentive Constraint, the Individual Rationality constraint of the certifier is slack. The solution to the problem is:

$$\beta_p = \min\left\{\frac{1-\delta}{\delta\lambda c_C}; 1\right\} \tag{3.22}$$

$$g_p = \min\left\{U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} + \left(\frac{1-\delta}{\delta\lambda}\right)^2; U - \bar{c} - \frac{1-\delta}{\delta\lambda}c_S\right\} \tag{3.23}$$

where the subscript p denotes the solution as "profit-maximizing".

The profit maximizing accreditation body offers this contract (β_p, g_p) if and only if $g_p - \frac{(\beta_p c_C)^2}{2} - K \geq 0$. This is the case whenever assumption (A2) of Proposition 2 is given.[38] This directly implies the following result:

Proposition 5. *Under sophisticated buyers, the entrance of a profit-maximizing accreditation body improves welfare, whenever the entrance of a national accreditation body would have been welfare improving (according to Proposition 2), but does not occur.*

Note that the profit-maximizing accreditation body always offers a quality of accreditation β_p that is at least as high as the quality β_n, which a national accreditation body would offer under the assumptions of Proposition 2. Whenever $\beta_n < \beta_p$, the cost of accreditation

[37]See the proof of Proposition 1.

[38]Compare conditions 3.33 and 3.39 in the proof of Proposition 2 in Appendix 3.8.4.3.

under the profit-maximizing accreditation body is higher. In these cases, welfare under the profit-maximizing accreditation body is lower than under a national body.

Illustration. Under the assumptions I made in my numerical example, entering a national accreditation body would improve welfare and yield a total welfare of $W \approx 0.356$. If there were no national accreditation body, but only a profit-maximizing one, total welfare would be lower: The profit maximizing accreditation body sets $\beta_p = \min\left\{\frac{1-\delta}{\delta\lambda c_C}; 1\right\} = \min\left\{\frac{120}{19}; 1\right\} = 1$. Hence, total welfare is $\frac{11}{32} \approx 0.344$. The welfare loss is 0.012.

As an extension of the model, I could think about the possibility that a profit-maximizing accreditation body is more efficient so that, for example, their establishment cost of K is lower than those of the national body. Then using a profit-maximizing instead of a welfare-maximizing accreditation body might increase welfare. □

3.4.2 Accreditation competition

So far, I assumed that only one accreditation body could offer its services in the model's economy. In this section, I analyze how the above results would change if there were competition among different accreditation bodies.

3.4.2.1 Foreign versus national welfare maximizing

Assume we are in the case that the entrance of a national accreditation body is welfare improving according to Proposition 2.[39] Recall that sophisticated buyers will not accept accreditation that does not induce incentive compatibility of a High-Quality Equilibrium. If the foreign accreditation quality β_f is instead such that the buyers can enforce a High-Quality Equilibrium, they accept both foreign and national accreditation. This result corresponds to conditions (1) and (2) in Proposition 4. The certifier then chooses the accreditation body with the lower fee. This is the foreign accreditation body if and only if

$$\frac{\beta_n^2 c_C^2}{2} + K \geq \frac{\beta_f^2 c_C^2}{2} + t \tag{3.24}$$

This condition corresponds to condition (3) in Proposition 4. Under these conditions, there is no demand for accreditation from the national body, and the establishment costs K of the

[39]Otherwise, no accreditation body can improve welfare since the national accreditation body sets the optimal quality of accreditation.

national body are sunk costs. Thus, welfare would be higher if the national body did not exist in the first place, as in Proposition 4.

In any other case, there is no demand for foreign accreditation. In all cases, society is better off under monopoly accreditation, in some cases (Proposition 4) with a foreign accreditation body, and others with their own.

3.4.2.2 Profit maximizing versus welfare maximizing

When competing with a national accreditation body, the profit-maximizing body must also price at cost. Thus, the profit-maximizing body will set the same accreditation quality as the national body to satisfy the Incentive Constraint. The national and the profit-maximizing accreditation bodies' accreditation contracts are then identical. The buyers and the certifier are indifferent. Sophisticated buyers enforce a High-Quality Equilibrium if and only if they would have done so under monopoly accreditation.

However, total welfare is lower since society spends double the cost of establishing accreditation bodies. Thus, society is better off with monopoly accreditation under a national or foreign accreditation body, according to Proposition 4.

If welfare-maximizing accreditation bodies did not exist, total welfare might be higher under the competition of two profit-maximizing accreditation bodies than under a monopoly profit-maximizing accreditation body.

3.5 Model: Accreditation and unsophisticated buyers

So far, I assumed that the buyers were sophisticated according to Definition 2. In most empirical cases, this assumption will not hold. It might be that some very institutionalized buyers spend many resources to understand the certification and accreditation market. They could then get an idea of the accreditation quality from different accreditation bodies.[40] However, many such professional buyers do not have sufficient knowledge to distinguish between effective and ineffective certification and accreditation (Heras-Saizarbitoria and Boiral, 2019). Therefore, I will now relax the assumption of sophisticated buyers. Instead, I assume that the buyers are unsophisticated according to the following definition:

[40] This is a more plausible assumption if I extend the model such that the buyers are long-lived downstream firms that have to ensure a High-Quality Equilibrium to keep their reputation towards their consumers.

Definition 4. *Unsophisticated buyers know the game, except for the certifier's and the supplier's costs. They know the type of the accreditation body and whether the certifier gets accredited or not but do not observe the accreditation contract* (β, g).

With this definition, the analysis still focuses on a case in which the buyers have a good level of knowledge about certification, which is unlikely for final consumers, but more reasonable for institutionalized professional buyers.

The rest of the game stays the same as in the above sections. In the following, I will analyze how the above results change under unsophisticated buyers and how the state should intervene. The first result says that with unsophisticated buyers, Proposition 1 does no longer hold:

Proposition 6. *Even if accreditation is available, unsophisticated buyers enforce a High-Quality Equilibrium if and only if* $\delta \geq \underline{\delta}_{NA}$.

Proof. See Appendix 3.8.5.1. □

This proposition directly implies the following:

Corollary 3. *If buyers are unsophisticated and there is no regulation of the accreditation market, there is no demand for accreditation.*

Note that this holds even when demand for accreditation by a national accreditation body would improve welfare, in which case sophisticated buyers would demand accreditation. With unsophisticated buyers, a national accreditation body or the state establishing it has to make an effort to inform buyers about the effect of accreditation. In other words, they must educate unsophisticated buyers to transform them into sophisticated buyers. Such education does not necessarily need to be the state's responsibility. Industry associations, media, and accreditation bodies could also do it. Another option would be to make accreditation compulsory.

3.6 Model: Accreditation bias

So far, I have assumed that buyers are either sophisticated or unsophisticated (Definitions 2 and 4). Another critical case is when buyers cannot judge if accreditation is incentive compatible in a specific case but prefer accreditation. This preference can, for example, be due to a belief created by marketing or because governments in other countries have made accreditation compulsory. To be precise, I define such a bias as follows:

Definition 5. *"Buyers with accreditation bias" know the game, except the certifier's and the supplier's cost. They know the type of the accreditation body and whether the certifier gets accredited but do not observe the accreditation contract* (β, g)*. They believe that accreditation is always welfare improving, i.e., that* $\underline{\delta}_{NA} > \delta \geq \underline{\delta}_A$

The rest of the game stays the same as in the above sections. In the following, I analyze how the above results change under buyers with accreditation bias and how the state should intervene.

3.6.1 Monopoly accreditation

Lemma 4. *Buyers with accreditation bias enforce a High-Quality Equilibrium if and only if* $\delta \geq \underline{\delta}_A$.

Proof. See Appendix 3.8.6.1. □

Lemma 4 means that Proposition 1 does no longer hold. It further implies that the entrance of a national accreditation body is welfare-improving whenever $\delta \geq \underline{\delta}_A$. That means that Proposition 2 holds even if assumption (A1) in the proposition is not satisfied: Buyers with accreditation bias have no willingness to pay for certified products when the certifier is not accredited. Thus, total welfare is zero if no accreditation body is available. Hence, the entrance of the national accreditation body improves welfare whenever it induces the certifier not to cheat.

Proposition 4 still trivially holds: When transportation costs for a foreign accreditation body are much lower than the establishment cost of a national body, and if the foreign body is as effective as the national one, it is still better to invite their services instead of establishing a national body.

The outcome under a profit-maximizing accreditation body changes when buyers have accreditation bias:

Proposition 7. *Under a monopoly profit-maximizing accreditation body, buyers with accredi-*

tation bias cannot enforce a High-Quality Equilibrium whenever

$$\delta < \min\{\underline{\delta}_{NA}; \frac{1}{\lambda 2(1-c_S)+1}\}$$

Proof. See Appendix 3.8.6.2. □

Note that Proposition 7 only establishes a sufficient condition for the non-existence of a High Equilibrium. In contrast, I could not derive a sufficient condition for the existence as in the previous sections. Proposition 7 implies that Proposition 5 does no longer apply:

Corollary 4. *With buyers with accreditation bias, the entrance of a profit-maximizing accreditation body does not always improve welfare when a national accreditation body would have improved welfare but did not enter.*

Consequently, a profit-maximizing accreditation is no longer a good replacement for a national accreditation body since the buyers' naive beliefs allow the profit-maximizing body to extract profit with zero-quality accreditation.

Illustration. Under the assumptions I made in my numerical example so far, the condition in Proposition 7 would come down to: $\delta < \min\{\frac{10}{11}; \frac{20}{23}\} = \frac{20}{23}$. Thus, the case under accreditation bias, where the entrance of a national body would improve welfare, but the entrance of a profit-maximizing body would not, is $\delta \in [\frac{80}{91}; \frac{20}{23}]$. The discount factor employed in the illustration so far ($\frac{19}{22}$) would satisfy this condition. □

3.6.2 Accreditation competition

Buyers with accreditation bias cannot distinguish accreditation bodies with a quality β which ensures a High-Quality Equilibrium from those with a lower quality. Thus, competition between a national accreditation body and a foreign accreditation body with lower quality and fees would prevent a High-Quality Equilibrium. It would then be welfare improving to prevent such accreditation bodies from offering their service in one's country.

If they compete, a profit-maximizing body can no longer set a higher accreditation fee than the national body. They solve the following problem:

$$\max_{\beta,g} g - \frac{(\beta c_C)^2}{2} \tag{3.25}$$

Such that the profit-maximizing accreditation body can compete with the national body:

$$g_p \leq g_n \tag{3.26}$$

Thus, the accreditation fee of the profit-maximizer is the same as the national body's fee g_n (defined in Proposition 2). Their accreditation quality β is zero. Their long-term profit is then

$$\sum_{t=0}^{\infty} \delta^t (1-\lambda)^t \left(g - \frac{(\beta c_C)^2}{2} \right) = \frac{g_n}{1-\delta(1-\lambda)} > 0 \tag{3.27}$$

Consequently, buyers cannot enforce the High-Quality Equilibrium whenever they can without accreditation, that is, if $\delta < \underline{\delta}_{NA}$. If that does not hold, the regulator should forbid competition with a profit-maximizing accreditation body or inform buyers sufficiently to eliminate their accreditation bias, at least for profit-maximizing accreditation bodies.

3.7 Discussion and Conclusion

This paper studied the effect of accreditation on the credibility of certification. I derived the conditions under which accreditation ensures certifiers' credibility by deterring fraud. The distinction between sophisticated, unsophisticated buyers and buyers with accreditation bias emphasized one fundamental condition: Buyers must be sufficiently informed to demand accreditation from certifiers, or accreditation must be compulsory.[41] That can explain why voluntary Business-to-Business and Business-to-Consumer standards are usually regulated differently. Accreditation of certifiers of voluntary Business-to-Business standards like the ISO 9000 management system standards is not compulsory. Instead, firms learn about accreditation, for example, through industry associations and consultants (Heras-Saizarbitoria and Boiral, 2019). Owners of Business-to-Consumer standards usually make accreditation of certifiers compulsory. An example is the Marine Stewardship Council (MSC). Most consumers who buy products with the MSC label[42] do not know about accreditation. They typically do not even know who certifies the producer's compliance with the MSC standard. Imagine consumers learned about fraudulent MSC certification. Then, MSC would lose its reputation even if only one certifier cheated. By forcing the certifier to obtain accreditation, the MSC may prevent this from happening.

To verify this mechanism, it would be helpful to extend my model to account for the role of

[41] See Corollary 3.

[42] See Figure 3.5 in Appendix 3.8.7.

standard owners. Moreover, adapting the model to allow for a mix of different buyer types would be useful. Especially for Business-to-Business standards, it is likely that some firms have enough insider knowledge to judge the quality of accreditation bodies while the majority have not. I could model this heterogeneity by allowing a fraction of sophisticated buyers among mostly unsophisticated buyers.

Such an extension would also allow better to evaluate the outcome under a profit-maximizing accreditation body. With sophisticated buyers, the model has shown that a profit-maximizer reduces fraud in the certification market whenever a welfare-maximizing accreditation body could do so but is not available.[43] If buyers naively trust in any accreditation, this result does not remain valid.[44] This can explain some of the certification fraud documented by Heras-Saizarbitoria and Boiral (2019) In future work, it would be interesting to analyze what fraction of sophisticated buyers is necessary to give accreditation bodies the right incentives. In any case, welfare is lower under a profit-maximizing accreditation body: If they deter the certifier from cheating, they do so with a higher accreditation quality than needed. Such excessive quality creates unnecessary social costs. The model thus explains that some countries have banned accreditation bodies not appointed by the state.[45] Where unappointed accreditation bodies are not banned, they usually operate as non-profits to be internally recognized (ISO, 2017). This situation may result from a mix of sophisticated buyers and buyers with accreditation bias: Sophisticated buyers might know that the presence of buyers with accreditation bias could incentivize for-profit bodies to cheat. This knowledge could decrease demand such that accreditation bodies prefer a non-profit status even when not state-regulated or altruistic.[46]

I could improve the model also by allowing for buyers with heterogeneous private valuations. In addition, I might consider buyers as retailers with long-term relations to final consumers. Such an assumption would require adapting the model so that the retailer gains some profit which they would lose if they lost their reputation. The retailer would then observe the stage game payoff and act similarly to the certifier: They would pay for certification and accreditation only when consumers were sophisticated enough to require them. Therefore, I expect the model's results will not change qualitatively. More fruitful might be an extension

[43]See Proposition 5.

[44]See Proposition 7.

[45]This is, for example, the case in the European Union (Reg. (EC) No 765/2008).

[46]Glaeser and Shleifer (2001) explain that even non-altruistic companies in a sector like accreditation may prefer to operate as non-profits since this gives them lower incentives to cheat and thus a better reputation among buyers.

of the model that includes imperfect competition of certifiers. The restriction to a single certifier and supplier does not account for the scale effects of accreditation, which might increase the effectiveness of accreditation.

An important result of this paper is that the optimal accreditation quality is, on average, higher in developing than advanced economies.[47] In other words, accreditation bodies in developing countries should exert more effort to monitor certifiers. However, establishing accreditation bodies in developing countries can be particularly challenging and costly (UNIDO et al., 2017). Therefore, my model suggests that developing and small economies should only establish their own national accreditation body if they cannot find a foreign accreditation body of sufficiently high quality and sufficiently low transportation cost.[48] This explains the creation of regional accreditation bodies in Southern Africa and the Caribic.[49]

The model's results on national and foreign accreditation bodies relied on the assumption that these bodies are welfare-maximizing. I assumed that they maximize the welfare of their own economy and that all players reside in their economy. In international trade, buyers often do not reside in the same country as the producer, the certifier, and the accreditation body.[50] Then buyers would not contribute to national welfare. This situation would not change the results of my model since I modeled the case of atomistic buyers.[51] Extending the model to buyers with market power would allow me to analyze further the incentives of welfare-maximizing accreditation bodies. National bodies might no longer want to provide sufficient accreditation quality. The same argument would apply if they could lower their quality when accrediting abroad.[52] The model could then be extended to account for the role of reputation building of accreditation bodies and the political economy of international peer reviews.[53]

My model also indicates that, even if accreditation bodies have the right incentives and buyers are sophisticated, accreditation cannot always prevent certification fraud.[54] This can be an

[47]See Proposition 3.

[48]See Proposition 4.

[49]These are the Southern African Development Community Accreditation Service (SADCAS) and the CARICOM Regional Organisation for Standards and Quality in the Caribbean (CROSQ, 2014; UNIDO et al., 2017).

[50]See for example Figure 3.0b in Appendix 3.8.2).

[51]Hence, the buyers made zero surplus.

[52]This would require relaxing Assumption 2.

[53]An interesting question in such a setup would be whether national accreditation bodies have the right incentives to monitor other accreditation bodies rigorously.

[54]See Proposition 2.

additional explanation for the fact that there are cases of fraudulent certifiers accredited by national bodies.

The econometric analysis of the effect of the international recognition of the Uruguayan accreditation body suggests that national accreditation bodies positively affect the demand and supply of certification. However, the result of the difference-in-differences analysis is only correct if no significant omitted factors changed the certification market in Uruguay compared to the control country Nicaragua between 2010 and 2016. To be sure of that, more qualitative and quantitative investigation is required.[55]

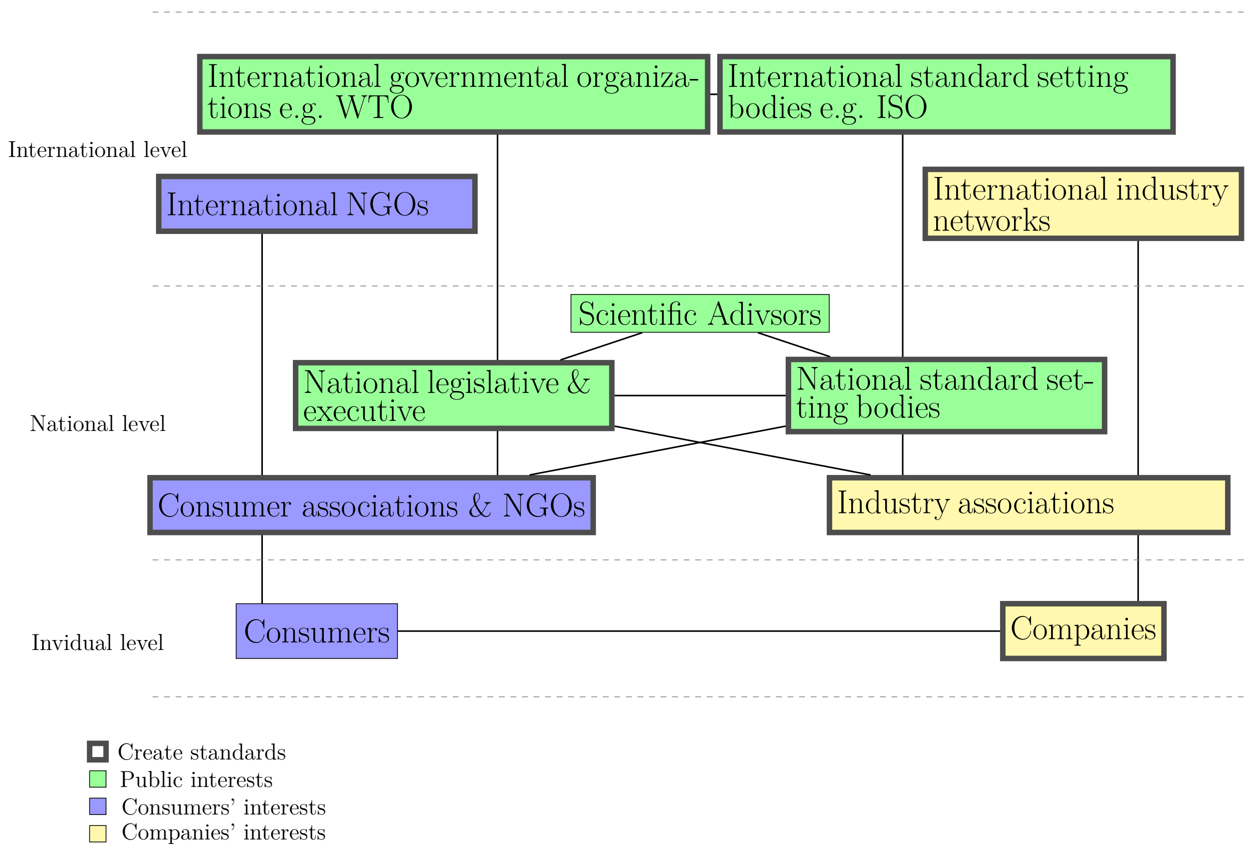
To conclude, this paper provides a solid starting point to understand the effects of accreditation and its regulation. It shows that the intervention of states or standard owners in the certification and accreditation market is often justified. However, the globalization of supply chains and the certification market can render a national focus on accreditation inefficient. More research is needed to find optimal ways of combining public and private as well as national and international endeavors in accreditation.

[55]It would be helpful to conduct further difference-in-differences analyses for other country pairs to verify the external validity of the result. Finding the correct country pairs and data is a challenge. Not many panel data on certification demand are conducted parallelly for different countries before and after a change in accreditation regulation. Moreover, the countries need to have a common pre-trend in certification. I found relevant data for El Salvador and Ecuador, but they did not have a common pre-trend. A systematic analysis would require collecting relevant data from accreditation bodies worldwide.

3.8 Appendix

3.8.1 Introduction

Figure 3.3: Public and private standard-setting processes



3.8.2 Context and motivating evidence

Figure 3.4: International and regional recognition of national accreditation bodies (Data: Websites of the IAF and the regional accreditation groups AFRAC, ARAC, EA, IAAC, PAC)

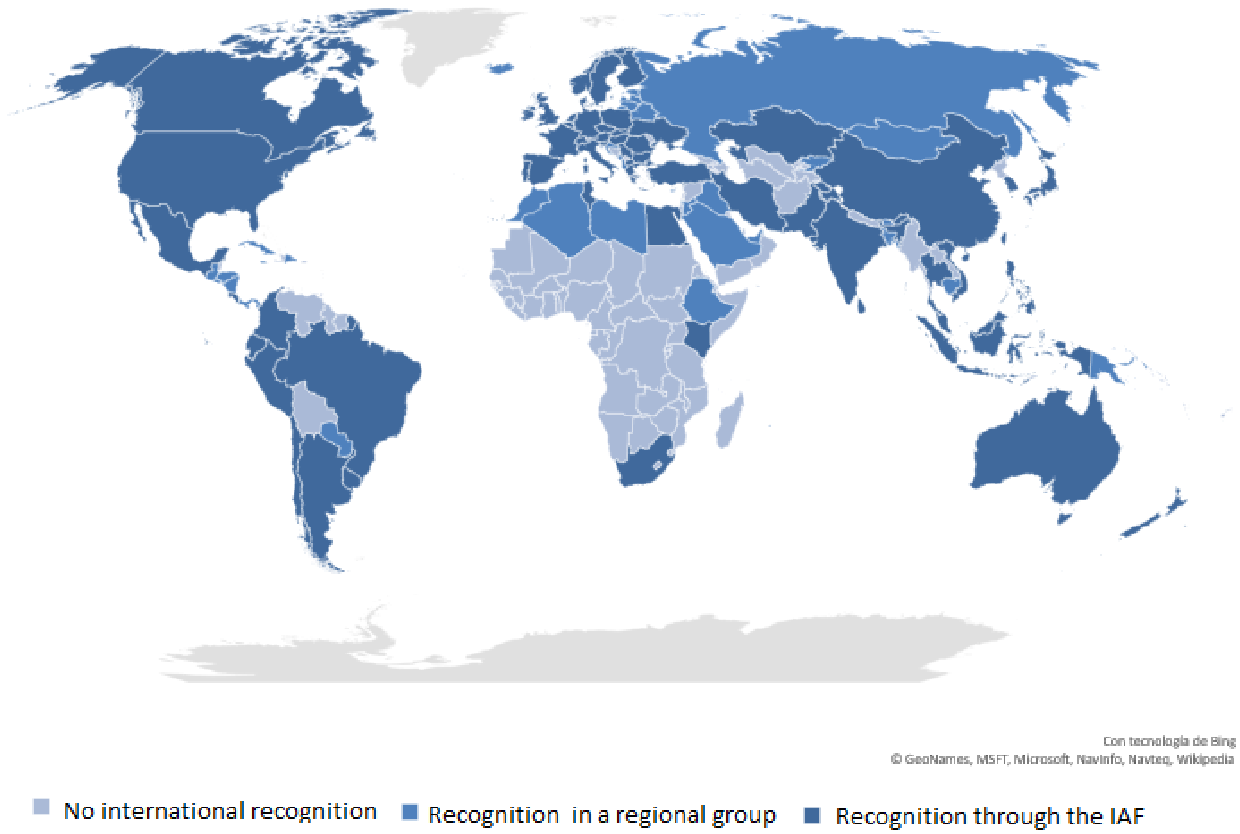


Table 3.2: Descriptive statistics of the treatment and control group

	Treatment group 2010		Difference in means *Treat.-Control*	
	Mean	SD	2010	2016/17
Dependent variable				
Internationally recognized certification	0.223	0.418	−0.035	0.175
Control variables				
% domestic shareholders	85.580	34.286	−0.944	−2.106
Years sector experience manager	24.554	13.358	2.645	3.692
	Treated		Control	
	2010	2017	2010	2016
Sample size	112	112	120	120

Table 3.3: Test of common trend

	(1)	(2)
Treated x (Year=2010)	0.0279	0.0305
	(0.35)	(0.38)
Year=2010	0.0213	0.0209
	(0.36)	(0.34)
Constant	0.191***	0.215
	(6.06)	(1.55)
Observations	340	340
Controls	No	Yes
Company FE	Yes	Yes
F-stat.	0.5	0.31

Notes: Estimates from regression of specification 2.3 in column (1), but with data from 2006 and 2010, using 2010 as the *Post*-treatment period, adding company controls in column (2) (percentage of domestic shareholders and the years of sector experience of the top manager). t statistics in parentheses, robust to heteroskedasticity. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.8.3 The baseline model

3.8.3.1 Proof of Lemma 1

I solve the game by backward induction:

No matter what the buyers play in stage 3, the supplier accepts a certification contract (v, f) in stage 2 if and only if their surplus is positive:

$$S_S \geq 0 \iff f \leq p_{\tilde{v}} - vc_S$$

In stage 1, the certifier offers $(0, p_{\tilde{v}})$ since $p_{\tilde{v}} - 0 - 0 > p_{\tilde{v}} - c_S - c_C$.

Thus, buyers cannot enforce true quality $v = 0$ and the social optimum in a one-shot game.

3.8.3.2 Proof of Lemma 2

Lemma 2 is proven in three steps by three additional Lemmas that are presented and proven below.

Lemma 5. *If $\delta \geq \underline{\delta}_{NA}$, the buyers can enforce a High-Quality Equilibrium by playing the* *following grim trigger strategy: If there were only high signals in the past, they buy the product at any price $p_{\tilde{v}} \leq U_{\tilde{v}}$ for a product of quality label $\tilde{v}$. Otherwise, their willingness to pay is zero.*

Proof. Let $\delta \geq \underline{\delta}_{NA}$. Assume the buyers play the grim trigger strategy described in Lemma 5. This assumption implies that their willingness to pay under certification ($\tilde{v} = 1$) equals their valuation U_1 for a product of true quality $v = 1$, whenever there were no low signals in the past. Otherwise, they do not trust a high-quality label $\tilde{v} = 1$.

Thus, if there was any low signal in the past, the seller could not make any positive profit, neither without certification ($\tilde{v} = 0$) nor from obtaining a quality label $\tilde{v} = 1$. The only certification contract they would accept is thus $(v, f) = (0, 0)$, where the true quality level v is zero, and they obtain certification for free. Hence, all players make zero profit.

If there were only high signals in the past, the stage game at period $t \geq 1$ is solved by backward induction as follows: In stage 3, the buyers are willing to pay any price $p_{\tilde{v}} \leq \tilde{v}$ for a product of quality label $\tilde{v}$. In Stage 2, if and only if $f \leq p_{\tilde{v}} - vc_S = U_1 - vc_S$, the seller accepts the certification contract and offers a product of quality label $\tilde{v} = 1$ for price $p_1 = U_1$. This contract gives them a payoff of $S_S \geq 0$. Otherwise, they can only offer a product of quality label $\tilde{v} = 0$ for price $p_0 = 0$. In Stage 1, the certifier offers a contract $(v, f) = (1, U_1 - c_S)$ if and only if there is no profitable one-shot deviation to the strategy of playing $(1, U_1 - c_S)$ each period.[56] The most profitable one-shot deviation of playing $(0, U_1)$ in period t is not profitable if and only if:

$$\sum_{t=0}^{\infty} \delta^t (U_1 - c_S - c_C) \geq U_1 + \delta(1 - \lambda) \sum_{t=0}^{\infty} \delta^t (U_1 - c_S - c_C) \tag{3.28}$$

This condition yields the condition given in 2.

The proof also shows why it is justified to limit attention to public strategies: The buyers have no private information and can hence only play a public strategy. Hence, the highest price $p_{\tilde{v}}$, which the buyers would accept to pay, depends only on public information. Consequently, the payoffs of the certifier and the supplier depend only on the history of public signals, such that their relevant behavior strategies should only depend on the public history so that their strategies are public. □

[56] See Mailath and Samuelson (2006) for the one-shot deviation principle

Lemma 6. *(a) The buyers can enforce a High-Quality Equilibrium only if $\delta \geq \underline{\delta}_{NA}$. (b) Only by playing a grim trigger strategy can the buyers enforce a High-Quality Equilibrium for any $\delta \geq \underline{\delta}_{NA}$*

Proof. Accepting prices $p_{\tilde{v}} > 0$, even after low signals in the past, would increase the expected long-term profit after a deviation and would thus increase the minimum discount factor that is required to achieve the social optimum such that it is larger than $\underline{\delta}_{NA}$. Note further that $\underline{\delta}_{NA}$ is decreasing in U_1, the willingness to pay for a certified product $\tilde{v} = 1$ if there have been only high signals in the past. Thus, the buyers would need to increase their willingness to pay to lower the minimum discount factor required to achieve the social optimum. Such an increase, however, would not be rational: the buyers would pay more than their valuation of a product of the highest true quality v. □

Lemma 7. *The buyers prefer a High-Quality Equilibrium over all other equilibria.*

Proof. The buyers make zero profit in any equilibrium since (1) they would never buy if they made a loss, and (2) the proof of Lemma 1 and 5 show that the certifier forces the supplier to extract all surplus from the buyers. Hence, the buyers obtain the same profit in all equilibria. Since I assumed that they then prefer high quality $v = 1$ in each period, they favor any equilibrium in which they buy a product of truly high quality. □

Lemma 5, 6 and 7 imply Lemma 2.

3.8.4 Model: Accreditation and sophisticated buyers

3.8.4.1 Proof of Lemma 3

The proof is analogous to the proof of Lemma 2, except that the condition that ensures that the certifier has no profitable deviation from offering a high-quality certification contract $(v, f) = (1, U_1 - c_S)$ changes:

$$\sum_{t=0}^{\infty} \delta^t (U_1 - c_S - c_C - g) \geq (U_1 - \beta c_C - g) + \delta(1-\lambda) \sum_{t=0}^{\infty} \delta^t (U_1 - c_S - c_C - g) \tag{3.29}$$

This yields the condition given in Lemma 3.

3.8.4.2 Proof of Proposition 1

Given the setup of the game, the buyers know the values of δ, $\underline{\delta}_A$, and $\underline{\delta}_{NA}$. Hence, they can enforce a High-Quality Equilibrium by playing the grim trigger strategy described in Lemma 5, with a slight adaption: If $\underline{\delta}_A > \delta \geq \underline{\delta}_{NA}$, they do not buy a certified product ($\tilde{v} = 1$) at a positive price $p_{\tilde{v}} > 0$ if the certifier is accredited. If $\underline{\delta}_{NA} > \delta \geq \underline{\delta}_A$, they do not buy a certified product ($\tilde{v} = 1$) at a positive price $p_{\tilde{v}} > 0$ if the certifier is not accredited. The certifier can only profit if they comply with the buyers' "conditions". Then, Proposition 2 and Lemma 3 apply, and the "if"-part of Proposition 1 follows. The proof of the "only if"-part is analogous to the proof of Lemma 6.

3.8.4.3 Proof of Proposition 2

It is easy to see that it is optimal to price at cost, i.e., to set $g_n = \frac{(\beta_n c_C)^2}{2} + K$.[57] Consequently, the Break Even constraint in Definition 3 is binding. Replacing the accreditation fee g_n in the Incentive Constraint, the remaining problem is to set β_n as low as possible but high enough to satisfy the Incentive Constraint. Let η be the Lagrange multiplier of the Incentive Constraint. Then the solution to this problem is

$$\beta_n = \eta \frac{1-\delta}{\delta \lambda c_C (1+\eta)} \tag{3.30}$$

If $U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} - K \geq 0$, the Incentive Constraint is slack, $\eta = 0$ and thus $\beta_n = 0$. However, then $\delta \geq \underline{\delta}_{NA} \geq \underline{\delta}_A$ and the entrance of the accreditation body is not welfare improving (see Corollary 1).

Assume instead:

$$U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} - K < 0 \tag{3.31}$$

This assumption corresponds to (A1) in Proposition 2. Then the Incentive Constraint is not slack at $\beta = 0$. Then $\eta > 0$ and the optimal quality of accreditation is the lowest $\beta_n \in [0, 1]$ that satisfies the Incentive Constraint:

$$U - \bar{c} - \frac{1-\delta}{\delta\lambda}(\bar{c} - \beta_n c_C) - \frac{(\beta_n c_C)^2}{2} - K \geq 0 \tag{3.32}$$

[57] This is done in many countries, for example in Germany (Deutsche Akkreditierungsstelle, 2017). In most countries and all EU member states, accreditation bodies are at least non-profit organizations (European co-operation for Accreditation (EA), 2019).

The derivative of the left-hand side of this quadratic inequality with respect to β_n is:

$$\frac{1-\delta}{\delta\lambda}c_C - \beta_n c_C^2 \begin{cases} \geq 0 & \forall \beta_n \leq \frac{1-\delta}{\delta\lambda c_C} \\ < 0 & \forall \beta_n > \frac{1-\delta}{\delta\lambda c_C} \end{cases}$$

The second order condition is negative, so the maximum is achieved at $\frac{1-\delta}{\delta\lambda c_C}$, which is always positive.

Thus, I have to check two cases to see whether the accreditation body can achieve Incentive Compatibility (and thus a High-Quality Equilibrium):

If $\frac{1-\delta}{\delta\lambda} \leq c_C$, we are in Case 1 and need that the Incentive Constraint is satisfied at $\beta* = \frac{1-\delta}{\delta\lambda c_C}$, since the maximum is achieved at this point.

If $\frac{1-\delta}{\delta\lambda} > c_C$, we are in Case 2 and need that the Incentive Constraint is satisfied at $\beta* = 1$, since the maximum under our restriction $\beta \in [0,1]$ is achieved at this point.

Case 1: At $\beta* = \frac{1-\delta}{\delta\lambda c_C}$, the Incentive Constraint becomes:

$$2(U - \bar{c} - \frac{1-\delta}{\delta\lambda}(\bar{c} - K) + \frac{1-\delta}{\delta\lambda} - \frac{1}{2}\left(\frac{1-\delta}{\delta\lambda}\right) \geq 0$$

Case 2: At $\beta* = 1$, the Incentive Constraint becomes:

$$2(U - \bar{c} - \frac{1-\delta}{\delta\lambda}c_S - K) - c_C^2 \geq 0$$

In both cases, there is no solution to the problem in Definition 3 if the Incentive Constraint of the given case is not satisfied. The lack of a solution means that under accreditation, buyers cannot enforce a High-Quality Equilibrium: $\underline{\delta}_A > \delta$ and the entrance of the accreditation body is not welfare increasing.[58] To minimizes potential welfare loss, the optimal quality of accreditation is zero: $\beta_n = 0$.

Note that the slope of the left-hand side of the Incentive Constraint is positive for any $\beta \leq \beta_n$ in both Case 1 and 2. Given that the Incentive Constraint is satisfied in both cases, a solution $\beta \leq 1$ necessarily exists that binds the Incentive Constraint. The existence of such a solution

[58]It can either be the case that a High-Quality Equilibrium can be achieved without accreditation (if $\delta \geq \underline{\delta}_A \iff U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} \geq 0$) or that no High-Quality Equilibrium can be achieved in any case (if $\delta < \underline{\delta}_A$)

implies

$$2(U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} - K) + (\frac{1-\delta}{\delta\lambda})^2 \geq 0 \tag{3.33}$$

so that the possible solutions are

$$\beta_1 = \frac{1-\delta}{\delta\lambda c_C} + \frac{1}{c_C}\sqrt{2(U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} - K) + (\frac{1-\delta}{\delta\lambda})^2} \tag{3.34}$$

$$\beta_2 = \frac{1-\delta}{\delta\lambda c_C} - \frac{1}{c_C}\sqrt{2(U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} - K) + (\frac{1-\delta}{\delta\lambda})^2} \tag{3.35}$$

It is obvious that $\beta_1 < \beta*$ both in case 1 and that $\beta_1 > 1$ in case 2. I therefore therefore verify that $\beta_2 \geq 0$: Assumption (3.33) implies

$$\beta_2 \geq 0 \Longleftrightarrow \frac{1-\delta}{\delta\lambda c_C} \geq \frac{1}{c_C}\sqrt{2(U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} - K) + (\frac{1-\delta}{\delta\lambda})^2} \tag{3.36}$$

$$\Longleftrightarrow \left(\frac{1-\delta}{\delta\lambda}\right) \geq 2(U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} - K) + \left(\frac{1-\delta}{\delta\lambda}\right)^2 \tag{3.37}$$

$$0 \geq 2(U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} - K) \tag{3.38}$$

This inequality is satisfied due to assumption (A1). Thus $\beta_2 \geq 0$.

Hence, there is a solution to Equation 3.32 if and only if $\beta_2 \leq 1$. Satisfaction of this inequality is trivial in case 1. Thus, there exists a solution in case 1 which only relies on assumption (A1) and $\frac{1-\delta}{\delta\lambda} \leq c_C$

In case 2, a solution exists if:

$$2(U - \bar{c} - \frac{1-\delta}{\delta\lambda}c_S - K) - c_C^2 \geq 0 \tag{3.39}$$

This inequality corresponds to the Incentive Constraint assumption (A2) in Proposition 2.

Thus, under assumptions (A1) and (A2) $\beta_n = \beta_2$ and the entrance of the national accreditation body is welfare improving according to Corollary 1. [59]

[59] If (A2) is not given, this problem has no solution. This means that under accreditation, no High-Quality Equilibrium can be achieved: $\underline{\delta}_A > \delta$ and the entrance of the accreditation body is not welfare increasing. The optimal accreditation quality is zero to minimize potential welfare loss: $\beta_n = 0$.)

We will have the following accreditation fee:

$$g_{n2} = \min\left\{ \left(\frac{1-\delta}{\delta\lambda}\right)^2 - \frac{1-\delta}{\delta\lambda}\sqrt{2(U-\bar{c}-\frac{1-\delta}{\delta\lambda}\bar{c}-K)+(\frac{1-\delta}{\delta\lambda})^2} + (U-\bar{c}-\frac{1-\delta}{\delta\lambda}\bar{c}-K); \frac{c_C^2}{2}+K \right\} \tag{3.40}$$

Finally I obtain way I write assumption (A2) in Proposition 2 in the following way: Recall that, next to assumption (A1), we need to be in case 1 or 2.

Case 1 requires:

$$\frac{1-\delta}{\delta\lambda} \leq c_C \Longleftrightarrow \delta \geq \frac{1}{\lambda c_C + 1}$$

Case 2 requires:

$$\delta < \frac{1}{\lambda c_C + 1}$$

and

$$2(U-\bar{c}-\frac{1-\delta}{\delta\lambda}c_S - K) - c_C^2 \geq 0 \Longleftrightarrow \delta \geq \frac{1}{\frac{\lambda}{c_S}(U-\bar{c}-\frac{c_C^2}{2}-K)+1}$$

3.8.4.4 Proof of Corollary 2

$$\frac{\partial\beta}{\partial\lambda} = -\frac{1-\delta}{c_C\delta\lambda^2}\left(1+\frac{\bar{c}-\frac{1-\delta}{\delta\lambda}}{\sqrt{2(U_1-\bar{c}-\frac{1-\delta}{\delta\lambda}\bar{c})+(\frac{1-\delta}{\delta\lambda})^2}}\right)$$

$$\frac{\partial\beta}{\partial\delta} = -\frac{1}{c_C\delta^2\lambda}\left(1+\frac{\bar{c}-\frac{1-\delta}{\delta\lambda}}{\sqrt{2(U_1-\bar{c}-\frac{1-\delta}{\delta\lambda}\bar{c})+(\frac{1-\delta}{\delta\lambda})^2}}\right)$$

It is easy to see that both derivatives are negative whenever $\bar{c} \geq \frac{1-\delta}{\delta\lambda} \Longleftrightarrow \lambda \geq \frac{1-\delta}{\delta\bar{c}}$. Note that this condition is satisfied whenever

$$\lambda \geq \frac{1-\delta}{\delta c_C} \Longleftrightarrow \delta \geq \delta \geq \frac{1}{\lambda c_C + 1}$$

If $\delta < \frac{1}{\lambda c_C + 1}$, the derivatives are negative if and only if

$$\frac{1-\delta}{\delta\lambda} - \bar{c} < \sqrt{2(U_1-\bar{c}-\frac{1-\delta}{\delta\lambda}\bar{c})+(\frac{1-\delta}{\delta\lambda})^2} \Longleftrightarrow 2(U_1-\bar{c}) > \bar{c}^2$$

This is a necessary condition for

$$\delta \geq \frac{1}{\lambda \frac{1}{c_S}\left(U_1 - \bar{c} - \frac{c_C^2}{2} - K\right) + 1}$$

which must be satisfied under assumption (A2) in Proposition 2, whenever $\delta < \frac{1}{\lambda c_C + 1}$.

3.8.4.5 Proof of Proposition 4

As argued above, the Incentive Constraint that ensures a High-Quality Equilibrium is only satisfied if $\beta_f \geq \beta_n$. In the rare case that β_1, defined in Appendix 3.8.4.3 is lower than 1, Incentive Compatibility would not be achieved for qualities above quality level β_1. This result is shown in the proof of Proposition 2 in Appendix 3.8.4.3. Hence, the second condition is $\beta_f \leq \beta_1$, which holds trivially when $\beta_1 > 1$. If that is the case, total welfare under the foreign accreditation body is higher than under the national body if and only if:

$$U_1 - \left(\bar{c} + \frac{(\beta_f c_C)^2}{2} + t\right) > U_1 - \left(\bar{c} + \frac{(\beta_n c_C)^2}{2} + K\right) \tag{3.41}$$

This yields condition (3) in Proposition 4.

3.8.5 Model: Accreditation and unsophisticated buyers

3.8.5.1 Proof of Proposition 6

Unsophisticated buyers do not know whether accreditation or no accreditation provides the certifier with enough incentives not to cheat. Hence, their best possible strategy is the grim trigger strategy from Lemma 5, as I have shown in the baseline model. The certifier decides whether or not to obtain accreditation by comparing the long-term payoffs in both cases.

If $\underline{\delta}_{NA} > \delta \geq \underline{\delta}_A$, then always obtaining accreditation would induce a High-Quality Equilibrium but is not incentive compatible. The certifier would only obtain accreditation if they had no incentive to deviate to no accreditation and cheating, i.e., if:

$$\sum_{t=0}^{\infty} \delta^t (U_1 - \bar{c} - g) \geq U_1 + \delta(1-\lambda) \sum_{t=0}^{\infty} \delta^t (U_1 - \bar{c} - g) \iff \delta \geq \frac{\bar{c} + g}{\lambda(U_1 - \bar{c} - g) + (\bar{c} + g)} \geq \underline{\delta}_{NA} \tag{3.42}$$

This inequality is a contradiction to the assumption $\underline{\delta}_{NA} > \delta$.

If $\delta \geq \max\{\underline{\delta}_{NA}, \underline{\delta}_A\}$, the certifier has no incentive to cheat no matter if they are accredited or not.

If $\delta < \min\{\underline{\delta}_{NA}, \underline{\delta}_A\}$, the certifier has the incentive to cheat no matter if they are accredited or not.

If $\underline{\delta}_A > \delta \geq \underline{\delta}_{NA}$, then never obtaining accreditation induces a High-Quality Equilibrium, and there is no profitable deviation to accreditation:

$$\sum_{t=0}^{\infty} \delta^t (U_1 - \bar{c}) \geq (U_1 - \beta c_C - g) + \delta(1-\lambda) \sum_{t=0}^{\infty} \delta^t (U_1 - \bar{c} - g) \tag{3.43}$$

This is trivially implied by the condition for $\delta \geq \underline{\delta}_{NA}$

$$\sum_{t=0}^{\infty} \delta^t (U_1 - \bar{c}) \geq (U_1 - \beta c_C - g) + \delta(1-\lambda) \sum_{t=0}^{\infty} \delta^t (U_1 - \bar{c} - g) \tag{3.44}$$

3.8.6 Model: Accreditation bias

3.8.6.1 Proof of Lemma 4

Buyers with accreditation bias believe that $\underline{\delta}_{NA} > \delta \geq \underline{\delta}_A$. That means that they expect that a certifier that is not accredited will cheat. Thus, they accept to buy a certified product ($\tilde{v} = 1$) at a positive price $p_{\tilde{v}} > 0$ if and only if the certifier is accredited and there has been no low signal in the past. Thus, the certifier can only profit if they obtain accreditation. Hence, they will always seek accreditation, so they will not cheat if and only if $\delta \geq \underline{\delta}_A$.

3.8.6.2 Proof of Proposition 7

Buyers with accreditation bias are willing to buy a certified product ($\tilde{v} = 1$) at a positive price $p_{\tilde{v}} > 0$ whenever the certifier is accredited. A monopoly profit-maximizing accreditation body may then have an incentive to cheat. They will set the quality of accreditation β only so high that they ensure the Incentive Compatibility of the certifier (β_p) if they have no profitable deviation. The most profitable deviation would be to set $\beta = 0$. The only restriction on the accreditation fee g would then be the Individual Rationality constraint of the certifier: $U_1 - \bar{c} - g_p \geq 0$. Thus their deviation profit would be $U_1 - \bar{c}$.

In the following, I show that this deviation is profitable whenever $\delta < \min\{\underline{\delta}_{NA}; \frac{1}{\lambda 2(1-c_S)+1}\}$. This condition is sufficient, not necessary.

Recall β_p and g_p from section 3.4.1.3.

I first consider case 1 from the proof of Proposition 2 in Appendix 3.8.4.3, where

$$\frac{1-\delta}{\delta\lambda} \leq c_C$$

such that

$$(\beta_p, g_p) = \left(\frac{1-\delta}{\delta\lambda c_C}, U_1 - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} + \left(\frac{1-\delta}{\delta\lambda}\right)^2\right)$$

Then the deviation to $(0, U_1 - \bar{c})$ is profitable if and only if:

$$\sum_{t=0}^{\infty}\delta^t(U_1 - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} + \frac{1}{2}\left(\frac{1-\delta}{\delta\lambda}\right)^2) \geq U_1 - \bar{c} + \delta(1-\lambda)\sum_{t=0}^{\infty}\delta^t(U_1 - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} + \frac{1}{2}\left(\frac{1-\delta}{\delta\lambda}\right)^2) \tag{3.45}$$

This is equivalent to:

$$\delta\lambda\left(U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} - K\right) - \frac{(1-\delta)^2}{\delta\lambda}\left(\bar{c} - \frac{1-\delta}{2\delta\lambda}\right) - (1-\delta)K + \frac{(1-\delta)^2}{2\delta\lambda} \geq 0$$

Whenever $\delta < \underline{\delta}_{NA}$, then $\left(U - \bar{c} - \frac{1-\delta}{\delta\lambda}\bar{c} - K\right) < 0$ (see Appendix 3.8.4.3). Then the above inequality cannot hold, whenever[60]

$$\frac{(1-\delta)^2}{\delta\lambda}\left(\bar{c} - \frac{1-\delta}{2\delta\lambda} - \frac{1}{2}\right) < 0$$

Since we are in the case $\frac{1-\delta}{\delta\lambda} \leq c_C$, a sufficient condition for that is:

$$c_S + \frac{1}{2}\left(\frac{1-\delta}{\delta\lambda} - 1\right) \geq 0 \Longleftrightarrow \delta < \frac{1}{\lambda 2(1-c_S) + 1}$$

In the rest of the proof, I show that whenever the deviation would be profitable in case 1, it would also be profitable in Case 2 where

$$\frac{1-\delta}{\delta\lambda} < c_C$$

[60] This is a sufficient, not a necessary condition.

Note first that Incentive Constraint 3.45 can also be rewritten as:

$$\delta\lambda\left(U-\bar{c}\right)-\left(1-\delta(1-\lambda)\right)\left(\frac{1-\delta}{\delta\lambda}\left(\bar{c}-\frac{1-\delta}{2\delta\lambda}\right)+K\right)\geq 0$$

Note that to derive the sufficient conditions in Case 1 above, I used the fact that $\bar{c}-\frac{1-\delta}{2\delta\lambda}\geq c_S+\frac{1-\delta}{2\delta\lambda}$. Using this implies that whenever the following condition is not satisfied, the Incentive Constraint 3.45 is not given either:

$$\delta\lambda\left(U-\bar{c}\right)-\left(1-\delta(1-\lambda)\right)\left(\frac{1-\delta}{\delta\lambda}\left(c_S+\frac{1-\delta}{2\delta\lambda}\right)+K\right)\geq 0$$

In Case 2, the deviation would not be profitable if

$$\delta\lambda\left(U-\bar{c}\right)-\left(1-\delta(1-\lambda)\right)\left(\frac{1-\delta}{\delta\lambda}c_S+\frac{c_C^2}{2}+K\right)\geq 0$$

Comparing the two constraints, we see that whenever $c_C\geq\frac{1-\delta}{\delta\lambda}$, the Incentive Constraint for Case 1 is easier to achieve and vice versa. This is precisely the case when we are in Case 1, and vice versa. Thus, the conditions sufficient for violating Incentive Compatibility in Case 1 are also sufficient in Case 2.

3.8.7 Conclusion

Figure 3.5: The label of the Marine Stewardship Council

Bibliography

(2008-08-13). Regulation (EC) No 765/2008 of the European Parliament and of the Council of 9 July 2008 setting out the requirements for accreditation and market surveillance relating to the marketing of products and repealing Regulation (EEC) No 339/93. *Official Journal of the European Union*, L 218:30–47.

AFNOR Group (2018). Liste des certifiés et evalués. https://certificats-attestations.afnor.org/certifie-detail/155023171193161227238041231176011123017243117151. [Online; accessed July-19-2018].

AfriCert (2017). Accreditation. http://africertlimited.co.ke/accreditation/. [Online; accessed May-23-2018].

Akerlof, G. A. (1970). The Market for “Lemons”: Quality Uncertainty and the Market Mechanism*. *The Quarterly Journal of Economics*, 84(3):488–500.

ANSI-ASQ National Accreditation Board. (2018). CB Directory. BSI Global-1251149 - BSI Assurance UK Limited. http://anabdirectory.remoteauditor.com/. [Online; accessed July-20-2018].

Arrow, K. J. (1963). Uncertainty and the Welfare Economics of Medical Care. *The American Economic Review*, 53(5).

Auld, G. and Gulbrandsen, L. H. (2010). Transparency in nonstate certification: Consequences for accountability and legitimacy. *Global Environmental Politics*, 10(3):97–119.

Auriol, E. and Picard, P. (2006). *Infrastructure And Public Utilities Privatization In Developing Countries*. The World Bank.

Bar-Isaac, H. and Shapiro, J. (2011). Credit ratings accuracy and analyst incentives. *American Economic Review*, 101(3):120–24.

Biglaiser, G. (1993). Middlemen as experts. *The RAND journal of Economics*, pages 212–223.

Biglaiser, G. and Friedman, J. W. (1994). Middlemen as guarantors of quality. *International Journal of Industrial Organization*, 12(4):509–531.

Bolton, P., Freixas, X., and Shapiro, J. (2012). The credit ratings game. *The Journal of Finance*, 67(1):85–111.

Bonroy, O. and Constantatos, C. (2014). On the economics of labels: How their introduction affects the functioning of markets and the welfare of all participants. *American Journal of Agricultural Economics*, 97(1):239–259.

CNAS (2012). Regulations of the People's Republic of China on Certification and Accreditation. https://www.cnas.org.cn/english/lawsandregulations/12/718389.shtml. [Online; accessed July-20-2018].

Cofrac (2018). AFNOR Certification. Accréditation(s) associée(s). https://www.cofrac.fr/fr/organismes/fiche.php?entite_id=41000578. [Online; accessed July-19-2018].

Competition Appellate Tribunal (2014). Appeal No. 03/2013. Order 17th January 2014. "http://accab.org/images/publication/15102014081908.pdf". [Online; accessed July-19-2018].

CROSQ (2014). Accreditation. https://www.crosq.org/index.php/quality-infrastructure/accreditation. [Online; accessed July-20-2018].

Darby, M. R. and Karni, E. (1973). Free competition and the optimal amount of fraud. *The Journal of Law and Economics*, 16(1):67–88.

Deutsche Akkreditierungsstelle (2017). Entgeltliste für die Akkreditierung von Konformitätsbewertungsstellen - KBS - (Laboratorien, Inspektions-, Zertifizierungs- und Verifizierungsstellen, Anbieter von Eignungsprüfungen und Referenzmaterialhersteller) im Ausland für Länder außerhalb der Gültigkeit der EU - Verordnung 7 65/2008. https://www.dakks.de/sites/default/files/dokumente/71_sd_001_entgeltliste_ausland_version_1.5.pdf. [Online; accessed July-20-2018].

Dranove, D. and Jin, G. Z. (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature*, 48(4):935–63.

Duflo, E., Greenstone, M., Pande, R., and Ryan, N. (2013). Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India. *The Quarterly Journal of Economics*, 128(4):1499–1545.

Dutch Accreditation Council (RvA) (2013). Policy rule for Cross-Frontier Accreditation . https://www.rva.nl/en/documents/rules-and-decisions. [Online; accessed June-13-2019].

Eguino, K. (2018). Personal communication.

European co-operation for Accreditation (EA) (2019). Relations with European Commission. https://european-accreditation.org/about-ea/relations-with-european-commission/#european-accreditation-infrastructure. [accessed June-14-2019].

European Commission (2014). CERTIF 2013-02– Requirement to seek accreditation in the Member State of establishment. Note to the expert group on the internal market for products., European Commission. Enterprise and Industry Directorate-General.

European Commission (2017). Report from the Commission to the European Parliament, the Council and the European Economics and Social Committee on the implementation of Regulation (EC) No 765/2008 of the European Parliament and of the Council of 9 July 2008 setting out the requirements for accreditation and market surveillance relating to the marketing of products and repealing Regulation (EEC) No 339/93. Report, European Commission, Brussels, B.

Farhi, E., Lerner, J., and Tirole, J. (2013). Fear of Rejection? Tiered Certification and Transparency. *RAND Journal of Economics*, 44(4):610 – 631.

Faure-Grimaud, A., Peyrache, E., and Quesada, L. (2009). The ownership of ratings. *The RAND Journal of Economics*, 40(2):234–257.

Fletcher, P. (2011). *Impact Investment: Understanding Financial and Social Impact of Investments in East African Agricultural Businesses.* Kampala: Pearl Capital Partners.

Freedom House (2017a). *Freedom of the Net 2017.* Washington: Freedom House.

Freedom House (2017b). *Freedom of the Press 2017.* Washington: Freedom House.

Galland, J.-P. (2017). Big third-party certifiers and the construction of transnational regulation. *The ANNALS of the American Academy of Political and Social Science*, 670(1):263–279.

Giannakas, K. (2002). Information asymmetries and consumption decisions in organic food product markets. *Canadian Journal of Agricultural Economics/Revue Canadienne D'Agroeconomie*, 50(1):35–50.

Glaeser, E. L. and Shleifer, A. (2001). Not-for-profit entrepreneurs. *Journal of Public Economics*, 81(1):99–115.

Hanushek, E. A. and Wößmann, L. (2007). *The role of education quality for economic growth*. The World Bank.

Heras-Saizarbitoria, I. and Boiral, O. (2019). Faking iso 9001 in china: An exploratory study. *Business Horizons*, 62(1):55–64.

Institute for Economics and Peace (2019). Global Peace Index 2019. http://visionofhumanity.org/indexes/global-peace-index/. [accessed June-14-2019].

International Accreditation Forum (2018a). IAF FAQ. About Accreditation. http://www.iaf.nu/articles/FAQ/288. [Online; accessed May-2-2018].

International Accreditation Forum (2018b). IAF MLA. http://www.iaf.nu/articles/IAF_MLA/14. [Online; accessed May-2-2018].

IOAS (2018). About IOAS. Quality. https://ioas.org/about-ioas/quality/. [Online; accessed July-20-2018].

ISO (2012). Conformity assessment — Requirements for bodies certifying products, processes and services. Standard, International Organization for Standardization, Geneva, CH.

ISO (2015a). *ISO/IEC 17021-1:2015. Conformity assessment – Requirements for bodies providing audit and certification of management systems – Part 1: Requirements*. Geneva: International Organization for Standardization.

ISO (2015b). Quality management systems — Fundamentals and vocabulary. Standard, International Organization for Standardization, Geneva, CH.

ISO (2017). *ISO/IEC 17011:2017. Conformity assessment – Requirements for accreditation bodies accrediting conformity assessment bodies.* Geneva: International Organization for Standardization.

Kastl, J., Pagnozzi, M., and Piccolo, S. (2018). Selling information to competitive firms. *RAND Journal of Economics*, 49(1):254 – 282.

Kirmes, R. (2019). Forschungsfrage zur Auslandstätigkeit der DAkkS. Personal communication.

Kitzmueller, M. and Shimshack, J. (2012). Economic perspectives on corporate social responsibility. *Journal of Economic Literature*, 50(1):51–84.

Lizzeri, A. (1999). Information revelation and certification intermediaries. *The RAND Journal of Economics*, pages 214–231.

Mailath, G. J. and Samuelson, L. (2006). *Repeated games and reputations: long-run relationships*. Oxford university press.

Mathis, J., McAndrews, J., and Rochet, J.-C. (2009). Rating the raters: are reputation concerns powerful enough to discipline rating agencies? *Journal of Monetary Economics*, 56(5):657–674.

MSC (2019). Certification bodies. https://www.msc.org/for-business/certification-bodies. [Online; accessed April-26-2019].

National Institute of Standards and Technology (2018). Compliance FAQs: Fastener Quality Act (FQA). https://www.nist.gov/standardsgov/compliance-faqs-fastener-quality-act-fqa. [Online; accessed July-20-2018].

OECD NCP (2018). FINAL STATEMENT of the German National Contact Point for the OECD Guidelines for Multinational Enterprises at the Federal Ministry for Economic Affairs and Energy in response to a complaint filed by the ECCHR, Garment Workers Unity Forum, Comrade Rubel Memorial Center, medico international e.V., FEMNET e.V., Bonn and Ms Raima Jahan, Mr Mahmudul Hasan Hridoy, Ms Rikta Khatun Joshna, Ms Morjina Begum and Ms Jesmin Akhter against TÜV Rheinland AG and TÜV Rheinland India Pvt. Ltd. https://www.bmwi.de/Redaktion/EN/Downloads/M-O/nks-abschliessende-erklaerung.pdf?__blob=publicationFile&v=2. [Online; accessed April-26-2019].

Olken, B. A. and Pande, R. (2012). Corruption in developing countries. *Annual Review of Economics*, 4(1):479–509.

Opp, C. C., Opp, M. M., and Harris, M. (2013). Rating agencies in the face of regulation. *Journal of Financial Economics*, 108(1):46 – 61.

Pannhausen, C. and Untied, B. (2010). Shifting development cooperation strategies to target regional agricultural trade for economic development in Sub-Saharan Africa. *Trade Negotiations Insights*, 9(7):4 – 5. [Online; accessed May-23-2018].

Quality Council of India (2015). Effectiveness of QMS of ISO 9001:2000 Certified Organizations (India). http://www.qcin.org/case-studies.php. [Report; accessed July-6-2018].

Ref. Ares(2014)2246326 (2014). Guidance papers on accreditation. Technical report, European Commission. Enterprise and Industry Directorate-General.

Skreta, V. and Veldkamp, L. (2009). Ratings shopping and asset complexity: A theory of ratings inflation. *Journal of Monetary Economics*, 56(5):678 – 695. Carnegie-Rochester Conference Series on Public Policy: Distress in Credit Markets: Theory, Empirics, and Policy November 14-15, 2008.

Stahl, K. and Strausz, R. (2017). Certification and market transparency. *Review of Economic Studies*, 84(4):1842 – 1868.

Strausz, R. (2005). Honest certification and the threat of capture. *International Journal of Industrial Organization*, 23(1-2):45 – 62.

Swinnen, J., Deconinck, K., Vandemoortele, T., Vandeplas, A., et al. (2015). *Quality Standards, Value Chains and International Development*. Cambridge: Cambridge University Press.

Transparency International (2018). Corruption Perceptions Index 2017. https://www.transparency.org/news/feature/corruption_perceptions_index_2017#research. [Online; accessed June-25-2018].

TÜV Rheinland AG (2015). 2014 Balance Sheet: TÜV Rheinl and Generates Strong Growth Worldwide. https://www.tuv.com/fr/france/apropos_de_nous/nouvelles/nouvelles-pdf_237075.jsp. [Online; accessed April-26-2019].

TÜV Rheinland AG (2016). Group Management Report. https://www.tuv.com/media/01 _presse_2/english_publications/corporate_report/TUEV_Rheinland_Corporate_Report_2015_Financial_Report.pdf. [Online; accessed April-26-2019].

TÜV Rheinland AG (2017). TÜV Rheinland Releases 2016 Annual Results. https://www.tuv.com/en/usa/about_us/press_1/press_releases/newspdf_usa_319106.jsp. [Online; accessed April-26-2019].

UNIDO, ILAC, and IAF (2017). *Setting up Accreditation Bodies in Developing Countries. A Guide to Opening the Door for Global Trade.* United Nations Industrial Development Organization, Vienna, Austria.

Van Leeuwen, B. (2017). *European standardisation of services and its impact on private law: paradoxes of convergence.* Bloomsbury Publishing.

World Bank (2017). Enterprise Surveys. Indicator Descriptions. http://www.enterprisesurveys.org/~/media/GIAWB/EnterpriseSurveys/Documents/Misc/Indicator-Descriptions.pdf. [Online; accessed April-22-2019].

World Bank (2019). GDP per capita growth (annual %). https://data.worldbank.org/indicator/NY.GDP.PCAP.KD.ZG?end=2017&locations=NI-UY&start=2002&view=chart. [Online; accessed June-11-2019].

World Justice Project (2019). World Justice Project (WJP) Rule of Law Index. http://data.worldjusticeproject.org/. [accessed June-14-2019].