Institute of Agricultural Policy

Justus-Liebig-University Giessen

The Characteristics of Fairtrade Research: A Meta-Review on the Designs of Studies that examine Fairtrade Certification Effects on hired Farm Workers

Master-Thesis

submitted by

Johanna Veslemøy Triebel

Sustainable Transitions M. Sc.

First Examiner: Prof. Dr. Martin Petrick

Second Examiner: Eduard Bukin

Auckland, 3rd November 2024

Table of Contents

List of Tables III

List of Figures III

Abbreviations III

Abstract IV

1 Introduction 1

1.1 Research Problem 1

1.2 Research Motivation 1

1.3 Research Objectives and Outline 1

2 Problem Background 3

2.1 Background Knowledge on Fairtrade 3

2.2 Background Knowledge on Hired Farm Workers 7

2.3 Contribution of Fairtrade to the situation of hired Farm Workers in developing Countries 9

3 State of Research 11

3.1 State of Primary Research: The Impact of Fairtrade on hired Farm Workers – A Literature Review 11

3.2 State of Meta-Research: The Impact of Fairtrade on hired Farm Workers 17

3.3 Digression: Assessment of Fairtrade's performance in general 19

3.4 Scientific Standards in Fairtrade Research 21

3.4.1 Primary Research 21

3.4.2 Meta Research and Theory Building 22

3.4.3 Impact Evaluations 24

3.4.4 Generalizability in Field Research 25

3.4.5 Assessing the Studies' Quality in this qualitative Meta-Research 27

4 Methods 29

4.1 Descriptive Review: The Study's Frame 29

4.2 Literature Research 30

4.2.1 Research Scope 30

4.2.2 Literature Search 31

4.3 Criteria Coding 32

4.3.1 Development of the Criteria Catalogue 33

4.3.2 Explanation of the Catalogue Criteria 34

4.4 Editing of the Data 42

5 Results 42

5.1 Descriptive Results 42

5.1.1 Part 0: The Study's Frame 42

5.1.2 Part 1: The logical Derivation of the Research Approach 45
5.1.3 Part 2: The Study's Design 46
5.1.4 Part 3: The Study's Methods 49
5.1.5 Part 4: The Data Editing, Analysis and Results 49
5.1.6 Part 5: The Study's Quality 51
5.2 Examination of Correlations 52
6 Discussion on the Findings 55
6.1 Correlations 55
6.2 Quality of the Studies 57
6.3 Generalizability 60
6.4 Research Constraints 61
7 Conclusion 62
8 Publication bibliography 1
9 Appendix 6
9.1 Appendix 1 – Overview on the Catalogue Criteria with the respective Categories, Excel Abbreviations and Results 6
9.2 Appendix 2 – Template to be filled out for each Study 15
9.3 Appendix 3 – Results summarized in Excel Table 20

List of Tables

Table 1 - Research Objectives and Questions (own figure, cf. Bissinger (2019)) 2
Table 2 - Extract from the pricing database of Fairtrade International, with two different coffee products (own figure, cf. Fairtrade International (2024c)) 4
Table 3 - Extract from the Fairtrade Standard Catalogue for Small-Farmers (Fairtrade International 2019) 10
Table 4 - Overview of the studies included in the examination of this thesis (own figure) 15
Table 5 - Methodological Approaches in Fairtrade Research, with some Examples (own figure) 21
Table 6 - Studie's Quality and the referring Catalogue Criteria 28
Table 7 - Examples of Studies not included into the Research, with the respective Reasons. 32
Table 8 - Allocation of Studies split up on the Structure of Questioning and the Kind of Methods 45
Table 9 - Descriptions of the labour kind in the respective studies, and the tendency of the results 48
Table 10 - Distribution of studies by the kind of methods and the kind of examination regarding workers (own figure) 49
Table 11 - Publications with negative monetary results (own figure) 50
Table 12 - Studies conducted by Sociologists 53
Table 13 - Allocation of Study Results in Relation to the existence of a Theory 54
Table 14 - Allocation of Study Results in Relation to the Structure of Questioning 54
Table 15 - Characteristics that tend to emerge together with to positive/ negative results. 56

List of Figures

Figure 1 – Locating the study design in the meta-research field (own figure, cf. Paré et al. (2015) and Snilstveit et al. (2012)). 29
Figure 2 - Visualization of the Research Approach (own figure) 32
Figure 3 - Overview of the six main catalogue parts, with an exemplary section of the sub-criteria in the part Design (2) (own figure). 34
Figure 4 - Distribution of studies by publication year (own figure). The data refers to the criterion publication year (0.1.3). 42
Figure 5 - Distribution of studies by the scientific background of the researchers (own figure). 43
Figure 6 - Distribution of study results by continent of data collection (own figure) 43
Figure 7 - Distribution of studies by farm products (own figure). 44
Figure 8 - Word cloud on the problem backgrounds stated in the studies (own figure) 44
Figure 9 - Number of studies addressing the respective criterion (own figure). 58
Figure 10 - Criteria contributing to the assessment of internal validity, and the respective number of studies addressing the criteria (own figure). 59

Abbreviations

ToC – Theory of Change
RCTs – Randomized controlled trails
CEval – Centre of Evaluation

Abstract

The *Fairtrade International* institution aims to support farmers and workers in developing countries, with a hope of ensuring equitable product prices for farmers and decent working conditions. The impacts of the fairtrade system have been examined in many studies. However, a quick review of the literature indicates that the insights of those studies are often contradictory. Since fairtrade is a huge field of research, this thesis will focus only on fairtrade's impact on hired farm workers. It aims to explain and compare how the existing studies were carried out. By analysing the various characteristics of the studies, this paper will seek to demonstrate whether the results of certain studies are influenced by bias in the design. Furthermore, it is discussed how the quality of such a heterogeneous research field can be assessed, and on this basis, if it is possible to draw generalizable conclusions on the impact of fairtrade on hired farm workers. These objectives and questions are addressed with a descriptive approach. A systematic literature review is followed by a structured analysis of the characteristics of the studies. Therefore, a catalogue of criteria is established and used to categorize each study. The results are analysed using frequency distributions and contingency tables. 18 studies were analysed, with an average of 255 participants. Six of them were conducted in Africa, seven in Central and Latin America and three in Asian countries. Most of the studies did reveal small impacts of fairtrade, either negative or positive. To emphasize the contrast between the study characteristics, the category *result tendency* was introduced. This category demonstrates whether results are rather positive or rather negative, regardless of the strength of the effect. On this basis, the following characteristics occur mainly with positive tended results: a cross-sectional design (rather than longitudinal design), a sample size =< 100 (rather than a sample size > 100), a descriptive or explorative approach (rather than testing-hypotheses approach) and when not only workers but additional persons got examined (for example management staff of a company). These findings suggest that the design of the studies might have biased their results. In contrast, mixed-methods designs were revealed to be reasonably balanced in their research approaches and results, which proposes the use of mixed-methods designs for further research. Furthermore, it was revealed that most of the studies examine workers on plantations, only a few studies focussed on workers on small-farms. Moreover, on those small-farms the impacts of fairtrade were rather negative, whereas, on plantations, the impacts were often rather positive. These findings point to two important questions: why is the research on small farms that scarce? And why are the impacts of fairtrade on small-farms rather negative? It seems that fairtrade does not make a decisive impact in contributing to fairer trading conditions worldwide. This study concludes that further research is necessary, especially in the context of fairtrade on small farms. This is in order to attempt to understand the cause of the discrepancy and, the barriers and facilitators that make fairtrade work, or prevent it from working.

1 Introduction

1.1 Research Problem

Fair Trade is a worldwide movement, supporting fair trading conditions along the supply chain. The institution *Fairtrade International* offers a certification system for farms in developing countries. The farm has to meet certain production standards, e. g. regarding environmental protection and the wellbeing of workers. In return, *Fairtrade International* provides regulations around wages for workers and a pricing and labelling system that aims to improve the financial stability of the farm. *Fairtrade International* states that it "changes the way trade works through better prices, decent working conditions and a fairer deal for farmers and workers in developing countries" (Fairtrade International 2024b). Research has been carried out on the impacts of fairtrade and the results are ambiguous. Some authors conclude that fairtrade meets its own targets, while many others suggest negligible or even negative effects of fairtrade (Oya et al. 2018; Johannessen and Wilhite 2010; Mitiku et al. 2017). To address the cause of the discrepancy, the focus can be narrowed down to different parts of the fairtrade system. For example, the impact of fairtrade on hired farm workers, who are an important part of the fairtrade system. The body of research on this specific topic is small, with diverse conclusions on the impact of fairtrade.

1.2 Research Motivation

Fairtrade was a concept that I held a favourable attitude towards, believing that it was useful for empowering the poorest producers and workers in the food supply chain – until I started researching the topic. I realized that the scientific evidence on the impact of fairtrade is quite nuanced, and not always positive for farmers and workers. "Fairtrade is a system that leads to happy farm workers, provided with healthy working conditions and decent earnings". This is the idea portrayed to consumers who are prepared to pay higher prices on fairtrade products, and it is probably the expectation of many Western consumers. The ambiguity of the impact of fairtrade, as revealed by researchers, motivated this study to attempt to obtain a better understanding of the field. Many studies already seek to determine if fairtrade actually meets its own goals. However, this study aims to investigate the other studies into the efficacy of fairtrade – how they are performed and why their results are so different. By understanding the existing research better, the aim is to determine where to set the focus for further research and how future research methodologies can be improved.

1.3 Research Objectives and Outline

The main goal of this thesis is to investigate the methodologies and biases of studies about fairtrade's impact on hired farm workers. Respective studies are categorized and compared on the basis of their designs. The examination is performed using a catalogue of criteria, which contains categories for several study characteristics. Each criterion is then analysed with frequency distributions and

contingency tables. The descriptive review is used as the framework which contains literature research, systematic comparisons, and descriptive statistics. Similarities and differences between the studies are outlined, and possible biases of the studies´ designs are discussed.

The research objectives and questions are summarized in *Table 1*. The first objective focusses on the state of research in the field of fairtrade's impact on hired farm workers. By conducting a literature review, the studies used for this analysis are briefly described and their results are presented. The second objective involves a comparison of those studies, developing a catalogue of criteria to reveal the characteristics of the studies. The characteristics are then also used to assess the studies´ quality. Objective 3 discusses the results of the studies, whether they can be generalized, and what limitations to the generalizability can be identified.

Table 1 - Research Objectives and Questions (own figure, cf. Bissinger (2019))

No.	Research Objectives (O) and Questions (RQ)
O1	Review the literature on the effects of fairtrade certification on hired farm workers.
RQ1	*What kind of studies exist in the research field? What insights do they reveal?*
O2	Gain a clear picture of the studies.
RQ2	*How are the studies designed? Are there correlations who might let assume the existence of biases?*
RQ3	*How can the studies quality be assessed?*
O3	Conclude on the possibility of generalization of the studies´ results.
RQ4	*What limitations constrain the generalizability?*

The research objectives are approached by the research questions.

The remainder of the thesis is the following. In *chapter 2* the problem background is represented. The focus will be on the fairtrade system, the working conditions of hired farm workers in rural agriculture in developing countries, and the ways that fairtrade contributes to worker well-beings. *Chapter 3* lays out on the state of research in the research field. Firstly, a literature review is provided, on the studies that are analysed in this thesis, addressing *RQ1*. Furthermore, the state of meta-research is described and the scientific standards in the field of fairtrade research are discussed. *Sub-chapter 3.3* provides a short assessment of fairtrade's performance in general. *Chapter 4* explains the methodology used in this thesis. The approach of a descriptive review is explained, as well as the literature research, and the approach to develop a catalogue of criteria is shown. *Chapter 5* explains the results of the analysis. It is divided into two main parts: 1) the descriptive results, and 2) the examination of possible correlations (addressing *RQ2*). *Chapter 6* provides a discussion of the answers to the research questions, *RQ3* and *RQ4*. Eventually, *chapter 7* presents a conclusion of the findings of the thesis and contains recommendations for further research.

2 Problem Background

This chapter includes background information about the fairtrade system, the working conditions in rural agriculture in developing countries and the relationship between both. It shall contribute to a deeper understanding of the circumstance in which fairtrade studies take place.

2.1 Background Knowledge on Fairtrade

In this sub-chapter, at first it is described how the fairtrade label developed and how different spellings can be distinguished. Second, the fairtrade system is explained, with its two main pillars and the basis Theory of Change. Third, the evaluation of fairtrade is represented.

History of Fairtrade

In this paragraph, a brief description of the development of the fairtrade label is provided. The ambition to realize fair trading partnerships was a matter already in 1972, when the fair trade organisation *El Puente* was founded (Brot für die Welt 2024). But trade in a fair manner took shape for the first time, when the Dutch *Max Havelaar* foundation launched the first fairtrade label in the Netherlands in 1988 (Wikipedia 2023). And some years later, in 1991, the first German fairtrade labelling organization, *Transfair e. V.*, was founded. As the fairtrade movement grew, there was the aim to internationalize it, which happened in 1997, when *Fairtrade International* was founded by 17 national fair trade initiatives (formerly called *Fairtrade Labelling Organization International*) (Brot für die Welt 2024). Nowadays *Fairtrade International* is the greatest fairtrade labelling organization in Germany, owning 81 % of total sales in the fairtrade market in 2023 (Forum Fairer Handel 2023). The fairtrade movement started with coffee certification in Latin America. Nowadays, the certification can also be found on many other food products, with cocoa, bananas and coffee being the products with the greatest share on total fairtrade food production (Fairtrade Deutschland 2024). Besides the food sector, fairtrade certification can also be found in flower, cotton and gold production, as well as on handicraft products (ibid).

Distinguishing different Spellings

For the understanding of the term “fair trade”, it is helpful to distinguish two different spellings. The term “fair trade”, written with two words, refers to the overarching goal of fair trading relationships (Raynolds 2014). It is a worldwide movement, promoted by countless companies and organizations but independent of a certain labelling. In contrary, the term “fairtrade” (one word) refers to the system of fair labelling, especially focussed on the label of *Fairtrade International* (Raynolds 2014). Since there is no official agreement about this differentiation, recurring confusion exists (Gröne et al. 2020). Beyond that, it is not defined if the terms must be written with capital letters, so it is done in different ways of writing. As several authors make use of the differentiation already, so will the author of this

thesis. In the following *fairtrade* refers to the organization *Fairtrade International* only and *fairtrade system* refers to the certification system of *Fairtrade International* only. (Raynolds and Bennett 2015; Bissinger 2019; Nelson and Pound 2009)

The Fairtrade System

This part represents the aim and approach of *Fairtrade International*.

Fairtrade International states on its webpage, that it aims to change "the way trade works through better prices, decent working conditions and a fairer deal for farmers and workers in developing countries" (Fairtrade International 2024b). Another statement is included in the vision of fairtrade, namely: "A world in which all small producers and workers can enjoy secure and sustainable livelihoods, fulfil their potential and decide their future" (Fairtrade International 2016). Workers are explicitly mentioned here, but the specific aims stay rather blurry. For the fairtrade system two main pillars are used to approach the vision. At first, there is the fairtrade minimum price. This price is paid to the producers at a minimum, even if the world market price decreases even lower. If the world markets price exceeds the fairtrade minimum price, the producers would get the world market price. This pillar is meant to hedge fairtrade producers from high price volatility on the world market and offers investment possibilities due to financial planning certainty. A minimum price is set for every fairtrade product, but differs depending on the country and, in some cases, depending on the degree of processing. On its website, Fairtrade International publishes the pricing database, which contains the minimum price and the price premium for every product. As to see in *Table 2* the product *coffee* comes with different characteristics. Conventional coffee Arabica, which is in its natural form, has a fixed minimum price of 1.75 USD per pound. In comparison, organic coffee Robusta, which is washed already, has a fixed minimum price of 1.65 USD per pound. The prices for both coffee products are applied worldwide, with no differences between certain countries.

Table 2 - Extract from the pricing database of Fairtrade International, with two different coffee products (own figure, cf. Fairtrade International (2024c))

Product	Quality	Form	Country	Minimum Price USD/ pound	Price Premium USD/pound
Coffee	conventional	Arabica; natural	worldwide	1.75	0.2
Coffee	organic	Robusta; washed	worldwide	1.25 + 0.4 (organic differential)	0.2

The Products are sorted by their quality and form. For coffee, the quality categories are conventional and organic. The form includes the coffee bean type and the state of processing. The table furthermore contains the information on which region the minimum price and price premium get applied. In the case of coffee, they are applied worldwide. The minimum price and the price premium are stated in USD per pound, and for the organic coffee there is an additional organic surcharge. Conventional arabica coffee in its natural form obtains a minimum price of 1.75 USD per pound and a price premium of 0.2 USD per pound.

Organic Robusta coffee in washed form obtains a minimum price of 1.25 USD per pound, with additional 0.4 USD per pound because of the organic quality, and a price premium of 0.2 USD per pound.

The second pillar is called the price premium. This premium is paid additionally for every sold product. While the minimum price is paid directly to the farmer, the price premium is paid to the fairtrade cooperative a fairtrade farmer must be part of. This premium is meant to support local development. What exactly it is used for, will be decided democratically at the cooperative level, it mostly contributes to the improvement of facilities in the community or farmer teachings of agricultural knowledge and techniques. For all coffee products the price premium is set at a level of 0.2 USD per pound (*Table 2*). These two pillars are at the core of the fairtrade system. For receiving the label for certain products, a farm has to pass a process of adjustment, to ensure that all the requirements of fairtrade are going to be met. Furthermore, the owner has to pay the certification fee to fairtrade. The process will be monitored independently, and at a certain point of progress, farm products can be labelled and the farmer will receive the fairtrade minimum price.

With this certification system, fairtrade aims to fulfil its vision. But the question is: Are "secure and sustainable livelihoods, fulfilling of potential and own decisions on the future" (Fairtrade International 2016) the actual outcomes of this approach? In 2015, fairtrade published a *Theory of Change* (ToC), aiming to make the strategy reasonable and support their work with a theoretical foundation. This theory includes interventions of fairtrade, which lead to short-term outcomes, mid-term outputs and long-term impacts. For instance, regarding the wellbeing of hired workers, an intervention of fairtrade is to provide "support to workers and their organizations" in order to "increase freedom of association" as an short-term outcome, which leads to "decent work" as a mid-term output, which leads to "improved household income, assets and standard of living" as an long-term impact (Fairtrade International 2016). This chain of reasoning is reasonable at first glance, though it is not obvious if it holds true when fairtrade gets applied on a farm. Much research has been done in the last years, both insisted by fairtrade and independently, and the results recure to be ambiguous (Oya et al. 2018; Jodrell and Kaoukji 2020). Therefore, it remains to be a core task of fairtrade to develop reliable evaluation.

The Evaluation of Fairtrade

The evaluation of the fairtrade system focusses on several aspects along the production and consumption chain. One of the first examinations was done by Kunz (1987), who examined the impact of the so called *Dritte Welt Läden* (third world shops) in Germany on their potential of developmental education. German research on fairtrade is mainly focussed on the consumer side of the chain (Gröne et al. 2020). Other examinations focus on the trading process or analyse the distribution of benefits along the supply chain (Valkila et al. 2010; Maertens 2019). A huge body of examinations is done on the farmer and cooperative level, examining the impact of fairtrade certification (Chiputwa et al. 2015;

Dragusanu et al. 2022; Hoebink et al. 2014). And another important, though not well developed, part of the fairtrade research is the examination of the impact on hired workers (Raynolds 2014; Cramer et al. 2014). The latter will be elaborated in *chapter 3.1* more comprehensive.

Former and recent Impact Evaluations on Fairtrade

The first impact evaluation on fairtrade was executed in 2012 by the *Centre for Evaluation* in Germany (Brot für die Welt 2024; CEval 2012). Six cases had been examined, focussing on different products and countries. The case studies were designed quasi-experimental, with a target group and a matching comparison group for each case. Regarding workers welfare, the evaluation revealed in some cases positive effects of certification on health and safety issues, the existence of written contracts, access to micro-credits, and educational programmes. Though, no increase of wages was observed. (CEval 2012)

Another approach of *Fairtrade International* to examine its impacts, is the use of literature reviews. Irregular summaries every couple of years had been published for the last 25 years. The first literature meta-review of fairtrade examined a period from 1999 to 2009 (Nelson and Pound 2009). 33 studies were included, all of them are peer-reviewed and with the focus on the local producer level. Most of them focus on smallholder coffee farmers in Latin America, only two studies examine the impact of fairtrade on hired workers. It can be mentioned for example that there was very little evidence on the empowerment of workers and that workers gain more from the fairtrade benefits when they know about the fairtrade system and their rights. Nevertheless, workers got provided with interest-free loans. The authors emphasize that the insights cannot easily be generalized, due to the small number of examinations, which additionally focus on several different aspects of fairtrade's impact. Nelson and Pound (2009) explain that "the economic inputs of [Fairtrade] are intended to lead to positive impacts *for producers and workers*" (p. 6). It is noticeable that the intentions of fairtrade always include producers and workers to the same amount, though the interventions and the examinations of the interventions are mainly focussing on the producers. A greater focus on workers is recommended. (Nelson and Pound 2009)

The subsequent literature meta-review covers the years from 2009 to 2015 (Darko et al. 2017). Forty-five studies had been included, again most of them examining the coffee sector in Latin America and Africa. Regarding the wellbeing of workers, this review discovers that the impact of fairtrade on wages and working conditions of workers is limited. However, the authors indicate that indirect impacts of fairtrade are important for workers wellbeing as well, in particular a better knowledge network and increased infrastructure stability due to the fairtrade premium, among others. Furthermore, the review reveals that the impact of fairtrade is highly dependent on the local political and economic situation, without the chance for fairtrade to have any influence on it. Additionally, this review

addresses the importance of standardized working conditions for small farms. Fairtrade distinguishes between small farms, which are mostly family owned and the farmers are self-employed, and plantations, which include hired labour situations. In the beginning of fairtrade, only small farms were covered by the certification, but as the label grew on market share and importance, it was decided to include also plantations into the system (Murray et al. 2006). Subsequently, standards for working conditions on plantations were developed, but no for small farms. The authors criticise that only plantations were covered. It reveals that also smallholders employ workers, which are however not considered by the fairtrade requirements. The authors assume a huge number of "invisible workers", marginalised within the fairtrade system. Therefore, standard working conditions for small farm situations would be necessary. (Darko et al. 2017)

The most recent literature meta-review covers the period from 2015 to 2020 (Jodrell and Kaoukji 2020). This research is based on the *Theory of Change* (ToC), which was developed in 2015, and includes 151 studies, 30 of them focussing on hired workers directly. The ToC aims to improve the monitoring, evaluation and learning process to discover respective mid-term and long-term impacts. The studies had been analysed with a questionnaire and rated on their quality. Each study contributes to the evaluation results depending on the rating score. The review reveals improved working conditions due to fairtrade certification, for example renumeration for overtime and adequate breaks, among others. Also, a higher job satisfaction was discovered. Fairtrade's requirements on empowerment and wages are said to be rather contributory, and again the circumstances in the local settings are crucial. However, it is criticised that hired workers in the small farming context are still not in the focus of fairtrade. (Jodrell and Kaoukji 2020)

2.2 Background Knowledge on Hired Farm Workers

Working Conditions in rural Agriculture in developing Countries

In this chapter, a brief insight in rural working structures is represented, since in those working situations fairtrade certified production takes place. Worldwide, there are about 450 million wage workers working in the agricultural sector (International Labour Organization 2024). Some of them are employed full-time or part-time but the majority is employed seasonally, sometimes only on a casual or temporary agreement (Hurst 2007). The payment is done in cash, but also in-kind payment is common, or a combination of both (ibid). Payment can be differentiated in hourly payment and performance-based, and a combination of both is common as well (Valkila and Nygren 2010). Furthermore, employment situations can include full year employment but with payment only during harvest season (Meemken et al. 2019). The workers are employed on small farms by the farmer itself, and on plantations by the plantation company or by a labour contractor (Hurst 2007). Since agricultural work requires a great number of workers but only seasonally, making use of labour contractors is increasingly common (ibid). Workers, who are hired by a contractor often are migrant workers from a

different region or even from different countries (ibid). Wage-work is also done by farmers, who do own a small farm, but however do not generate enough income with it (ibid). The share of woman working in agriculture is estimated to be about 20 to 30 percent, with some variations worldwide (Hurst 2007).

This description on working conditions in rural agriculture is to be seen as approach to the complex field, but is not to be taken as complete neither to be generalized. There is a vast variation on the working conditions and kinds of employment among plantations as well as among small farms, even if they are located in the same region (Cramer et al. 2014; Sender et al. 2005).

Distinguishing certain employment relations

Fairtrade distinguishes employment relations in *hired labour organizations* and *small-scale producers*. *Hired labour organizations* are organizations that depend on hired work force on a permanently basis, for example flower or tea plantations. Small-scale producers are defined to be independent of hired workers. Fairtrade expect that the work on small-scale production sites is mainly executed by family members. However, Cramer et al. (2017) reveal that actually in small farm structures a dependency on hired labour exists, which is not sufficiently addressed by the fairtrade standards. As presented in the paragraph above, the employment conditions are more complicated than the two standard systems of fairtrade might let assume. Fairtrade seems to simplify the rural situation and fails to include decent requirements in their certification standards. For instance, conditions for casual workers, other than to keep record of them, are not addressed in the standards of fairtrade (Fairtrade International 2019, 2024a).

Furthermore, in the literature on hired farm workers a variety of descriptive terms is used. In the fairtrade standards, workers are addressed as *workers*, with sub-categorizations like *temporary*, *seasonal*, *permanent* or *migrant* workers, among others (Fairtrade International 2019). Dragusanu et al. (2022) distinguish between *skilled workers*, which are said to be farmers mostly, *unskilled workers*, which refers to hired labour and *non-farm occupations*, involving the working tasks sales, storage, transport or processing. The companies addressed by the term *hired labour organizations* are also referred to as *plantations* or *estates*. Small farms are described as *small-scale producers*, *small farmers* or *small sites*, among others. However, regarding the studies analysed in this thesis, the employment situation is not always described in detail. Therefore, in the remainder of this thesis the term *hired farm workers* is used for any of the worker categories and employment situations, and will only be specified if necessary.

Challenges of those working Conditions

Some of the challenges, rural workers are facing, emerge by those working conditions. Seasonal employment leads to high unemployment rates in off-season, which endangers decent living

conditions of wage workers. Men are more likely to migrate to non-agricultural work, which is less easy for women, as they do most of the care work in households. Therefore, women are more dependent on local agricultural work and hence are more vulnerable to exploitive employments. Moreover, there is a continuing gender pay gap, with women being paid less than men. Another vulnerable group are the migrant workers, as mentioned, often hired by contractors. Contractors are suspected to abuse their authority, by overcharging hired workers for transport, housing and food, or by holding back wages. Wage levels, employment security, health and social protection are often endangered in those employment conditions. To establish a minimum wage in rural work situations is another challenge, since contracts are rather informal and difficult to track, and a high labour surplus exists due to unemployment. Nevertheless, the International Labour Organization (2023) makes effort towards a minimum wage and emphasizes it in its policy recommendations. (Hurst 2007)

2.3 Contribution of Fairtrade to the situation of hired Farm Workers in developing Countries

About 1.9 million out of 450 million wage workers in agriculture are employed under fairtrade certification standards (Fairtrade International 2024d). Fairtrade focusses on two main working situations; self-employed small farmers, who are organized in so called *producers organisations*, and hired labour conditions on plantations. For both settings fairtrade developed standards, which contain alle the requirements a farm has to fulfil in order to get the certification. The standards for producer organizations are described in extracts hereinafter. Regarding wage workers, the standards include requirements for the case that a farmer employs ten or above employees (Fairtrade International 2019). The standards include requirements on working hours, breaks, sick and maternity leave, overtime, contracts and wages, among others. *Table 3* shows some of the requirements exemplarily.

Table 3 - Extract from the Fairtrade Standard Catalogue for Small-Farmers (Fairtrade International 2019)

Priority	Explanation
Wages	
Core Year 0	You and your members set salaries for workers according to CBA regulations where they exist or at regional average wages or at official minimum wages for similar occupations whichever is the highest. You specify wages for all employee functions and employment terms, such as piecework.
Subcontracted workers	
Core Year 0	If you or your members employ migrant or seasonal workers through a contracting agency or person, you put effective measures in place to ensure that their hiring and working conditions also comply with this Standard.
Permanent employment	
Development Year 3	Where possible you and your members assign all regular work to permanent workers.
Maternity leave, social security and other benefits	
Development Year 6	You and your members set maternity leave, social security provisions and non-mandatory benefits according to national laws or according to CBA regulations where they exist, or according to the agreement signed between the workers' organization and the employer, whichever is the most favourable for the worker.

The table shows the subcategories of the standard catalogue on the left side, as well as the priority of each requirement. The right column presents what exactly is required in each category.

Core requirements have to be fulfilled in year zero or year one after certification. *Development* requirements are scheduled for year three or year six. For example, wages have to meet collective bargaining agreements (CBA) or the regional wage average from year zero on. The one, which is the most favourable for the workers, is to choose. Requirements for subcontracted workers are also to meet from year zero on. The farm has to ensure, that all the requirements of the standard are in force also for subcontracted workers. Tasks, that have to be fulfilled regularly, should be done by permanently employed workers. "Where possible" is quite blurry wording, but it is named at least (cf. *Table 3*). This requirement has to be fulfilled within year three after certification. Also, "maternity leave and social security" are a development requirement and is scheduled for year six (cf. *Table 3*). Again, the requirement is rather blurry, referring to national laws or collective bargaining. The standards for small farmer organizations include several more requirements, also on environmental management or the use of hazardous material, among others. In small farm organizations, the farmers get the fairtrade minimum price, but the premium will be paid to the organization and the farmers have to decide democratically how to use it.

A separate standard catalogue was developed for greater companies, plantations, with a huge number of hired workers. The catalogue addresses the same topics, though the exact requirements differ in parts. On plantations, a fairtrade premium committee administers the fairtrade premium. The

committee consists of management staff, but also worker representative, who have to be elected democratically. (Fairtrade International 2024a, 2019)

The Auditing within the Fairtrade System

The requirements of both standards are audited by *FLOCERT,* the Fairtrade Labour Organization Certification Company, which is the certification entity in the fairtrade system. A farm that wants to get certificated, would contact *FLOCERT* and after an auditing process, *FLOCERT* will issue the certification. The certificate will hold for three years, and has to be renewed afterwards, which would be accomplished by a new audit. (Fairtrade International 2024e)

Challenges of the Fairtrade System

The fairtrade system comes along with certain challenges from which two shall be presented here exemplarily. First of all, while fairtrade was originally meant to empower hired farm workers, in reality fairtrade has to remain in the tolerance borders of farm owners with its requirements in order to offer a serious trade alternative to farmers (Makita 2012). This compromise leads to blurry requirements on working conditions and wages (Raynolds 2014). Moreover, the standards only address employment situations with ten or more employees. Small farmers with less than ten permanently hired workers do not have to follow those standards. Secondly, several authors ascertain unequal power relation between workers and plantation owners. Makita (2015) and Besky (2014) reveal a reinforcement of a patron-client relationship due to fairtrade certification, which is historically rooted on tea plantations due to former colonial structures. Likewise did Raynolds (2014) discover paternalistic patterns on flower farms to be deepened by fairtrade. All three authors criticise unequal power relations but at the same time refer to positive effects that arise from those relations, like secure and reliable living and working conditions. Moreover, Oi and Idson (1999) emphasise a positive link between the farm size and wage levels, which contributes to livelihood improvement of workers on plantations. It is to see that the fairtrade certification system is quite complex with strengths and weaknesses regarding the empowerment of workers.

3 State of Research

3.1 State of Primary Research: The Impact of Fairtrade on hired Farm Workers – A Literature Review

About the effects of fairtrade, there is already a lot of research. The debate around the impact of fairtrade began almost when fairtrade was conducted (Howse and Trebilcock 1996). Already in the 1990s several papers about the impact of fairtrade had been published (The Fair Trade Institute 2017). Topics like “The Fair Trade Fraud” (Boyard 1992) and “TransFair coffee for a good feeling” (Korner 1994) suggest a critical examination of the fairtrade concept already in the early years, though an access to those papers is not possible anymore. Since then, a variety of research has been carried out,

approaching from different perspectives. It includes, but is not limited to, the impact of fairtrade on the buying habitude of consumers (Arnot et al. 2006), the influence of the fairtrade system on the movement of the world market, or the actual share of revenues along the supply chain (Valkila et al. 2010; Maertens 2019). Currently, only a small part of the research focuses on the effects for hired workers on the producers' farms (Krumbiegel et al. 2018). And those studies offer a wide range of results. The following literature review represents the study variety in the research field.

Positive insights on the effects of fairtrade on hired farm workers, for example, are revealed by Krumbiegel et al. (2018) and Bennett (2018). Krumbiegel et al. (2018) compare fairtrade and non-fairtrade pineapple plantations in Ghana. Using semi-structured interviews, they discover higher wages on fairtrade farms, as well as higher job satisfaction. In comparison to other sustainability standards, Bennett (2018) reveals rather positive performance of fairtrade in establishing the respective national legal wage and the possibility of collective bargaining. Rather positive results are also found by Raynolds (2022), who examines the implications of fairtrade on the Ecuadorian flower market. A mixed methods approach is used and reveals that workers empowerment increases by fairtrade certification. The implemented worker committees strengthen the negotiation power and daily life facilities get improved by the fairtrade premium. Worker committees are able to negotiate working conditions and wages. Despite of this, the ability of fairtrade to sustainably ensure the voice of workers on the management level is limited, and still organizational control from fairtrade is needed to maintain workers empowerment (ibid). In Raynolds (2014) the focus and insights are rather similar, and in Raynolds (2012) the main ability of fairtrade, which is supposed to be the requirement of worker committees, was emphasized as well.

Results are revealed to be unclear, because they are mixed in their insights on the performance of fairtrade. Ruben and van Schendel (2009) compare two nearby banana farms in Ghana, one with fairtrade certification, one without. In this specific case, the workers on fairtrade farms work fewer daily hours and therefore earn less wage. However, they benefit from more additional social payments and assets. Furthermore, due to less working hours, they earn some additional money by own agricultural production. In conclusion, fairtrade workers seem to be better off, but at the same time are confronted with higher risks due to more diverse activities (Ruben and van Schendel 2009). Another perspective is presented by Said-Allsopp and Tallontire (2014), who focus on the effect of fairtrade on women workers. The question is, if the *worker committee*, which decides how to handle the fairtrade price premium, could be more empowering for women workers than the *gender committee*, which is a second committee that is applied on the respective farms. With a qualitative approach, Said-Allsopp and Tallontire (2014) compare the abilities of those two committees. They conclude, that *worker committees* can play an important role for women "to perceive themselves as agents of change" (p. 15). Hence, fairtrade can have positive impacts on women workers

empowerment. Yet, the issue remaining is that workers still are uninformed on the possibilities they have due to the fairtrade membership (ibid). Correspondingly Siegmann et al. (2019) represent a comprehensive mixed methods research on the development of fairtrade structures over time, focussing on labour conditions and the performance of workers collective agencies. A comparison with non-fairtrade farms was conducted. They try to answer the question, how the opportunities of workers on certified farms can be improved by fairtrade. This study was conducted on tea farms in India and Sri Lanka. And likewise, Valkila and Nygren (2010) discover missing awareness of workers on fairtrade farms about their rights and options. Data was collected by semi-structured interviews in a longitudinal design. They conclude on limited abilities of fairtrade to improve working conditions, such as labour rights. Another study with rather positive results is van Rijn et al. (2020), focussing on the improvement of hired labour conditions due to certification by comparing fairtrade and non-fairtrade farms. They found a greater work satisfaction on fairtrade farms, as well as a greater worker representation. The ability of fairtrade to influence worker`s wages seems to be limited in this examination as well, though van Rijn et al. (2020) found relevant side effects, like providing workers with in-kind benefits, greater job-security, and enable private savings.

Rubio-Jovel et al. (2024) reveal that producers being certified with the fairtrade label only, pay six percent less wage to their workers than the control group and on fairtrade farms nine percent fewer efforts were made to reduce health risks for workers. However, to hold several certifications, with fairtrade being one amongst others, both health protection and wages are revealed to increase due to the certifications. They conduct a quasi-experimental design, comparing several certification schemes with a non-certified control group based on their performance to reach the *Sustainable Development Goals* (SDGs). Hence, the impact of fairtrade on hired farm workers was only one issue amongst others, examined in this study.

Makita (2012) and Besky (2014) pursue another theoretical approach, focussing on fairtrade's performance on still existing colonial relationships. Makita (2012) carries out a qualitative case study on a tea plantation in India. The focus lays on the patron-client relationship, that still exists on this plantation. Makita (2012) concludes that workers welcome fairtrade certification, as it improves their working conditions. Although, fairtrade seems not to be able to change the unequal patron-client relationship into equal work-relationships (ibid). A similar focus is chosen by Besky (2014), laying on old colonial work relationships on Darjeeling tea plantations. Even though many of them are fairtrade certified, the residents do not seem to be very concerned on ensuring democratic systems on the farms, that enable worker to improve their wages (ibid). That fairtrade "hardly affects traditional employment modalities" is also discovered by Meemken et al. (2019, p. 635). They focus on the small farm sector, differentiating between workers hired by the cooperative and workers hired by the farmers themselves. Meemken et al. (2019) conduct interviews with about 1000 respondents in total,

comparing working conditions on fairtrade and non-fairtrade cooperatives and farms. With a regression analysis, they find out that cooperative workers indeed gain from fairtrade certification, whereas farm workers do not profit.

Oya et al. (2018) present one of a view systematic reviews in the research field. They examine the effects of certifying systems on agricultural production. Among others, they examine the effects of fairtrade on the socio-economic outcomes of wage workers. Four of six studies offer negative effects of fairtrade on wages, two studies discover effect estimates around zero. Five of those six examinations are also included in the analysis of this thesis, namely Cramer et al. (2014), which includes four of those examinations and Dragusanu et al. (2014). Schuster and Maertens (2016) will not be included in the analysis of this thesis, because they focus on the effects of private standards in general, though including fairtrade, yet, not examining it separately. Cramer et al. (2014) discover a significant negative effect of fairtrade on the wages of wage workers. Research was carried out in Ethiopia on coffee and flower productions and in Uganda on coffee and tea productions, with a contrastive survey covering three places for each good in both countries on a longitudinal basis. In each case, workers on a fairtrade farm got less wage than workers on non-fairtrade farms. Cramer et al. (2017) lay the focus on the poorest rural people, which mainly are hired workers. They state, that this group is so marginalised, it even does not get mentioned in most of the studies examining the impact of fairtrade on production farms. Contrastive comparisons were conducted and the findings are, that fairtrade does not improve the lives of marginalised hired farm workers (Cramer et al. 2017). Furthermore, they discuss possible research designs in development research and conclude, that experimental designs are not possible to conduct, because of impossible randomization. However, Granville and Telford (2012) offer an experimental approach with a randomized controlled trails (RCTs) design for fairtrade and non-fairtrade wine farms in South Africa. They emphasize the difficulties and are rather self-critical with their approach. The results are ambiguous, with workers on fairtrade farms seeming to be a little better paid. And last but not least, Dragusanu et al. (2022) examine the impact of fairtrade certification in the coffee production in Costa Rica. They conduct a longitudinal panel setting, examining the coffee mill dynamics and conducting semi-structured interviews. They conclude that the income of hired workers does not increase due to the certification. Furthermore, they suggest that the ambiguous research insights, on the relation between fairtrade certification and product price, rely on the fact that most of the studies are “moderately-sized, cross-sectional comparisons”, what they tried to complement with a different approach (ibid, p. 6).

A short overview on the current state of research was provided in this chapter. It is to see that the approaches and insights on the research field are very diverse. With the help of the categorization explained in *chapter 4.3*, the differences are meant to be better comprehendible.

Table 4 - Overview of the studies included in the examination of this thesis (own figure)

Author (Publication year) – Study Title	Approach	Monetary results	Non-monetary results	Results tendency	Premium mentioned
Bennett (2018) Voluntary Sustainability Standards: A Squandered Opportunity to Improve Workers' Wages	Explorative approach Cross-sectional design, qualitative methods	Positive	None	Positive	No
Besky (2014) The Darjeeling Distinction: Labor and Justice on Fair-Trade Tea Plantations in India	Explorative approach Longitudinal design, qualitative methods	Negligible	Negative	Negative	Positive
Cramer et al. (2017) Fairtrade and Labour Markets in Ethiopia and Uganda	Descriptive approach Longitudinal design (cohort study), mixed-methods	Negative	N/a	Negative	Negative
Cramer et al. (2014) Fairtrade, Employment and Poverty Reduction in Ethiopia and Uganda	Testing approach Longitudinal design (cohort study), mixed-methods	Negative	Negative	Negative	No
Dragusanu et al. (2022) The Effects of Fair Trade Certification: Evidence from Coffee Producers in Costa Rica	Testing approach Longitudinal design (panel study), mixed-methods	Negligible	Negative	Negative	No
Granville and Telford (2012) Empirical Evidence from South Africa	Testing approach Cross-sectional design, quantitative methods	Negligible	Negative	Negative	No
Krumbiegel et al. (2018) The Role of Fairtrade Certification for Wages and Job Satisfaction of Plantation Workers	Testing approach Cross-sectional design, mixed-methods	Positive	Positive	Positive	Positive
Makita (2012) Fair Trade Certification: The Case of Tea Plantation Workers in India	Descriptive approach Cross-sectional design, qualitative methods	N/a	Positive	Positive	Positive

Author (Publication year) – Study Title	Approach	Monetary results	Non-monetary results	Results tendency	Premium mentioned
Meemken et al. (2019) Effects of Fairtrade on the livelihoods of poor rural workers	Testing approach Cross-sectional design, mixed-methods	Positive	Positive	Positive	No
Oya et al. (2018) The effectiveness of agricultural certification in developing countries: A systematic review	Testing approach Meta-analysis, mixed-methods	Negative	N/a	Negative	No
Raynolds (2022) Can certification increase trade fairness and worker empowerment? Lessons from Fairtrade International-certified plantations in Ecuador	Descriptive approach Cross-sectional design, mixed-methods	N/a	Positive	Positive	Positive
Raynolds (2014) Fairtrade, certification and labour: global and local tensions in improving conditions for agricultural workers	Explorative approach Cross-sectional design, qualitative methods	Negligible	Positive	Positive	Positive
Raynolds (2012) Fair Trade Flower: global Certification, Environmental Sustainability, and Labor Standards	Descriptive approach Cross-sectional design, qualitative methods	Negligible	Positive	Positive	Positive
van Rijn et al. (2020) Does certification improve hired labour conditions and wageworker conditions at banana plantations?	Impact evaluation Cross-sectional design, mixed-methods	Negligible	Positive	Positive	Positive
Ruben and van Schendel (2009)	Testing approach Cross-sectional design, quantitative methods	Positive	Positive	Positive	Positive

Author (Publication year) – Study Title	Approach	Monetary results	Non-monetary results	Results tendency	Premium mentioned
The impact of Fair Trade in banana plantations in Ghana: income, ownership and livelihoods of banana workers					
Rubio-Jovel et al. (2024) SDGs trade-offs associated with voluntary sustainability standards: A case study from the coffee sector in Costa Rica	Testing approach Longitudinal design (panel study), methods: n/a	Negative	Negative	Negative	No
Siegmann et al. (2019) Fairtrade certified tea in the hired labour sector in India and Sri Lanka: Impact Study and baseline data collection	Impact evaluation Cross-sectional design, qualitative methods	Negligible	Positive	Positive	Positive
Valkila and Nygren (2010) Impacts of Fair Trade certification on coffee farmers, cooperatives, and laborers in Nicaragua	Testing approach Cross-sectional design, qualitative methods	Negligible	Negligible	Negligible	Positive

The table shows all studies included in the examination of this thesis. Furthermore, the results of the studies are presented. They are split up into results linked to workers' wages and results linked to non-monetary benefits. Both parts sum up to the collum Results tendency, which shows the tendency of the results. As monetary and non-monetary results never occur contradictory to each other, the tendency equals the main direction of the results. The comprehensive explanation is to be found in sub-chapter 4.3.2.5. Additionally, it is shown if the fairtrade premium is examined in the study, and if it is found to contribute in a positive or negative way.

3.2 State of Meta-Research: The Impact of Fairtrade on hired Farm Workers

In this chapter, research field of fairtrade meta-research will be introduced. The state of research on the meta-level, regarding the impact of fairtrade on hired farmworkers, seems to be poorly developed. Just a few authors address this research gap. Besides the reviews commissioned by fairtrade itself (Nelson and Pound 2009; Darko et al. 2017; Jodrell and Kaoukji 2020), and the independent researches of Terstappen et al. (2013) and Oya et al. (2018), regarding hired farm workers exclusively there is no such review to date, and all of the authors address the impact of fairtrade on hired farm workers only fragmentary. Since the elicitation of primary data is scarce as well, this is hardly surprising. Another review is probably provided by Ruben (2013), though this document is not to be found online and a request via E-Mail was not answered.

Research commissioned by Fairtrade

Fairtrade International commissioned irregular summaries every couple of years for the last 25 years. The first literature meta-review of fairtrade examined a period from 1999 to 2009 (Nelson and Pound 2009). Thirty-three studies had been included, all of them were peer-reviewed and with the focus on the local producer level. Most of them focussing on smallholder coffee farmers in Latin America, only two studies examined the impact of fairtrade on hired workers. Exemplarily it can be mentioned that there was very little evidence on the empowerment of workers and that workers gain more from the fairtrade benefits when they know about the fairtrade system and their rights. Nevertheless, workers got provided with interest-free loans. The authors emphasize that the insights cannot easily be generalized, due to the small number of examinations, which additionally focus on several different aspects of fairtrade's impact. Nelson and Pound (2009) cite "the economic inputs of [fairtrade] that are intended to lead to positive impacts *for producers and workers*" (p. 7). The inputs named are the minimum price, the price premium, micro-finances, long-term contracts, and sales chances on export markets; it is noticeable that such intentions of fairtrade always include producers and workers to the same amount, though the interventions and the examinations of the interventions are mainly focussing on the producers. A greater focus on workers is recommended. (Nelson and Pound 2009)

The subsequent literature meta-review covers the years from 2009 to 2015 (Darko et al. 2017). Forty-five studies had been included, again most of them examining the coffee sector in Latin America and Africa. Regarding the wellbeing of workers, this review discovers that the impact of fairtrade on wages and working conditions of workers is limited. In contrast, the authors indicate indirect impacts of fairtrade being important for workers wellbeing, which are a better knowledge network and increased infrastructure stability due to the fairtrade premium, among others. Furthermore, the review reveals that the impact of fairtrade is highly dependent on the local political and economic situation, which fairtrade as hardly any influence on. Besides that, the importance of standard working conditions for small farms get addressed. This is not obvious, because of the development of the system. Fairtrade distinguishes between small farms, which are mostly family owned with farmers being self-employed, and plantations, which include hired labour situations. In the beginning of fairtrade, only small farms had been covered by the certification, but as the label grew on market share and importance, it was decided to include also plantations into the system (Murray et al. 2006). Subsequently, requirements for working conditions on plantations were developed. This review criticises that only plantations were covered. It reveals that also smallholders employ workers, which are however not considered by the fairtrade requirements. The authors assume a huge number of "invisible workers", marginalised within the fairtrade system. Therefore, standard working conditions for small farm situations would be necessary. (Darko et al. 2017)

The most recent literature meta-review covers the period from 2015 to 2020 (Jodrell and Kaoukji 2020). This research is based on the *Theory of Change* (ToC), and includes 151 studies, 30 of them focussing on hired workers directly. The ToC aims to improve the *monitoring, evaluation and learning process*, in order to discover the chain from an intervention to the respective long-term impact. The studies had been analysed with a questionnaire and rated on their quality. A study got a certain impact on the evaluation results, dependent on its rating score. The review reveals improved working conditions due to fairtrade certification, for example renumeration for overtime and adequate breaks, among others. Also, a higher job satisfaction was discovered. Fairtrade's requirements on empowerment and wages are said to be rather contributory, and again the circumstances in the local settings are crucial. Hired workers in the small farming context are still not focussed on by fairtrade. (Jodrell and Kaoukji 2020)

Further, independent Research

The systematic review of Oya et al. (2018) was already discussed in *chapter 3.1*. The authors focus on certification systems in general, conducting a theory-based, mixed-methods approach. The first part includes the analysis of effect size estimates of quantitative impact studies. A second part focusses on a thematic synthesis to approach the respective qualitative literature. Regarding the wages of hired farm workers they reveal negative, but no significant impacts of fairtrade certification. (Oya et al. 2018)

Another review was published by Terstappen et al. (2013), who did not aim to draw new insights on the impact of fairtrade, rather present the literature scope at that time. Studies were analysed with the help of a data matrix, to examine the methods and key findings. The authors focussed on the social dimension of fairtrade's impact, including the topics *gender*, *health*, *labour* and *equality*. At that time the imbalance of qualitative and quantitative approaches was obvious, with 40 qualitative studies, 33 studies with mixed-methods and only 4 quantitative studies included. Regarding the situations of hired workers, Terstappen et al. (2013) find that fairtrade farmer organizations do not realize the concerns of workers, and therefore do not improve their situations. Due to the authors this might be because workers are not be considered in the fairtrade standards.

This chapter represents the state of the meta-research in the fairtrade research field that addresses hired farm workers in some extent. As to see, the meta-research on fairtrade is not quite comprehensive. However, from the existing studies it can be concluded that hired farm workers were and still are neither in the focus of fairtrade nor in the focus of fairtrade meta-research.

3.3 Digression: Assessment of Fairtrade's performance in general

To gain insights on the general performance of fairtrade is not the actual aim of this thesis. Since the research field is still small and many studies are of questionable quality, meta-analyses do not offer

reliable results. Nevertheless, it is an obvious interest to ask for the performance of fairtrade. The following paragraph provides some of the insights discussed in the current debate.

It is crucial to differentiate the target of the impact of fairtrade. When examining the product price farmers receive for one sold product, many studies confirm higher product prices for fairtrade products (Méndez et al. 2010). However, fairtrade certified farmers are mainly not able to sell all their certified products on the fairtrade market. Papers reveal amounts from 11 % to 60 % that can be sold on fairtrade distribution channels (Dragusanu et al. 2022; Krumbiegel et al. 2018; Valkila and Nygren 2010). Therefore, farmers do not always actually gain a net benefit due to the certification, in some cases certification cost even seem to exceed revenues (Dragusanu et al. 2022).

When examining the impact of fairtrade certification, it can also be focussed on worker's wages. Some authors reveal increasing wages (cf. *Table 4*, for example Ruben and van Schendel (2009) and Bennett (2018)), but the main part of studies in the research field does not reveal any effects on wages, neither positive nor negative. This is hardly surprising, since the fairtrade standards do not include precise requirements for worker's wages (Fairtrade International 2019). It is a challenge of the system to set up decent requirements for workers, while still staying achievable for farmers. Only if farmers decide to get certified, the requirements can have an impact at all; a tension that cannot be solved easily. What many studies though reveal are non-monetary benefits for workers, and also for farmers. For example, the requirement of a worker committee on a certified farm is meant to improve the bargaining power of the workers, as well as acting as mediator between employer and employees (Raynolds 2022). Furthermore, farmers gain from knowledge exchange in fairtrade farmer networks, and the fairtrade premium is in many cases used to improve local infrastructure. Also, health care requirements are set up, like maternity leave and sick days. Such development in correlation with fairtrade certification are to find by many examinations (Raynolds 2022; Siegmann et al. 2019; Krumbiegel et al. 2018). Furthermore, an ongoing debate exists on the question if development on certified farms is caused by fairtrade or just correlated to it. Longitudinal studies with pairwise matching methods try to approach this topic (Dragusanu et al. 2022).

On the consumer level it is to keep in mind that higher prices at the point of sale do not necessarily lead to an equal rise in product prices for farmers. Valkila et al. (2010), among others, found a high price volatility on the consumer side, so that retailers could increase the farmers fairtrade price to increase their own profits. Therefore, the importing country would retain a remarkable share of the product price.

And the possibilities of fairtrade, to empower small, marginalised farms decreased when it was decided in 2015 to include plantations into the certification system. The sudden increase of the fairtrade market size, though now reaching a wider consumer market, favours greater companies whereas for small-

farms it is more difficult to sustain (Davenport and Low 2015). As to see, it is not possible to draw distinct conclusions on the impact of fairtrade. The debate is very complex and so have to be the possible improvements.

3.4 Scientific Standards in Fairtrade Research

While the subsequent chapter introduced the existing meta-research on the impact of fairtrade on hired farm workers, this chapter addresses the quality of fairtrade research. It provides the theoretical basis to approach research objective O3: *Conclude on the possibility of generalization of the studies´ results,* and the research question *RQ4*: *What limitations constrain the generalizability?* Fairtrade research is clustered in four different study types (*Table 5*). Each is discussed on its possibility to generalize from individual results to conclusions on the general performance of fairtrade. The main emphasize is laying on the qualitative meta-research, since this is the research approach of this thesis.

Table 5 - Methodological Approaches in Fairtrade Research, with some Examples (own figure)

		Methods	
		Qualitative Approach	Quantitative Approach
Research level	Primary research	**Qualitative primary research** E.g. Besky (2014) Impact evaluations possible	**Quantitative primary research** E.g. Granville and Telford (2012) Impact evaluations possible
	Meta research	**Qualitative meta-research** N/a	**Quantitative meta-research** E.g. Oya et al. (2018)

This table represents a broad cluster on the methodological approaches in the research field. Primary research is for example provided by Besky (2014) and Granville and Telford (2012), the former with a qualitative approach, the latter with a quantitative one. In primary research also impact evaluations are carried out. Meta-research is for example provided by Oya et al. (2018) with a quantitative approach. An example for qualitative meta-research could not be found yet, and this thesis shall contribute to fill out this gap.

Insights into the impacts of fairtrade are achieved by primary and meta research with qualitative and/ or quantitative methods. Each approach aims to deepen and to further mark out the understanding of fairtrade's impact.

An additional approach to examine the impacts of fairtrade on hired farm workers are *impact evaluations*. This term commonly refers to an examination that has been commissioned by a stakeholder to answer a specific, reality-relevant question. Impact evaluations can take place as primary research, and can be qualitative or quantitative.

3.4.1 Primary Research

Qualitative and quantitative research both contribute to a better understanding of the effects, though with different approaches. Quantitative research has the aim to discover the strength of an effect. It can be judge by the strength of internal and external validity, who contribute to the reproducibility of

a study. An example is the study of Granville and Telford (2012), who conduct a RCTs approach on the impact of fairtrade. A main critique on quantitative research, especially on randomized controlled trails (RCTs), in empirical research is represented by Basu (2013), who emphasize the fact that significant quantitative results do not automatically reflect causality. RCTs are able to describe a certain effect in a certain situation, but one should be careful to conclude on causality too fast (Basu 2013; Deaton 2010). This is because context analyses are rare in quantitative approaches, and without a decent understanding of context variables causality might be concluded falsely.

Qualitative research includes the context of the dependent and independent variables, and aims to understand side-effects or barriers and facilitators of the effects. It is eligible in order to provide context to quantitative results, who otherwise cannot be interpreted usefully (Darko et al. 2017). Qualitative research cannot be that easily assessed since the qualitative methods are much less standardized. An approach to ensure quality in qualitative research is the method of triangulation, which aims to validate a fact by verifying it from different perspectives (Seale 1999). An example in fairtrade research is Raynolds (2022), who triangulated data by aligning statements of workers with those from management and governmental representatives. Qualitative research is criticised to expound on the context so comprehensively that it is not able to provide any transferable lessons (Pawson 2002a). An example is Besky (2014) with detailed, rather subjective, description of tea plantations.

3.4.2 Meta Research and Theory Building

The existing meta-research on the impact of fairtrade on hired farm workers is represented in *chapter 3.2*. All presented studies reveal poor impact of fairtrade on hired farm workers and it was shown that there is no research focussing on hired farm workers exclusively. Therefore, a discussion is necessary on the feasibility and necessity of meta-research in this specific field.

Quantitative Meta-Research

There is no quantitative meta-study with the focus on hired farm workers published yet. The quantitative studies in primary research are scarce and often are poor in reporting, so that information necessary to conduct a meta-analysis are hard to gather (Cramer et al. 2014). It will be an enrichment for the research field when a meta-analysis takes place, but at first more quantitative primary research with comprehensive reporting is necessary. Therefore, in the following it will be focussed on qualitative meta-research.

Qualitative Meta-Research

Most of the examinations in fairtrade research are qualitative, and there is an ongoing debate whether such data can be summarized and generalized by a systematic research (Denzin and Lincoln 2018). Opponents observe the problem of decontextualization (ibid), since the appliance of the methods and

the results depend a lot on the specific situation, and an abstraction of the conclusion might therefore not be suitable. Additionally, a debate exists on whether it is reasonable to judge on the quality of qualitative research at all (Seale 1999). Nevertheless, the demand on systematic reviews of qualitative data can be perceived; however Thomas and Harden (2008) emphasize the thought of conducting qualitative meta-research while keeping its weaknesses in mind. Additionally, Paré et al. (2015) developed a typology for reviews, stating that any kind of primary data can be summarized in a review.

Simultaneously, the debate on qualitative meta-research, it is conducted in many cases. Pawson (2002a) refers to Towner et al. (1996)[1], who introduced an approach on qualitative studies based on an analysis template with certain criteria to be filled out for each study. Pawson (2002a) concludes that this approach on the one hand is very insightful, because it provides many details on the studies and provides a good overview. On the other hand, it fails to be a comprehensive review. Because of the many details it provides, there is no common basis for a comparison. Therefore, Thomas and Harden (2008) improved the approach of Towner et al. (1996): After filling out a criteria template in the wording of the certain study, an additional step is made to create *analytical themes*, which include an interpretation and unification of the original wording by the reviewer, in order to "generate new interpretive constructs, explanations or hypotheses" (Thomas and Harden 2008, p. 1). As to see, qualitative meta-reviews are conducted already and would be useful in fairtrade, since there is many qualitative research.

Pawson (2002b) introduces *realist synthesis*, which is meant to overcome the weaknesses of meta-analysis (no consideration of the context, comparison not useful) and narrative review (too much context, no comparison possible) by focussing on theory building. The role of theory will be expanded in the following.

The Role of Theory

The importance of theoretical background in empirical research is emphasized by several authors (Deaton 2010; Pawson 2002a).

Pawson (2002b) introduces theory building as a step-by-step progress by analysing the primary studies. Each study should be considered in its context in which the study took place, and the respective outcomes. In each case, so to say small part-theories are built up, which will then sum up to a main theory, with the several part-theories included. All together a comprehensive picture should be revealed which shows the possibilities and constraints of an intervention. However, Pawson (2002b) does not answer the question how detailed those part-theories should be. If they are focussed on too

[1] The reference cannot be found online anymore.

small details of a certain context, they might not fit a main theory. No clue is given on how to decide on the scope of those part-theories.

What Pawson (2002b) is furthermore emphasising on, is the theoretical approach of understanding the intervention as object, not subject. Normally, it is checked if the intervention "works", however the author argues, that it is not the intervention "working", but the subjects who interact with the intervention, and which are enabled by the intervention on a certain degree (p. 345). This degree depends on the context variables, and those should be examined by theory building. (Pawson 2002b)

Additionally, Deaton (2010), who is examining development research, emphasizes the importance to understand the underlying mechanisms of an intervention, rather than to focus on the study results. The mechanisms would reveal the contexts in which an intervention works (ibid). Deaton refers to the field of behavioural economics, in which the abstraction of underlying mechanisms beyond the certain situation is already practiced (ibid).

This sub-chapter represents different approaches of meta-research and their abilities for research on hired farm workers. The insights of Paré et al. (2015), Thomas and Harden (2008) and Pawson (2002a) and Pawson (2002b) are used to derive the research approach for this thesis. The general research procedure is explained in *chapter 4*, referring to those authors.

3.4.3 Impact Evaluations

In this sub-chapter, impact evaluations in the fairtrade research are introduced. Their strengths and weaknesses are discussed, as well as their relevance in the research field.

In empirical science, *impact evaluation* is a fixed term distinguished from "pure" scientific research. Impact evaluations are meant to assess interventions, in order to give policy recommendations on them. For policy makers it is important to know if an intervention works and causes the expected results, or if different effects appear, in the worst-case effects that do more harm than good (Glewwe 2020). Impact evaluations therefore are commissioned by policy makers or other stakeholders, who are interested in a specific question regarding the performance of an intervention. Impact evaluations are not fixed in their approaches, they are used with several methods for data collection, and sometimes are criticised to lack on quality. A reason might be that the research motivation does not result on basis of the current research state in order to fill a scientific research gap, but is meant to reveal insights relevant for immediate actions in relation to the research object (Stockmann 2019). It is therefore criticised often to not be embedded in the scientific context and therefore to miss a theoretical basis. An example would be the work of CEval (2012). The *Centre of Evaluation (CEval)* conducted one of the first impact evaluations on fairtrade's impacts. Their report did not include any discussion on the scientific context, yes, they did refer to a few external publications, but not even a reference list is provided. Of course, impact evaluations are recommended to meet scientific quality

standards, and, in fact, several authors elaborate that and provide instructions (Glewwe 2020; Stockmann and Meyer 2014). Another challenge comes with the aim of addressing real-life situations. The task of impact evaluations is to conclude the research with distinct recommendations. Therefore, research results must be simplified, which leads to the risk of overinterpreting the results (Glewwe 2020). This stands in contrast to scientific research, which often reveals ambiguous results, without the need of drawing distinct conclusion. Furthermore, since impact evaluations are mostly commissioned, and financed, by a stakeholder of the intervention, they are criticised to be biased in their results.

Several impact evaluations have been commissioned by fairtrade. In CEval (2012), six estates have been examined, of different size and production and in four different countries. The case studies were designed quasi-experimental, with a target group and a matching comparison group for each case. Regarding workers welfare, the evaluation revealed in some cases positive effects of certification of health and safety issues, of the existence of written contracts, access to micro-credits, and educational programmes. Though, no improvement on the wages was observed (CEval 2012). Other impact evaluations, examining the impact of fairtrade in hired farm workers, are for example Siegmann et al. (2019) and van Rijn et al. (2020), who are included in the analysis of this thesis.

Those impact evaluations provide insights for fairtrade on its own effects. Though, as shown above, they come along with certain weaknesses. For a comprehensive insight on a research field, proper scientific research should be included in the impact considerations. For example, the approach of impact evaluations could be embedded into the scientific field and conducted with scientific standards more thoroughly. However, stakeholder commissioned research plays a role in such a small research field since the independent financing possibilities are limited.

3.4.4 Generalizability in Field Research

The previous paragraphs referred to the different scientific approaches and their quality, within the research field of fairtrade. What they have in common is the desire to understand cause-effect-chains, and eventually provide generalizable insights to foster intervention development. Especially impact evaluations are confronted with the need of easily accessible answers, comprehensible for non-scientific stakeholders. Meanwhile the “pure” social science is not so much driven by the aim of generalization, nevertheless researchers are keen to understand cause-effect chains. The impact of something on something else is the core focus of research. The desire to understand cause-effect-chains is linked to the attempt to generalize the insights of one examination to other situations. The matter of generalizability is discussed within the important quality criteria of reproducibility. Of course, one single study result is not to be generalized and applied as such to all other similar situations. But if the study can be reproduced in other situations and the results remain the same, a certain number

of studies is needed to draw conclusion for some general hypothesis, regarding the respective research object. Generalizability is one core goal of science.

Generalizability and Cause-Effect-Chains depend on internal and external Validity

However, it depends on the external validity of a study, to which extent it can contribute to generalization. Additionally, it depends on the internal validity of a study, namely to which extent it can reveal the underlying cause-effect-chains. Internal validity of a study is strong when the effects of an intervention can be causally related to the same. This is the case when the research environment can be controlled by the researcher as much as possible, which is important when conducting quantitative methods. It is improbable to fully meet this requirement in field research, which is why those methods have disadvantages and need to be complemented by qualitative approaches. External validity addresses the quality of study results in terms of their possibility to be plausible in other situations as well. A challenge is to extract the interaction between context-specific variables and the intervention. For this reason, qualitative approaches can be suitable as they offer a more comprehensive understanding of the complexity of a research field than quantitative ones. The less context-specific distortions take place, the higher is the external validity. Drawing conclusions on the external validity is crucial, if an evaluator gives recommendations, whether an intervention should be applied in other contexts as well. External and internal validity are suspected to trade-off against each other but do not necessarily do so. (Glewwe 2020)

This contributed connection between qualitative and quantitative studies is a strong argument for mixed method approaches and possibly also for combined meta-research of qualitative and quantitative methods.

Do internal and external Validity compulsively lead to Generalizability?

Without doubt, the reliability of study results is dependent on the internal and external validity. However, they are not for themselves ensuring generalizability. Supposing, the researchers conduct a study with high internal validity. They eliminate all relevant distortion factors and an intervention can be indeed shown to cause respective effects. This result is interesting in the certain situation, but it cannot be transmitted to any other situation. The results and the high internal validity count for this certain research situation only (Glewwe 2020). If the researchers aimed to reproduce the study in other contexts, external validity would have to be increased as well. Therefore, distortion factors do not have to be eliminated, but must be defined and considered in the approach. For example, when examining the effects of fairtrade on hired farm workers this effect might be biased by different employment situations. Full-time employment might have a different side-effect, than temporary casual employment. For internal validity, the employment status could be unified in order to eliminate the effect. However, in field research it will mostly not be possible to change the employment status of the workers. Rather, in the study design different employment conditions could be considered by

comparing worker groups with different employment statuses. Eventually, the distortion variable *employment status* would not be eliminated but included in the study approach. And the more side-effects are considered, the higher is the external validity (Glewwe 2020). And with many side-effects included, the study can be reproduced to several situations that meet the included effects. It is therefore shown that high external validity contributes to the reproducibility of a study. However, reproducibility does not lead to a useful generalization necessarily. The repetition of a study only proofs the extent of external validity of the study approach. But in order to draw generalizable statements of the examined intervention, conclusions out of the underlying mechanisms should be made. An empirical study could be followed by a theory building review.

Another though is provided by Basu (2013), who argues that generalizability can never be provided by empirical research, since the research is always examining past events. It is part of the human behaviour of intuition that research insights are transferred to the future, for example by providing policy recommendations. It is based on the experience that a variable causes the same effect repeatedly, though actually, if this were still the case tomorrow, we do not know today. While this thought is true in theory, it can be hypothesized that relying on that never-exactly-certain generalizability in practice lead people to deeper knowledge than drawing no conclusions at all.

3.4.5 Assessing the Studies´ Quality in this qualitative Meta-Research

It is a challenge to evaluate the quality of studies in qualitative meta-research. This is, at first, when qualitative as well as quantitative primary research is included. Both kinds follow different approaches and are therefore hardly to be assessed on the same quality criteria. Secondly, it is challenging to judge on the design and methods because it is not possible to discover the situation, in which the study had been conducted, again, to evaluate if the design and methods were appropriate. The only clue is the study itself, subjectively reported by the researchers, who most likely argue how well suited their design and methods are. Third, the reporting of the studies is differently comprehensive. Some studies describe their approach extensively, others only mention key points. Therefore, it is required to, on the one hand, balance the quality criteria to be not too detailed, in order to avoid empty spaces in the data collection when studies do not offer so many details. On the other hand, the quality criteria should not be too superficial, because no insights would be provided. For example, it is examined how the researchers handle *divergent results (5.1.3)*. It could be required in the analysis to know how many divergent results existed and if the researchers adjust their approach because of this. But that information is not stated in most of the cases. It would also be possible to not examine this point at all, but this would not give any insight. The quality criteria, that is included in the catalogue, is focussed on the reporting: *Is it* mentioned*, how divergent results had been handled?*

With these three challenges in mind, the approach described in the example is applied for several of the quality criteria: Instead of examining the quality of the study, it is revealed if the authors themselves discuss those criteria. The approach is that if authors discuss quality criteria, this is a quality criterion for itself. It is actually of great worth and desirable to get a comprehensive discussion on the quality provided, especially when the quality of the approach is improvable. Such a discussion enables the reader to weigh the results and draw own conclusion of it. *Table 6* provides an overview on the quality criteria used in this research. The terms *internal* and *external validity* are used for all studies, quantitative and qualitative ones. While they would otherwise refer to quantitative research approaches, they are used for this thesis in a very broad sense, explained in *chapter 3.4.4*. Internal validity refers to the internal congruency of a study, whereas external validity refers to the reproducibility of the study and its results in different contexts. Some of the catalogue criteria can be assigned to internal and/ or external validity. Some refer to the quality assessment on basis of the reporting within the study. An overview is presented in *Table 6*, which is referred to in the discussion of the quality in *chapter 6*.

Table 6 - Studie's Quality and the referring Catalogue Criteria

	Quality of Reporting	Quality of Approach	
	reproducible due to transparency	Internal Validity	External Validity
0.3 Financier linked to fairtrade		x	
1.2 Theory mentioned			x
1.3.1 Hypothesis mentioned		x	
1.3.2 Hypothesis fit research question		x	
2.4.2 Worker randomization		x	x
2.6.1 Persons examined additionally		x	
2.6.2 interested in positive sector reputation		x	
2.7 Place of data collection		x	
3.2 Indicators explained	x		
3.3 Methods fit research question		x	
4.5 Implications fit questioning	x		
4.6 Useful in terms of impact evaluation		x	x
5.1.1 Quality criteria addressed	x	x	x
5.1.2 Confounding variables	x		
5.1.3 Divergent results handling	x		
5.1.4 Reflection of internal and external validity	x		
5.2 Explanation on results			x
5.3 Reasonability		x	
5.4 Peer-reviewing		x	
5.5 Practical relevance stated			x
5.6 Transparency	x		

The quality of the studies can be assessed by examining the approach of the study and by examining the reporting of the approach. Some of the criteria can be assign to the one or the other. The criteria, contributing to the quality of the study approach, are separated in the ones contributing to internal validity and the ones contributing to external validity. Some of the criteria can be assign to more than one quality assessment.

4 Methods

In this chapter, the methods used in this thesis are represented. The meta-research is located in the general meta-research field, on basis of Paré et al. (2015), and the descriptive approach is explained (*chapter 4.1*). The characteristics of this approach are based on King and He (2006), which at the same time provides the structure for the following procedure. The literature research, explained in *chapter 4.2*, is inspired by the approach of Oya et al. (2018). The development of the catalogue of criteria is explained in *chapter 4.3*. The criteria and the approach are based on Thomas and Harden (2008), among others. Eventually, in *chapter 4.4* the editing of the data is explained.

4.1 Descriptive Review: The Study's Frame

Literature reviews are used frequently in order to represent an overview on a certain scientific field. There are several kinds of reviews, who differ in their strengths and weaknesses. To speak with Guzzo et al. (1987): "All forms of literature review have the potential to make a clear statement about what is *not* known" (p.29). To identify what exactly is known already by using a review depends on the quality of primary data and the explicit research questions (Snilstveit et al. 2012). Paré et al. (2015) developed a typology for reviews, which categorizes approaches for any kind of primary data. It is the same in primary research as it is in meta-research. When there is only a small body of studies, the research starts with qualitative approaches to discover the structure and scope of the field. If decent qualitative research had been conducted, hypothesis can be derived from it and can be tested with quantitative methods. Cooper et al. (2009) and also Borenstein et al. (2013) point out, that the differences between the respective reviews are not always clear cut, neither are the names scientifically consistent. Nevertheless, a rough classification is possible and was derived from different sources for the use within this thesis (see *Figure 1*).

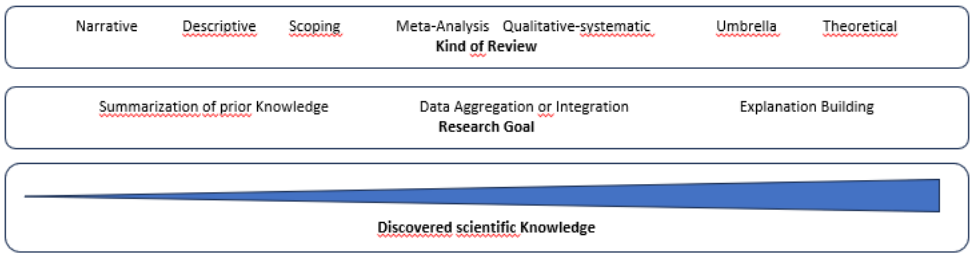

Figure 1 – Locating the study design in the meta-research field (own figure, cf. Paré et al. (2015) and Snilstveit et al. (2012)). Different kinds of reviews are presented. The narrative, descriptive, scoping reviews with the aim to summarize prior knowledge. The meta-analysis and qualitative-systematic review with the aim to aggregate or integrate data in order to produce new insights. And the umbrella review and theoretical review, with the aim of developing theories and contribute to explanation building. Therefore, the discovered scientific knowledge is rising from narrative to theoretical reviews.

As mentioned already, the *discovered scientific knowledge* on the impact of fairtrade on hired farm workers is rather small, mainly providing qualitative data or relatively poor-quality quantitative approaches. The aim is to provide representative and structured insights on the research field.

Therefore, the descriptive review will be used as approach for this research. King and He (2006) describe the following characteristics of a descriptive review:

1. It is **based on a systematic search**, which aims to be comprehensive,
2. and is therefore **representative for the state of research**.
3. The studies get **coded on specific criteria.**
4. A **frequency analysis can be conducted.**
5. It is possible to **discover distinct patterns**.

The characteristics point to a suitable working procedure, which is used for this thesis. Likewise, Oya et al. (2018) who already conducted a meta study in the research field of fairtrade, used a similar structure. The first step is literature research, addressing the characteristics one and two and is explained in *chapter 4.2*. The second step is the development of a catalogue of criteria, including the third characteristic. It is explained in *chapter 4.3*. The fourth and the fifth characteristic contribute to the part of results, which follows from the first two steps, and is represented in *chapter 5*. It is to specify, the term *distinct patterns* for this thesis is addressed with the term *correlations* in research question *RQ2*. With this insights, the focus for further research can be suggested, which is another goal of conducting a review (Guzzo et al. 1987).

There are some advantages of the descriptive review. It is systematically and provides a reliable overview on a topic. Though it is not narrowed down to numbers, as a meta-analysis does, but is able to include insights on underlaying concepts in a field and meanings between the lines. However, the descriptive review is weak in its possibilities to examine the effect strength of an intervention. Therefore, it is not the aim to close this thesis with distinct conclusions on the effectiveness of fairtrade on hired farm workers. This conclusion could be drawn by conducting a meta-analysis, since it is able to present the effect size estimates (Borenstein et al. 2013). But comprehensive primary data are necessary to conduct a meta-analysis, yet, are not sufficiently available for the research of this thesis. Moreover, to comprehensively understand the research field, more qualitative research like a descriptive review is required.

4.2 Literature Research

The literature research is in its approach based on Oya et al. (2018).

4.2.1 Research Scope

Since the research body on fairtrade research is huge, it was feasible for this thesis to narrow down the scope on examinations regarding hired farm workers. The studies must address fairtrade and hired farm workers directly. Studies that additionally address other certification systems than fairtrade only were included when the examination on fairtrade could clearly be distinguished. Studies that

additionally address other groups of persons than hired farm workers only were included when the examination on hired farm workers could clearly be distinguished. The term *hired farm workers* refers to all kinds of employment in the agricultural sector. Almost all commodities in all countries were included, except of fairtrade certified handicraft products. All kinds of publications were included. Published articles, either peer-reviewed or not, as well as grey literature, which included reports, impact evaluations and book chapters, among others. The requirement is that the articles are freely available and presented an examination approach with scientific methods. To include all kinds of literature aims to avoid publication bias (Wagner and Weiß 2022). Research was limited to German and Englisch literature, due to the abilities of the author. The year of publishing did not matter, though no earlier research than 2003 was available.

4.2.2 Literature Search

Literature research for this thesis started by preparing the proposal in January and the main research took place in March 2024. The search engines *Google Scholar*, *JustFind* and the AI-based search engine *Elicit* were used. Some data bases were used directly for search, namely *Forum Fairer Handel* (2024), *The Fair Trade Institute* (2017), the *Journal of Fair Trade* (ScienceOpen 2024), and the *International Journal of Rural Development* (Rural21 2024). The search terms that were used were “impact of fairtrade”, ”fairtrade effects” and “effectiveness of fairtrade”, as well as the search string “fairtrade AND worker* OR labo?r OR wage*”, among others. The first step was to choose articles by scanning the abstracts. Articles were included when an examination of the fairtrade impact on hired farm workers could be suggested. Additionally, the theoretical part and the references of matching studies were scanned to find further literature. In the beginning 26 studies were collected and it was expected to find 10 to 15 more. The second step was, screening the articles more thoroughly and keeping them included, when agricultural workers in the fairtrade context were addressed directly and distinguishably. It turned out that several studies did not match these criteria or were not freely accessible, moreover no additional studies could be found. For that reason, the first sample consists of 20 studies. During the analysis, two more had to be sorted out due to their focus. During the thorough analysis of the studies, some more references were still found, which then stayed excluded for time reasons. Studies excluded are exemplarily shown in *Table 7*, with the respective reasons.

Table 7 - Examples of Studies not included into the Research, with the respective Reasons.

Author (Publication year)	Title	Reason of Excluding
Frundt (2009)	Fair Bananas: Farmers, Workers, and Consumers Strive to Change an Industry	No access
Jaffee (2009)	'Better, but not great': the social and environmental benefits and limitations of Fair Trade for indigenous coffee producers in Oaxaca, Mexico	Hired workers not in focus, neither directly interviewed
Nelson and Martin (2013)	Assessing the poverty Impacts of sustainability Standards	No distinct differentiation between *Fairtrade International* and *Rainforest Alliance*
LeBaron (2018)	The Global Business of Forced Labour: Report of Findings	Time constraints

4.3 Criteria Coding

The second research objective of this thesis is to "gain a clear picture of the studies" (cf. *chapter 1.3*). This objective is approach with the help of a catalogue of criteria, which is explained in the following section. The approach can be visualized as follows:

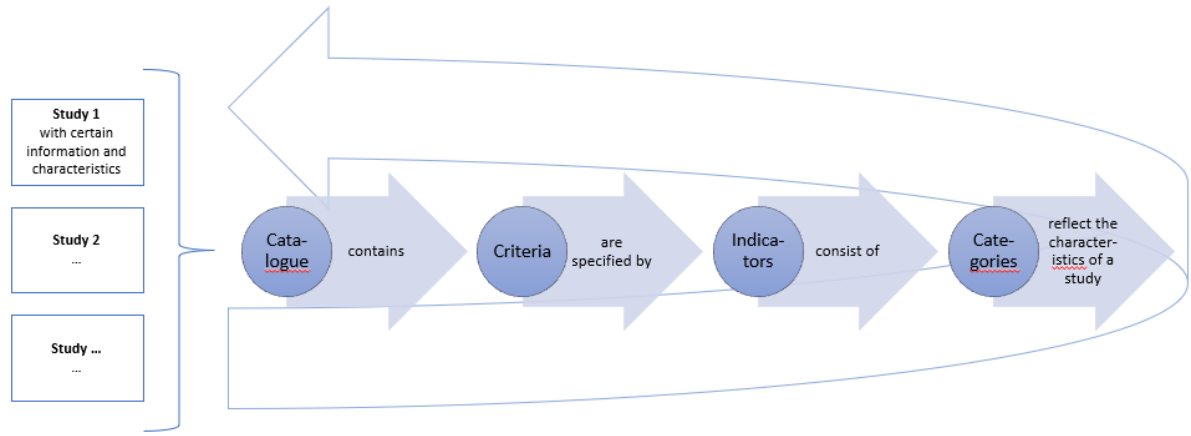

Figure 2 - Visualization of the Research Approach (own figure)

Each study contains certain information and characteristics. These shall be unified and sorted to compare them with the information and characteristics of other studies. Therefore, a catalogue is used. The criteria in this catalogue will be specified by indicators. And those indicators are specified by a range of categories. The categories in turn reflect the characteristics of the studies in a unified and comparable way.

4.3.1 Development of the Catalogue of criteria

The first draft of the catalogue of criteria was developed on the basis of the book *Handbuch Methoden der empirischen Sozialforschung* (Baur and Blasius 2022). It contains a comprehensive description on the conduct and characteristics of qualitative and quantitative empirical research. The chapters of Przyborski and Wohlrab-Sahr (2022), Stein (2022), Flick (2022) and Krebs and Menold (2022) were particularly used and will be referred to in the following explanation. This first draft was then revised and completed during the process of the analysis. The emphasis of Deaton (2010) on a solid theoretical foundation of studies in the development research were included, as well as the thoughts of Thomas and Harden (2008) and Pawson (2002a). The latter to deal with the challenge of comparing qualitative and quantitative designs on basis of the same criteria. Furthermore, Jodrell and Kaoukji (2020) developed a "study checklist" in order to score on the quality of fairtrade impact studies, which was used to confirm and expand the catalogue of criteria of this thesis. Eventually, some of the criteria had been developed by the reviewer to meet the specific characteristics of the studies of this research.

Exemplary Difficulties of the Catalogue Development

To develop the catalogue comes along with some challenges. First of all, it is difficult to categorize qualitative and quantitative studies with the same criteria. Przyborski and Wohlrab-Sahr (2022) consider that qualitative research does not always follow clear guidelines and standards. Therefore, there is a tendency to approach qualitative research with a sort of trimmed quantitative, standardized methods (ibid). Such an approach leads to poor quality of the results since qualitative research is not able to fulfil the requirements of standardization (ibid). The same problem emerges when trying to unify the characteristics of qualitative and quantitative nature. For example, whereas a quantitative interview contains standardized questions and is applied on a predetermined number of participants in a predetermined setting, qualitative approaches can contain interviews with open questions, can take place spontaneously and participants can be included continuously. However, there are also many criteria fitting both types of research, such as the *main research object (0.6)* or the *number of surveys (2.3.2)*. Therefore, some criteria can be applied easily, for some other criteria a suitable application in the multifaceted research field had to be discussed.

Another difficulty was to unify the wording of the studies. Sometimes, words are used in different contexts with different meanings. For example, some researchers state their studies to be *case studies*, when in fact they examined several objects and compared them. In the definition of this research, such "case studies" would be categorized as *comparative studies*. When confronted with this problem, the

analysis of the studies had to be revised until the categories were unified.

4.3.2 Explanation of the Catalogue Criteria

The aim of this catalogue of criteria is to describe the studies in their conduct, summarize them on their characteristics and compare them. As represented in *Figure 2*, the characteristics are addressed by the criteria. Each criterion was defined by an indicator or an indicator question and either can be answered openly or has some possible categories to answer. There are six main parts with several sub-criteria. *Figure* 3 represents the main parts, with the sub-parts of *Design (2)* as an example.

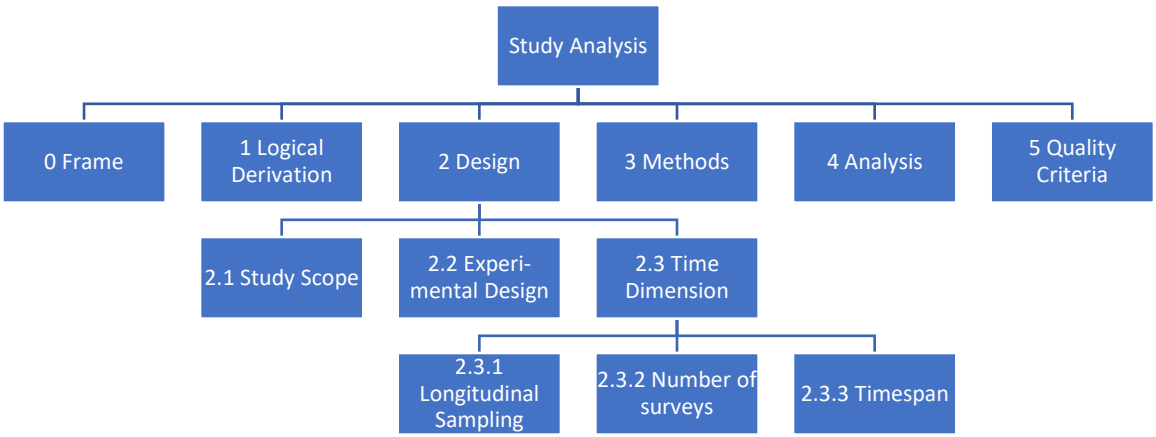

Figure 3 - Overview of the six main catalogue parts, with an exemplary section of the sub-criteria in the part Design *(2) (own figure).*

The six main parts are Frame, Logical Derivation, Design, Methods, Analysis and Quality Criteria. Those main parts contain several sub-criteria. For example, the part Design contains the sub-criteria Study Scope, Experimental Design and Time Dimension. The criteria Time Dimension is again divided in three sub-criteria, namely Longitudinal Sampling, Number of Surveys and Timespan. All those criteria contain different categories with which the information for each study can be collected. The complete overview is to find in Appendix 1.

This numbering design was chosen for allowing to include further criteria in the right place during the analysis. Though, the differentiation between the parts is not clear cut, for example part *0 Frame* includes the criterion *main research object (0.6)*, which could also be included into part *2 Design*. For that reason, the parts themselves do not have a huge expressiveness, but mainly serve practicability. The information of the catalogue is collected for each study in order to compare and analyse them afterwards.

4.3.2.1 Part 0: The Study's Frame

The first main part *Frame (0)* presents information on the frame of the study. The criteria are partly developed by the reviewer herself and partly based on the criteria provided by Pawson (2002a).

At first, the main publication information is noted, namely the *main author (0.1.1)*, *joint authors (0.1.2), publication year (0.1.3)* and the study's *title (0.1.4)*. Furthermore, since every researcher has an individual *scientific background (0.2)* this background is researched and categorized. The indicator is the current occupation or department description and the categories are *economics, sociology* and *interdisciplinary*. A researcher working at a department of economics or finance, or something similar, would be categorized *economics*, a researcher working at a department of development studies or sociology or international affairs, or something similar, would be categorized *sociology*, and if the description is not clear cut or differs between the researchers the category *interdisciplinary* would be chosen. The scientific background might influence the research approaches and result interpretations (Weischer 2007). It can be therefore analysed, if the researchers background correlates with certain study designs or even with certain study results.

The next characteristic refers to the *financiers (No. 0.3)*. It is checked, what kind of relationship towards the fairtrade system exists. If the financiers are close fairtrade stakeholders, or they are even the clients of the study, they might have a financial interest in certain study results which would possibly produce a bias. The indicator is defined by the question: *Is the financier directly linked to fairtrade?* It can be answered with the categories *yes, no* and *n/a* (no answer). The question is answered with *yes*, when the financier is part of fairtrade directly or is using money from *Fairtrade International*. The question is answered with *no*, when there is no obvious connection. And it is answered with *n/a* when no financier is mentioned. It can be analysed, if the relationship between the financier and fairtrade correlates with certain study results. If the study results are rather positive when the financier is a fairtrade-supporting stakeholder, then a bias would be supposed.

Subsequently, the *countries (0.4.1)* are collected and summarized based on the *continents (0.4.2)* they are located on. The names of the countries and continents are noted. The *product (0.5)*, which is produced by the observed farms, is noted as well as the *main research objective (0.6)*. Both criteria are first collected in open text form and secondly summarized with fitting keywords to standardize the answers. Criteria 0.4.1 to 0.6 are collected for the description of the research field, regarding the allocation of the studies: how many studies took place in which countries, which food sub-sectors and with which foci.

4.3.2.2 Part 1: The logical Derivation of the Research Approach

The second main part is called *Logical Derivation (1)* and focusses on the derivation of the research approach. It is to ask, if the approach is consistent, for only a reasonable approach can provide reliable

results. Criteria are derived on the basis of Deaton (2010), Jodrell and Kaoukji (2020), Stein (2022), and Bortz and Döring (2006).

The *problem background (1.1)* is mostly part of the introduction of each study and contains a discussion on the importance of the research. It is approached by an open text analysis to extract keywords. Those keywords are sorted by themes and graphically presented by a word cloud. The keywords of the different studies might be quite similar as they all focus on the same research field. Similar keywords appear bigger, dependent on the frequency they are mentioned. This criterion refers to the description of the research field.

It is then checked if a *theory is mentioned (1.2).* A reasonable derivation of the research approach with a theory as foundation might increase the external validity of the research (Deaton 2010), as it was discussed in *chapter 3.4.2* already. The indicator is defined by the question: *Is a basis theory mentioned?* It can be answered with the categories *yes* and *no*. It is examined by an open text analysis. The question is answered with *yes*, when any kind of descriptive assumption on the underlying mechanisms of the examined object is existent. The question is answered with *no*, when there is no such assumption. This criterion is linked to the criterion *explanation on results (5.2)*, since both address the theoretical basis of a study. The latter is explained in the part *Quality Criteria (5)*.

A *hypothesis (No. 1.3)* presents a suggestion on how the research object would change throughout the study. Hypotheses specify a general theory and apply the mechanisms on a certain situation. They predict a certain behaviour of, or result on, the research object. A hypothesis will rather be included in quantitative studies, since it is a clear assertion, that can be confirmed or disproved. However, Weischer (2007) argues that a hypothesis is useful for any kind of empirical research because it avoids inductive reasoning and therefore might contribute to higher quality of the research. For this thesis, the existence of a hypothesis is examined by a qualitative text analysis. The indicator is defined by the question: *Is a hypothesis for the research mentioned* (*hypothesis mentioned 1.3.1)?* It can be answered with the categories *yes* and *no*. The question is answered with *yes*, when any kind of predictive assertion of the research object exists. The question is answered with *no*, when there is no such assertion. This criterion is used to draw conclusions about the quality of a study, referring to the part *Quality Criteria (5)*. Studies with a testing approach are expected to provide a hypothesis, whereas all other studies do not necessarily have to provide one.

Furthermore, it is asked if the *hypothesis fits the questioning (1.3.2)*, of the research questions. This criterion refers to the logical rigor and can be answered with the categories *yes* and *no*. It is answered with *yes* when hypothesis and question address the same problem and study approach, which is subjectively assessed by the reviewer.

The r*esearch questions (1.4)* are collected by an open text analysis. Also implicated research questions are included; they do not have to be mentioned interrogatively. This criterion is part of the study checklist provided by Jodrell and Kaoukji (2020). It is used to conclude on the next criterion: *kind of questioning (1.5)*.

The *kind of questioning (1.5)* is examined. A certain approach of questioning points to the following study design and both should match. The indicator is defined by the question: *How are the questions structured?* The categories' names are based on the categorization of Stein (2022). It is answered with *explorative* when the research questions point to an unsystematic examination of the research objective, with quite distinct methods. It is answered with *descriptive* when the research questions point to a qualitative approach, the objectives are clear cut and some kind of structure in the research approach can be observed. It is answered with *testing hypothesis*, when the research questions imply a quantitative approach with statistical calculations. It is answered with *prognosis*, when the research questions address possible future development. And it is answered with *evaluation*, when the research questions address the aim to draw general conclusions on examined effects. Alternatively, *evaluation* is chosen when the study is commissioned by a stakeholder (Bortz and Döring 2006), regardless the kind of questioning. This criterion is an attempt to reveal underlying deductive mechanisms of a research approach. The insights of this criterion contribute to a better understanding of the research field.

4.3.2.3 Part 2: The Study's Design

The next part focusses on the study *design (2)* and is based on the insights of Stein (2022). It contributes to the assessment of the *reasonability (5.3)* of a study and refers to the description of the research field. The criteria 2.5 to 2.7 additionally refer to the quality of a study, as it is explained below.

The *scope of the study (2.1)* is analysed. The indicator is defined by the question: *What is the study´s scope?* This can be answered with the categories *case study*, *comparative study*, and *discourse analysis*. It is examined if a certain focus leads to a certain study result.

Furthermore, it is on the *level of experimental characteristic (2.2)*. If it is actually reasonable to conduct experimental field studies is discussed in *chapter 3.4.1*. At this point, the data is collected without judgement. The indicator is defined by the question: *Is the design (quasi-)experimental or with an ex-post-facto design?* It can be answered with the categories *(quasi-)experimental* and *ex-post-facto-design*. It is examined if the experimental characteristics correlate with certain results.

The next characteristic is on the *dimension of time (2.3)*. The indicator is defined by the question: *What is the time structure of the study?* It can be answered with the categories *longitudinal study* and *cross-sectional study*. Another criterion regarding longitudinal studies is called *longitudinal sampling (2.3.1)*

and refers to the kind of sample, namely *cohort study, panel study* or *trend study*. Additionally, the *number of surveys (2.3.2)* are collected, as well as the timespan the study lasts *(timespan 2.3.3)*. A correlation between the time dimension and the results is examined.

Next, the *randomization (2.4)* is analysed. It is asked for *general randomization (2.4.1)*, with the categories *yes, no, partly* and *n/a*. The randomization of workers was particularly added during the analysis, in order to achieve clearer insights (*worker randomization 2.4.2)*. Furthermore, the *sample size (2.4.3)* is collected, at first in open text form and secondly as comparable number. A correlation to the study results is examined.

An important criterion is the kind of working situation that has been examined in the studies (*labour kind 2.5*). As shown in *chapter 2.2*, the employment situations in rural agriculture are diverse. Whether the kind of labour is stated in a study is an important quality criterion, since it is very useful to interpret the results. The information for this criterion is firstly collected in open text form and secondly summarized with suitable keywords.

Besides workers, in some studies other stakeholders got included, for instance farm owners, exporters, representatives from aid agencies or the government. These persons are addressed with the criterion p*ersons examined additionally (2.6.1)*. Additionally, the criterion is combined with the note, whether the respective persons are *interested in positive sector reputation (2.6.2)* or if they are probably influenced by a stakeholder which might be interested in a positive reputation. This accounts, for instance, for the *worker committee*, which is likely to be manipulated by the management (Raynolds 2012). Stakeholders that are suggested to be interested in a positive sector reputation are: persons in farm management positions, farm owners, exporters, worker committee members and governmental representatives. Stakeholders that are not suggested to be interested in a positive sector reputation are: aid agencies and other non-governmental organizations. It could be possible that answers of biased stakeholders would influence study results. Therefore, a correlation to the results is examined.

The *place of data collection (2.7)* could have an impact on the results as well. Workers might be more honest on their work experiences when a survey is not taken place at work (Dragusanu et al. 2022). Therefore, the categories are *private* and *work*. A correlation between the place of data collection and the study results is examined.

4.3.2.4 Part 3: The Study's Methods

The part of *methods (3)* aims to discover the methods and ask, if they apposite the research questions. It is based on the insights of Baur and Blasius (2022).

The *methods (3.1)* are first collected in open text form and then summarized with keywords. Two important keywords are *interviews* and *questionnaires*, which shall refer to different survey structures:

I*nterviews* refer to open structured and semi-structured surveys, whereas *questionnaires* refer to structured surveys. The methods are categorized in different *kinds (3.1.1)*, namely *qualitative, quantitative,* and *mixed* methods. For practical reasons, if surveys are included, they are categorized again with the criterion *structured/ open (3.1.2)*. This criterion only refers to the part of data collection where workers have been included. Meaning, when a study includes a qualitative analysis of governmental documents and a structured worker questionnaire, criterion 3.1.2 is answered with *structured* according to the method *questionnaire*.

Indicators (3.2) of the respective methods were first assessed on the criterion if they were mentioned *somehow*. This criterion turned out to be too superficial. It was changed into *Indicators explained (3.2)*. It is answered with *yes*, when some explanation on some indicators of the methods could be observed, regardless to the extent. This is a compromise, since information on the indicators are provided in different detail.

The criterion *methods fit the research question (3.3)* can be answered with *yes* and *no*. It is answered with *yes* when the methods seem to approach the research questions in a reasonable way. The decision is drawn subjectively by the reviewer.

The inclusion of a *cost-benefit-analysis (3.4)* is examined, based on Stockmann and Meyer (2014), who stated this criterion to be an characteristic of impact evaluations. It can be answered with *yes* and *no*, and is answered with *yes* if such an analysis in included.

4.3.2.5 Part 4: The Data Editing, Analysis and Results

In the next part the *data editing and analysis (4)* is examined and it is based on Lück and Landrock (2022) and Kuckartz and Rädiker (2022). The criteria 4.2.1 and 4.3 were also addressed by Pawson (2002a).

The *editing (4.1)* is examined. It is collected in open text form how data was edited after collection, and afterwards keywords are deduced. If no information on the data editing is provided, *n/a* is stated as answer to the criterion.

The information on the *analysis (4.2.1)* is collected in open text form as well. This criterion contains a short description on the kind of analysis and certain steps that were applied in the study. The descriptions are then summarized with keywords. In another criterion, the *used estimators (4.2.2)* are noted, which is mainly important for the studies with quantitative or mixed methods. Both criteria will refer to the *reasonability (5.3)* of the study.

To collect the study's results is kind of obvious and also suggested by Pawson (2002a). The criteria on the study results were differentiated during the analysis process. The results were first collected all in one. During the analysis the reviewer found it interesting to split the results up into *monetary*

results (4.3.1) and *non-monetary results (4.3.2)*, since in many studies both parts are mentioned separately. To be precise, fairtrade's monetary effects on loans and extra payments, e. g. maternity payment, for this thesis refer to *non-monetary results* because they are mostly mentioned together with, and indistinguishably from, non-monetary effects like health services or trainings. The criterion *monetary results* will only refer to the wages of workers. For both criteria, *monetary results* and *non-monetary results*, the same categories count. In the beginning, the categories *positive, negative, mixed, negligible, no answer*, and *no answer possible* were included. Results are categorized as *positive* (*negative*), when the effects of fairtrade on workers was observed to be *positive* (*negative*). Results are categorized as *negligible*, when no effect of fairtrade was observed. Results are categorized as *mixed*, when *positive* and *negative* effects of fairtrade were observed. Results are categorized as *no answer*, when the study provides no information on the results, although the study design theoretically could provide results. Results are categorized as *no answer possible*, when the study design does not produce and results on the effect of fairtrade on hired farm workers. During the analysis, many results revealed *mixed* findings. Therefore, in a second step, to gain a clearer picture, the reviewer decided on the category *mixed*, whether the results regarding workers are *rather positive* or *rather negative*. Those results are included in the criterion *tendency of results (4.3)*.

The *premium (4.3.3)* was added during the analysis, when it emerges to be interesting that the premium got emphasized a lot in some studies whereas got not mentioned at all in others. The categories are straight forward: the premium can either be *positively mentioned, negatively mentioned,* or *not mentioned*.

The *significance (4.4)* of a study is evaluated. The results can either be *significant, not significant, partly significant* or the significance criterion cannot be applied (*none*) due to the design of the study. The decisions are made in accordance to the reporting of the authors. If the authors state their findings to be significant, it is counted as such, even if it is not obvious how the significance had been calculated.

It is examined if *implications fit the questioning (4.5)* of the research questions. Glewwe (2020) emphasized the risk of overinterpreting study results. Whether it is appropriate to draw implications might be suggested in relation to the *kind of questioning (1.5)*. The implications are collected by a qualitative text analysis. The indicator question is: D*o the implications fit the questioning?* It can be answered with the categories *yes* and *no*.

If a study is *useful in terms of impact evaluation (4.6)* is examined because of the appliance of impact evaluations in the research field. This criterion is answered from subjective perspective based on the assessment of the reviewer. It might be interesting to discover, if studies can be used as impact evaluations, even if they were not conducted as such.

4.3.2.6 Part 5: The Study's Quality

Finally, the last part examines the *quality criteria (5)*. It is roughly based on Krebs and Menold (2022), for quality criteria in qualitative research, and Flick (2022), for quality criteria in quantitative research, as well as on Jodrell and Kaoukji (2020), for criteria regarding the limitations of studies.

An attempt is made to conclude on the quality of the studies. Therefore, some quality criteria were developed. It is asked if the authors discussed the quality of their approach (*quality criteria addressed 5.1.1*), if possible *confounding variables (5.1.2)* were mentioned, if there is an explanation on the *handling of divergent results (5.1.3)* and a *reflection on internal and external validity (5.1.4)*. All those criteria can be answered with *yes* and *no*, and are answered with *yes* when the respective criterion is to be found in the report of the study.

Furthermore, it is examined if there is an explanation on the results (*explanation on results 5.2*). This criterion refers to the theoretical basis of the study. If the authors set their approach on a *theoretical basis (1.2)*, it is interesting to interpret the results in relation to the theory as well. It can be concluded on the self-consistence and therefore contributes to the quality of the study. The criterion can be answered with *yes* and *no*.

An additional reflection on the self-consistence takes place with the criterion *reasonability (5.3).* It contains a subjective examination along the line of reasoning overall. The *problem background (1.1)* should prepare on the *research questions (1.4).* The questions should be approached by the general *study design (2),* and applied by the *methods (3).* The methods should implicate the respective data *analysis (4.2)* and with the analysis conclusions on the *results (4.3)* should be drawn. If the study offers four to five of those linkages, it is categorized as *reasonable*. If it offers two or three logical linkages, it is categorized as *partly reasonable*. If it offers none or one logical linkage, it is categorized as *not reasonable*.

If the study was *peer-reviewed* (*5.4*) can be answered with *yes, no* and *no answer*. This criterion refers to the study quality as well, though it must be treated with caution, since a publication bias is possible. This criterion is also included into the study checklist of Jodrell and Kaoukji (2020).

It is assessed if the authors themselves state their findings to be practically relevant (*practical relevance stated 5.5*). This criterion can be answered with *yes* and *no*.

And lastly, the *transparency (5.6)* is addressed. The approach is to count how often a criterion was stated as *n/a*. The number of *n/a*s in relation to the number of all criteria results in a percentage, which reveals how many information had been provided by the authors. This can allow conclusions about the transparency of the author's reporting.

4.4 Editing of the Data

The catalogue got applied as the reviewer read the studies and filled out the catalogue template, one for each study. The template is attached to the thesis in *Appendix 2*. Subsequently, an excel table was filled out, with each criterion in a respective column and the results for each study in a respective row. Aiming to make the information in the excel table more accessible, abbreviations had been used, as well as the numbers *1* and *0* for the categories *yes* and *no*. The abbreviations can be comprehended in *Appendix 1* and the excel table is to be found in *Appendix 3*. The results were then examined with the help of pivot tables. For the descriptive analysis, each criterion was analysed in a table with the number of studies split up on the respective categories. Where suitable, graphs were developed from the values.

5 Results

5.1 Descriptive Results

5.1.1 Part 0: The Study's Frame

In the analysis 18 studies from 15 different authors are included. Most of the studies got published between 2018 and 2022 (cf. *Figure 4*), the oldest ones in 2009 and the most recent one in 2023 (*publication year 0.1.3*).

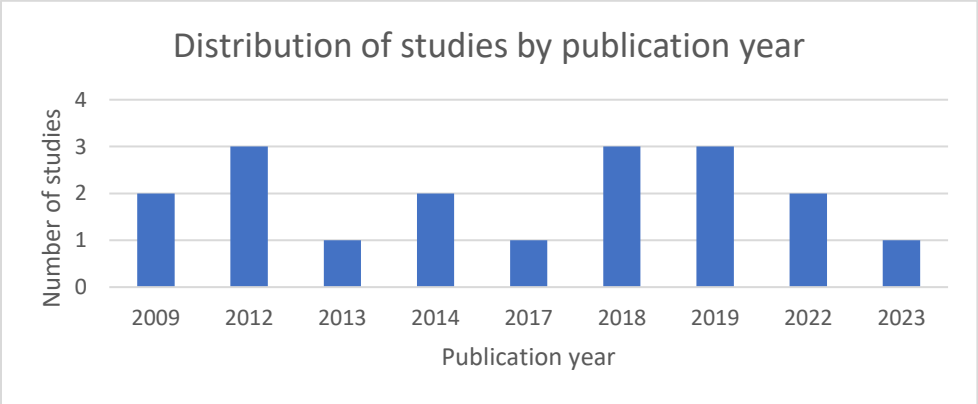

Figure 4 - Distribution of studies by publication year (own figure). The data refers to the criterion publication year (0.1.3).

Nine of the articles were conducted by researchers with a sociological *scientific background (0.2)*, five had an economical background and four studies were conducted from an interdisciplinary research team (*Figure 5*).

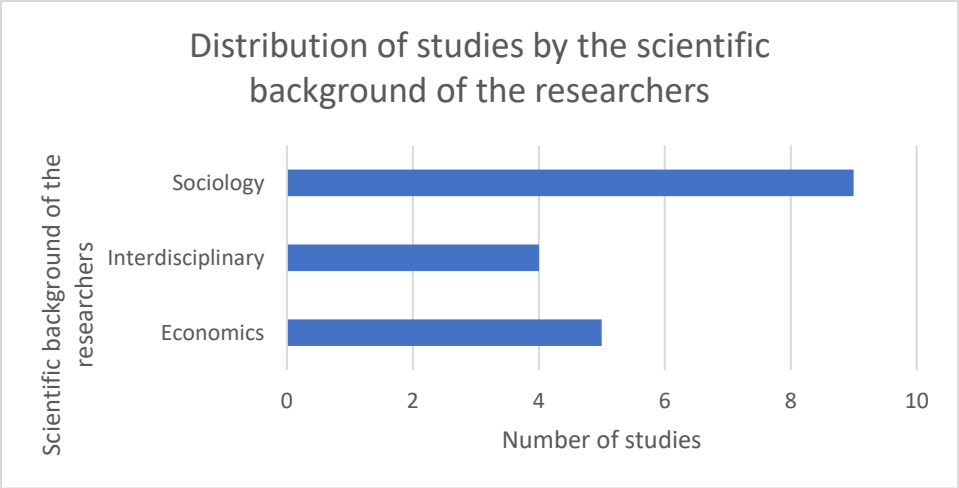

Figure 5 - Distribution of studies by the scientific background of the researchers (own figure).
The categories sociology, interdisciplinary and economics refer to the criterion scientific background (0.2).

Eight studies took place in African countries, seven in Latin American and Central American countries and three studies were conducted in Asia (*0.4.2 continent*). In Latin America and in Asia, more positive than negative impacts of fairtrade on hired farm workers could be observed.

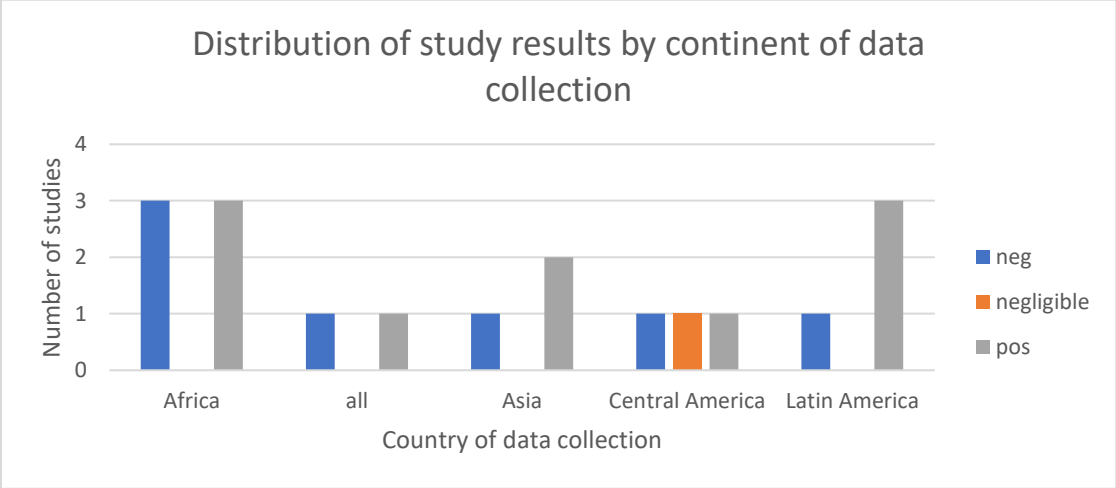

Figure 6 - Distribution of study results by continent of data collection (own figure)

Several *products (0.5)* were produced on the fairtrade farms, namely coffee, cocoa, tea, flowers, bananas and pineapples (*Figure* 7). There were no farms producing more than one of the observed products at a time; the part in *Figure* 7, addressing "tea, coffee, flowers", refers to one study that examined different farms with the respective products. Products were sold on fairtrade and on conventional distribution channels.

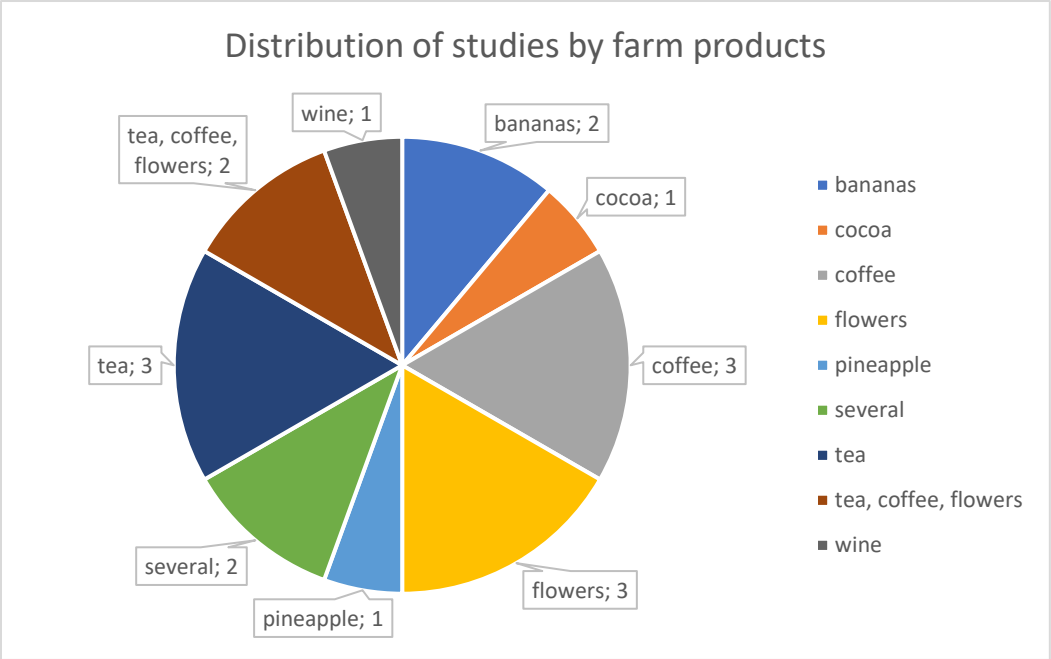

Figure 7 - Distribution of studies by farm products (own figure).
Speech bubbles containing the product and the number of studies that examine farms with the respective product.

Most of the studies, though not all of them, put hired workers in focus (*0.6 main research object*). Some studies actually examined voluntary sustainability standards or had a systematic approach, examining several different impacts of fairtrade. Furthermore, different *problem backgrounds (1.1)* were stated to be initial for the respective studies (*Figure 8*). Mostly, the concern is to examine the *impact of fairtrade*, which was stated in nine of the studies. In four studies, more detailed motivations were stated, for instance examining *patron-client-relations* (Besky 2014; Makita 2012) or discover *mechanisms of poverty reduction* (Cramer et al. 2017).

Figure 8 - Word cloud on the problem backgrounds stated in the studies (own figure)

Two of all studies were officially commissioned and funded by Fairtrade International, and for four of all studies it was not possible to discover the funding (*0.3 financier linked to fairtrade*). The remaining 12 studies should certainly be unbiased regarding the financier. Both of the commissioned studies reveal no effects on wages and (slightly) positive effects on non-monetary benefits.

5.1.2 Part 1: The logical Derivation of the Research Approach

Eleven studies developed their research approach on basis of one or several theories (*theory mentioned 1.2*) and three studies stated hypotheses for their research (*hypothesis mentioned 1.3.1*). In four cases, neither a hypothesis nor a basis theory is mentioned. With 15 of 18 studies providing some kind of content basis, the impression emerges that the derivation of the research approach is not too weak. Though, as any kind of theoretical reference was considered, the meaningfulness of this results might to be doubted. The subjective impression of the reviewer, when reading the articles, is that the theoretical derivation of the research approaches is rather blurry.

The *research questions (1.4)* were collected mainly for better understanding and to derivate the *structure of questioning (1.5)*. The structure of questioning was either mentioned in the article or derivated by the reviewer on basis of the article's information. Nine studies include some kind of statistical testing in their examination, though the underlying computation was not always presented. The explorative and descriptive studies sum up to seven, and the differentiation is not always clear cut. As expected, the structural approach fits the methods in most of the cases, as represented in *Table 8*. Qualitative methods are mostly used for descriptive and explorative study designs, whereas testing design almost always include quantitative methods.

Table 8 - Allocation of Studies split up on the Structure of Questioning and the Kind of Methods

	3.1.1 Methods kind			
1.5 Structure of questioning	mixed	n/a	qualitative	quantitative
descriptive			3	
evaluation	2		1	
explorative			2	
integrated analysis (descriptive and evaluation)	2			
testing	4	1	1	2

In the catalogue, the category *prognosis* was included on basis of Stein (2022). It turns out that no *prognosis* design was to be observed in the research field. However, the criterion *4.5 Implications mentioned* draws on it, revealing possible suggestions or recommendations of the authors, though only as part of the conclusion and not as overarching approach. Three evaluation studies are included, but two of it are impact evaluations in the sense that they are conducted due to the order of a client (Glewwe 2020). None of the studies include a *cost-benefit-analysis (3.4)*, which otherwise would be a characteristic of evaluations as well (Stockmann and Meyer 2014).

An inconsistency occurs at the category *testing hypothesis*. Nine studies were categorized as *testing hypotheses*, though the analysis reveals six of the nine studies to actually not state a hypothesis (*hypothesis mentioned 1.3.1*). This is remarkable regarding the quality of the studies, since proper statistical testing requires to state a null hypothesis and an alternative hypothesis beforehand. The reason might be the variety in wording. The author of this thesis tried to collect any possible wording of research objectives, but sorted them into the criterion *research questions (1.4)* rather than into the criterion *hypothesis mentioned (1.3.1)*, when not clearly stated as hypothesis. What can be concluded anyway is that the wording is not straight forward in many cases.

5.1.3 Part 2: The Study's Design

Another challenge occurs in the category *study scope (2.1)*. While many of the studies state themselves to be case studies, by definition of this thesis, most of them are actually comparative, when they include more than one estate in the examination. This depends on the reference point. When the single worker is the reference, none of the studies would be a case study, since all of them examine more than one person. If the reference point is on the estate level, there is only one case study that focus on a single farm only (Makita 2012). Almost all other studies are categorized as comparative studies, regardless what they state their selves to be. The exception is Oya et al. (2018), which is a systematic research and is categorized as *discourse analysis.*

The results on the criterion *experimental design (2.2)* are unambiguous. Of 18 studies, only one study state itself to conduct an *experimental design* (Rubio-Jovel et al. 2024), all other studies can be categorized as *ex-post facto designed*.

Regarding the *time dimension (2.3)*, 12 of the studies were designed cross-sectional, whereas five have a longitudinal design. Longitudinal studies might not be conducted so often, because they require more financial and time resources. Furthermore, no quantitative longitudinal studies can be observed, they are only of qualitative and mixed designs. Longitudinal studies include different kinds of *sampling (2.3.1)*, two conduct a cohort study, two a panel study and one longitudinal study had an explorative approach and no sampling was conducted. No trend design could be observed.

The *number of surveys (2.3.2)* was collected to conclude on the reliability of the data. Nine studies are based on one data survey each. Four studies rely on two or three surveys, which had been conducted in a *time span (2.3.3)* of one to four years. One longitudinal study includes nine surveys in nine years. Four studies did not specify the number of surveys. The results are consistent with the *time dimension (2.3)*, cross-sectional studies mainly conducted one single survey and all longitudinal studies include more than one survey. However, one cross-sectional study includes two surveys at different times, but the interviews were conducted with different target groups, namely stakeholders and workers.

The *randomization (2.4)* of the sampling was not informative initially, because the category *partly* was included. Studies used randomization partly when they purposively chose an estate by certain criteria, but randomly chose the workers to be interviewed. The reviewer decided that the randomization of workers is more important to the research questions than the randomization of other objects, like estate or country. Therefore, in a second step, a focus was put to the randomization of workers, which revealed a clearer picture. 10 out of 17 studies used randomization of workers. Six studies did not randomize the sample, one study did not give an information on it, and one study was meta-research and the category could not be applied.

Regarding the *sample size (2.4.3)*, five studies were excluded. Four studied did not state a sample size and one stated a vague percentage for the sample size of each estate surveyed. For the remaining 14 studies, the average sample size is 497, with a range from 36 to 1700 participants, and the median is 253,5. When the three very large samples were excluded (with 1700, 1700 and 1050 participants), the average sample size is 329.

The *labour kind (2.5)* was addressed in different detail. Some studies only refer to the kind of company, others to the characteristics of the workers. One study has to be excluded, because of its meta design (Oya et al. 2018). In total, 11 studies could be counted into the category *plantation workers*, three into the category *small-farm employees*. Other explanations on the criterion were presented by four studies, using the terms *casual wage workers*[2], *landless labourer* (Granville and Telford 2012) or *unskilled workers* (Dragusanu et al. 2022). Latter cannot be allocated to a certain kind of farm, as in both cases no additional persons were examined. The products of those farms are coffee (Dragusanu et al. 2022) and wine (Granville and Telford 2012), and at least wine is probably cultivated on a plantation, whereas coffee is also common to be grown on small farms. If for the study of Granville and Telford (2012) the category *plantation workers* would be suggested, and for Dragusanu et al. (2022) the category *small-farm employees*, a substantial higher proportion of labour was examined on plantations (12 examinations) than on small farms (5 examinations). It might be a reflection of the present problem that the extent of employment in the small-farm sector is still underestimated (Cramer et al. 2014; Dragusanu et al. 2022).

[2] The explanation *casual wage workers* was provided by Cramer et al. 2014, which is most likely the same sample as Cramer et al. 2017, and therefore probably refers to *small-farm employees* as well.

Table 9 - Descriptions of the labour kind in the respective studies, and the tendency of the results

	Main author	Tendency of results			
Labour kind		negative	negligible	positive	results
Plantations					12
landless labourers	Granville and Telford (2012)	1			1
plantation workers	Krumbiegel et al. (2018; Makita; Raynolds), Raynolds (2012), Raynolds (2014), Raynolds (2022; Makita), Makita (2012), Bennett (2018), Ruben and van Schendel (2009), Rubio-Jovel et al. (2024), van Rijn et al. (2020)	1		8	9
plantation workers; organizational	Besky (2014)	1			1
plantation workers; organizational; factory workers	Siegmann et al. (2019)			1	1
Small-farms					5
casual wage workers	Cramer et al. (2014)	1			1
small-farm employees	Cramer et al. (2017), Meemken et al. (2019), Valkila and Nygren (2010)	1	1	1	3
unskilled workers	Dragusanu et al. (2022)	1			1
Meta-design (excluded)					1
workers	Oya et al. (2018)	1			1
results		7	1	10	18

The table represents the description on the kind of labour used in the examined studies. The first column contains the description, the second column contains the studies, who used those descriptions on their report. Colum four, five and six contain the tendency which the studies´ results are pointing to. The last column summarizes the results. It is to see that the studies examining plantations are more than twice as much as those examining small-farms.

Persons that are examined additionally (2.6.1) are representatives from management, government or aid agencies, as well as exporters, farm owners and workers from the workers committee, among others. Twelve studies included additional stakeholders. In each case at least one of them could be suspected to be *interested in* a *positive sector reputation (2.6.2),* which is therefore answered with *yes* 12 times.

The *place of data collection (2.7)* was not mentioned in most of the cases. On 16 studies, this criterion was able to be applied, but ten studies did not provide an answer to this question. The remaining chose private situations for the surveys. In two studies it was mentioned that they also chose work situations but only additionally and just in a few cases.

5.1.4 Part 3: The Study's Methods

Almost all answers to *methods (3.1)* include interviews or questionnaires, which is common for sociological studies. Nine studies state a mixed methods approach, six are qualitative and two are quantitative. The mixed method approaches got split up by the criterion if workers were examined in a structured or open way (*structured/ open regarding workers 3.1.2*, *Table 10*), with the assumption that *structured* refers to quantitative methods, and *open* to qualitative methods. Five of the mixed method approaches reveal a structured approach regarding workers, two reveal an open approach. Therefore, the examination on workers with qualitative and quantitative methods is quite even distributed.

Table 10 - Distribution of studies by the kind of methods and the kind of examination regarding workers (own figure)

	Kind of methods				
Structured/ open examination of workers	mixed	n/a	qualitative	quantitative	results
both	1				1
none	1		1		2
open	2		5		7
structured	5	1		2	8
results	9	1	6	2	18

Indicators are *explained (3.2)* in 11 out of 18 cases. Though in most of those 11 studies, the explanation is not comprehensive.

The criterion *methods fit the research question (3.3)* is answered consistently with *yes*.

5.1.5 Part 4: The Data Editing, Analysis and Results

The *editing (4.1)* of the data is specified in three studies. In two studies the exact statistical program is mentioned, in one study just a "software for qualitative data analysis" is mentioned.

The criterion *analysis (4.2.1)* is answered in different detail. For example, in some studies *regression-analysis* are stated, some specify *analysis of covariance, ordinary least squares* or *chi-squared tests*, among others. *Used estimators (4.2.2)* are mentioned by eight studies. Ten studies do not specify the analysis, including all seven studies that applied open interviews (*3.1.2*) regarding workers.

The results in general were split up into *monetary results (4.3.1,* referring to wages) and *non-monetary results (4.3.2,* referring to all other benefits*)*, and the categories *mixed* results and *negligible* were split up as well as much as possible. With focus on monetary results, eight studies remained *negligible*, five studies discover *positive* effects of fairtrade certification on the monetary situation of workers, and four studies discover *negative* results. Indeed, Cramer et al. (2017) discover lower wages on fairtrade farms compared with non-fairtrade farms. Though, the negative result of Oya et al. (2018) is partly

based on the insights of Cramer et al. (2017), and the primary data of Cramer et al. (2014) and Cramer et al. (2017) are most likely identical. Another negative result is presented by Rubio-Jovel et al. (2024), who anyway discuss their findings to be independent from fairtrade, since all certification systems examined revealed negative results. However, the fairtrade certification was correlated with the lowest of all wages. When summing up the three examinations with the, most likely, same primary data set as one, there are in total two surveys that reveal negative correlations of fairtrade and worker's wages (cf. *Table 11*).

Table 11 - Publications with negative monetary results (own figure)

Publication		monetary results (4.3.1)
Cramer et al. (2014)	Most likely based on the same sample	Negative
Cramer et al. (2017)		
Oya et al. (2018)	Examination based on Cramer et al. (2014) and Cramer et al. (2017), among others	
Rubio-Jovel et al. (2024)		Negative

With focus on *non-monetary results (4.3.2)*, seven studies reveal *positive* findings, five reveal *negative* ones, and two studies offer *mixed* findings. *Non-monetary* effects included for example providence of health care, clean toilets, sickness pay and trainings on agricultural practices (Rubio-Jovel et al. 2024; Cramer et al. 2014; Krumbiegel et al. 2018). Reasons for negative results are discussed by some of the authors. Cramer et al. (2014) suppose that negative effects might also be caused by the farm size, since smaller farms in general perform worse, compared with plantations, regardless the certification. The negative effects of fairtrade on school enrolment, discovered by Dragusanu et al. (2022) might be caused by a shift in revenues within the coffee production chain. They observed a decrease on incomes of intermediaries in the coffee production and suppose that they therefore cannot afford school for their children anymore. Hence, this result might not be connected to the workers hired on agricultural farms.

After the differentiation, the results reveal that *monetary* and *non-monetary* results never occur contradictory. Either both characteristics were *positive* (or *negative*) or one characteristic had a distinct direction and the other one was *negligible* or not mentioned (*n/a*) in the report. Therefore, the results are categorized again with the criterion *tendency of results (4.3)*, containing the overall direction of the results. The categories are *positive, negative* and *negligible*, and are chosen regarding the most directive result. For example, Makita (2012) reveals no *monetary* results at all but positive *non-monetary* effects and will therefore be categorized as *positive* regarding the criterion *tendency of results (4.3)*.

As suggested, the *premium (4.3.3)* got mostly mentioned *positively* or was not in focus of the study and therefore got not mentioned at all. Ten studies emphasized that the premium has some kind of positive impact on hired farm workers, seven studies did not mention the premium directly. Only one study judged negatively on the impact of the premium (Cramer et al. 2017). The authors discovered and emphasized the fact that the premium in the small-farm context is only meant to support farmers but not workers. They are criticising the fairtrade system in this specific point.

The *significance (4.4)* revealed mixed findings. Out of all fourteen studies that address the significance of their results, six studies find their results to actually be significant, six did not reveal significant results and two reveal partly significant results. Two qualitative studies made statements on the significance, which is confusing. Either the wording is inaccurate, using *significant* as a synonym for *interesting* or *meaningful*, or the methods are differently to what it seems like in the report.

Implications fit the questioning (4.5) of the research questions in 11 out of 16 cases. Two studies did not mention any implications.

If a study was *useful in terms of impact evaluation (4.6)* is very difficult to answer, since no clear indicators were stated beforehand. And the criterion *cost-benefit-analysis included (3.4)*, which should be considered at this point, was found to be negative for all the studies and therefore does not make any difference. The examination seems to be arbitrary. Nine and nine studies are found to be useful and not useful regarding impact evaluation.

5.1.6 Part 5: The Study's Quality

Quality criteria are addressed by most of the studies. 12 studies directly address the quality of the study (*quality criteria addressed 5.1.1*), 15 studies mention *confounding variables (5.1.2)* and an *explanation on the results (5.2)* is provided by 11 studies. Only five studies mention the handling with *divergent results (5.1.3)* and four studies provide a discussion on the *internal and external validity (5.1.4)*.

The criterion *reasonability (5.3)* is answered quite subjectively by the reviewer. It reflects the overall impression of a study's conduct and line of reasoning. Accordingly, 12 studies are categorized *reasonable* and five are categorized *partly reasonable.* One study is categorized as *not reasonable* (Raynolds 2022), because it was unidentifiable what exactly the results of the study are and how the analysis of the data took place.

Regarding *peer-reviewing (5.4),* it is to see that most of the studies are peer-reviewed, though five are not peer-reviewed and in two cases it is not possible to determine.

The *practical relevance (5.5)* is stated by 12 studies, who emphasize their contribution to current challenges. Six studies draw no connection between their studies and current challenges.

The *transparency (5.6)* offers an insight on the comprehensiveness of the studies report. Six studies offer more than 95 % of all answers to the catalogue, and 14 studies offer more than 90 %. The least transparent was found to be Raynolds (2014), which still offers 85 % of all possible answers.

5.2 Examination of Correlations

It was suggested, at the beginning of this thesis, that the variety in study results is based on different conducts of the studies, which might have an influence on the results. Hence, with an analysis of the characteristics of a study in relation to certain results, correlations might be revealed. The sample is quite small and it will not be possible to draw distinct conclusion on causalities, though it might be possible to offer a lead for further research. The correlations are not examined with every criterion, but with those who seem to offer interesting insights.

Overall, the analysis revealed more positive than negative results on the impact of fairtrade certification on hired farm workers. Therefore, the Allocation of studies on the categories of another criteria in relation to the results within a contingency table has to be treated with caution. It is reasonable to focus on comparisons within a category rather than between different category-result-combinations. Many studies discover only small or none effects of fairtrade certification, and/or found effects but either monetary *or* non-monetary. Fairtrade certification therefore seems to be not very effective at all. Nevertheless, in order to provide a discussion and comparison on the studies, the criterion *tendency of results (4.3)* was developed. This criterion includes all effects, regardless the size, but does not refer to the results directly, but to the direction they are pointing to. Hence, a study with a small positive effect of fairtrade would be counted as *positive tended* in the following.

Regarding the *scientific background* (*0.2*) it can be observed that the categories *economics* and *interdisciplinary* contain an almost balanced number of *positive* and *negative tended* results. Whereas a research background of *sociology* reveals twice as much positive (six) as negative (three) results. This finding is also correlated with the *kind of methods used (3.1.1)*. As to see in *Table 12*, researchers with a sociological background published five qualitative studies, three with mixed methods and none quantitative ones. Four out of the five qualitative studies have a positive tendency of the results. The one exception is Besky (2014) who discussed the negative impact of colonial relationships between farm owners and workers and how they get strengthened by the fairtrade certification. All qualitative studies had been conducted by sociologists, except one with an interdisciplinary background.

Table 12 - Studies conducted by Sociologists

Main author	Publication year	0.2 Scientific Background	3.1.1 Methods kind	4.3 Tendency of Results
Bennett	2018	Sociology	qualitative	Positive
Makita	2012	Sociology	qualitative	Positive
Raynolds	2014	Sociology	qualitative	Positive
Raynolds	2012	Sociology	qualitative	Positive
Besky	2013	Sociology	qualitative	Negative
Raynolds	2022	Sociology	mixed	Positive
Siegmann	2019	Sociology	mixed	Positive
Oya	2018	Sociology	mixed	Negative
Rubio-Jovel	2023	Sociology	n/a	Negative

The table shows the main authors of each study, with the publication year of the study and three categories of the catalogue of criteria: the scientific background, the kind of methods used and the results tendency.

Half of all the studies contain a mixed methods approach (nine studies) with a nearly balanced distribution between positive and negative tended results. Since half of the sample includes mixed-methods, the finding of Oya et al. (2018), who discovered a lack of mixed-method approaches in the research on certification systems, cannot be confirmed in this specific context.

It was suggested that a bias occurs, if the *financier* of a study is *linked to fairtrade (0.3)*, because there might be an interest in positive results. Indeed, the two studies with an obvious link to fairtrade provide *positive tended* results, though only regarding *non-monetary* benefits for workers. Both of the studies did not find any effect of fairtrade on wages of workers. When the financier is not obviously linked to fairtrade, there were *negative tended* results observed nearly as much as *positive* ones.

The distribution is also quite balanced when examining if a *theory got mentioned (1.2)*. The existence of a theoretical basis is correlated with five *negative* and six *positive tended* results (cf. *Table 13*). No theory is mentioned in seven studies with two *negative* and four *positive tended* results. One could conclude that negative tended results occur more often when a theory got mentioned, whereas for positive tended results it did not make a big difference. However, since the theories mentioned within the studies are from different quality and differ in the extent of contribution to the studies approach, it is probably not recommended to draw conclusions on this comparison at all.

Table 13 - Allocation of Study Results in Relation to the existence of a Theory

	Tendency of Results			
Theory mentioned	negative	positive	negligible	Results
no	2	4	1	7
yes	5	6		11
Results	7	10	1	18

Number of studies that include/ do not include a basis theory, in relation to the tendency of the results.

The structure of the research questions is imbalanced towards *testing* studies (nine studies). The other half is distributed on *explorative, descriptive* and *evaluative* studies, and nearly all those studies reveal *positive tended* results. The *negative tended* results almost all occur in studies with testing research questions (*Table 14*).

Table 14 - Allocation of Study Results in Relation to the Structure of Questioning

	Tendency of Results			
Structure of Questioning	negative	positive	negligible	Results
descriptive	1	3		4
evaluation		2		2
explorative	1	2		3
testing	5	3	1	9
Results	7	10	1	18

The *scope of the studies (2.1)* is set into relation to the results. As most of the studies have a *comparative* scope, a comparison to other categories is difficult. The study conducted as discourse analysis, namely Oya et al. (2018), which was the only meta-analysis included, revealed *negative tended* results regarding the effects of fairtrade on hired farm workers, whereas the two case studies reveal a *positive tendency* (Makita 2012) and *no tendency* (Valkila and Nygren 2010) of the results.

The experimental approach of Rubio-Jovel et al. (2024) provides a *negative* tendency of the results. And all the longitudinal studies (5 out of 18 in total) reveal negative tendency of the results. In comparison, the results of 10 out of 12 cross-sectional studies direct to positive effects.

The randomization on the level of the workers in relation to the tendency of the results reveals no clear picture. When a randomization on the worker sample was applied, the results of seven studies point into a positive tendency (and three in a negative), whereas without randomization there were three *positive*, two *negative* and one *negligible tendency of results*.

Regarding the *sample size (2.4.3)* the tendency of results is mainly negative for samples with more than 100 participants: five out of eight studies reveal negative tended results. Almost all smaller samples, with 100 or less than 100 participants, reveal positive tendencies of the results (five out of six, the last

one reveals no effect at all). The four studies that did not give distinct information on the sample size come along with two positive and two negative tended results.

Regarding the *labour kind (2.4.5)*, 9 out of 12 studies examining workers on plantations reveal positive tended results. One examination on small-farm employees concludes on positive tended results, three examinations point into a negative tendency and one was negligible.

When *additional persons* were *examined (2.6.1),* in each study at least one was suspected to be *interested in a positive sector reputation (2.6.2)*. Eight of those 12 studies reveal positive tended results. Whereas, those studies who did not examine additional persons, reveal two positive and four negative tended results.

The *place of data collection (2.7)* could be suggested to reveal correlations to the results, but as all studies who give information on his criterion conducted the surveys in private settings, it is not possible to draw any further conclusions.

The relation of the *kind of methods (3.1.1)* and the results was already discussed above. An additional approach of the methods criteria is to distinguish between open and structured surveys regarding workers. If workers are surveyed in a structured way, a balanced picture of four negative and four positive tended study results emerge. With an open structure of the survey five positive, but only one negative tended result emerges.

The *premium* got *mentioned (4.3.3)* almost only linked to positive impacts for workers. And in relation to the results, it is to see that a positive mentioning of the premium is accompanied by positive tended study results (8 out of 10 studies). When the premium got not mentioned, or got mentioned in a negative way, six out of eight studies reveal negative tended results. It is to keep in mind that the positive tended results contain the non-monetary impacts of fairtrade and those are mostly caused by the premium. The criteria are therefore strongly correlated.

6 Discussion on the Findings

The results of this analysis reveal insights on the conduct of studies examining the impact of fairtrade certification on hired farm workers. The descriptive part of the analysis provides an overview on the characteristics of the studies and the interpretative part relates the characteristics to each other. In the following paragraphs, the quality of the studies will be discussed, the results will be related to the insights of other authors and suggestions will be drawn.

6.1 Correlations

This sub-chapter refers to the research question *RQ2: How are the studies designed? Are there correlations who might let assume the existence of biases?* In *chapter 5.2* correlations between study

characteristics and certain results are examined. The sample size used in thesis is too small to draw reliable conclusions, and the method cannot be used to reveal cause-effect-chains. However, some characteristics seem to emerge mainly in combination with certain results, which is presented in the following table:

Table 15 - Characteristics that tend to emerge together with to positive/ negative results.

Characteristics	Positively tended results / total category results	Characteristics	Negatively tended results/ total category results
plantations	9 / 12	small-farms	3/5
cross-sectional studies	10 / 12	longitudinal studies	5 / 5
sample size =< 100	5 / 6	big sample size > 100	5 / 10
descriptive/ explorative questioning	5 / 7	testing	5 / 9
Persons examined additionally	8 / 12		
qualitative methods	4 / 6		
Positive tendency of results		Negative tendency of results	

The table shows the allocation of results in relation to the number of studies in total referring to a category. For example, 12 studies are focussed on the examination of plantation. From those 12 studies, 9 studies reveal positive tended results. And 10 out of 12 cross-sectional studies reveal positive tended results, whereas all 5 longitudinal studies reveal negative tended results.

That the results would be rather positive for plantations than for small-farms was already supposed by Cramer et al. (2014), with the explanation that plantations can use scale effects to improve working conditions, whereas for small-farms it is relatively more expensive to implement improvements. Another important thought is that workers on small-farms are actually not in focus of the fairtrade system. The standards are not precise, in the fairtrade system it is assumed that small-farms are mainly family driven and do not depend on hired workers. This is proven to be wrong by several authors (Cramer et al. 2014), who examined the working situations on small-farms. It might also be possible, that fairtrade in fact is aware on the working situations on small-farms, but as well on the limited resources of small farmers for improvements. Fairtrade remains in this tension of setting up requirements to protect workers on the one hand and being dependent on the demand of the farmers on the other. Nevertheless, it seems like fairtrade might have indeed a positive impact on workers situation on plantations. However, an increasing focus of the fairtrade system on small-farms is recommended, which could be accompanied by research on the barriers and facilitators that influence small-farmers in improving workers situations. Further research with focus on the small-farm situations is highly suggested, since the research body is scarce.

What can be obtained additionally in *Table 15* is that some other characteristics mainly occur with positive tended results. Namely cross-sectional studies, a sample size smaller than, or equal to 100, whether persons, additional to hired workers, got examined in the study, a descriptive or explorative research approach, and qualitative methods. However, those characteristics only appear all together in two studies (Raynolds 2012; Makita 2012). All positive tended results are revealed by cross-sectional studies (10 studies), nine of them examining plantation workers. One exception is the research of Meemken et al. (2019), who focus on small-farm employees explicitly. Plantation workers are examined with longitudinal designs in two cases, which are the only two cases revealing negative tended results regarding plantation workers. Almost all negative tended results occur in studies with testing designs, with two exceptions that are revealed by descriptive and explorative approaches.

Quite unbiased insights are revealed when examining the mixed methods approaches only. There is a quite even distribution of longitudinal and cross-sectional designs, with very small and very large sample sizes, sometimes additional persons got included into the survey and sometimes not, and there are almost as many negative as positive tended results. There seems to be no imbalance in the studies conduct.

6.2 Quality of the Studies

This sub-chapter summarizes on the research question *RQ3: How can the studies quality be assessed?* The theoretical background is provided in *chapter 3.4.5*. The quality is assessed by several criteria. They are split up in two categories, referring to the quality of the reporting of the studies and the quality of the studies research approach.

Quality of Reporting

The criterion *transparency (5.6)* was the criterion that should mainly address the quality of reporting. The results include 14 studies with a transparency value > 90%. One could conclude that the reporting of the studies is mainly sound and reliable. But what has to keep in mind is, when developing the catalogue, it was tried to find criteria and categories that fit the studies as much as possible. Therefore, what the criterion *transparency* actually reflects, is the extent to which the catalogue was developed in accordance to the studies. Nevertheless, on the required information, the reporting in the research field is quite comprehensive. However, what is missing in most of the cases, is an explanation on the editing of the collected data, as well as a comprehensive explanation of the analysis. Both criteria are obviously very important to conclude on the quality of studies. Further research should emphasize this improvement.

Furthermore, what was not reflected in the criterion *transparency (5.6)*, are criteria with the categories *yes* and *no*, since *transparency* counted the number of given *n/a*'s, which was only applicable on eight criteria at all. But some additional criteria contribute to the quality of reporting, namely if *indicators*

got explained (3.2), if the *implications fit the questioning (4.5),* if *quality criteria (5.1.1)* and *confounding variables (5.1.2)* were addressed, if there was explained how *divergent results* were *handled (5.1.3)* and if a *reflection on internal and external validity took place (5.1.4).* Criterion 3.2, 4.5, 5.1.1 and 5.1.2 were each addressed in 11-15 studies (*Figure 9*), which is pointing to the attempt of studies to explain their approaches. Criterion 5.1.3 and 5.1.4 were only addressed in five and four studies (*Figure 9*). To address the *handling of divergent results (5.1.3)* contributes to the demand of a more comprehensive reporting on the analysis of the studies, mentioned above. The discussion on *internal and external validity (5.1.4)* should be increased, because it would help the reader to judge on the possibility of generalizing the reported results. This is of high importance, especially in the research field of fairtrade, since stakeholders, that are interested and/or dependent on the results, are not always trained with a scientific understanding. Authors in the research field should contribute to the lack of judgement by providing a sound discussion on the external and internal validity of their studies.

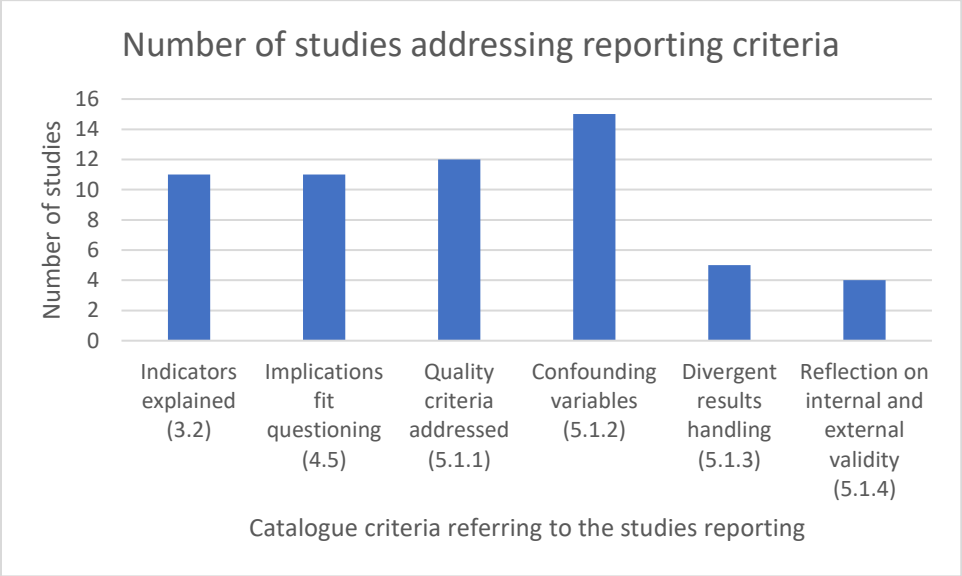

Figure 9 - Number of studies addressing the respective criterion (own figure).
For example, confounding variables are addressed by 15 studies, whereas only four studies reflect on internal and external validity. The chart represents all studies on which the criteria were assessed with yes.

Quality of Studies Research Approaches

On the quality of the studies, it can be concluded as well. Internal validity is met when the design is well suited to the aim of measuring the effects of the examined variables. Therefore, the methods should fit the research aim and the research context. The criterion *reasonability (5.3)* addresses those linkages, which was assessed by the reviewer to be met by most of the studies. Additional criteria that contribute to the assessment of the research approach are: if a *hypothesis is mentioned (1.3.1)* and if it *fits the research question (1.3.2),* if the *methods fit the research question (3.3)* and if the study got *peer-reviewed (5.4).* The latter can be used as confirmation of other authors on the reliability of the study. Furthermore, some possible bias criteria were addressed, namely: if the *financier is linked to*

fairtrade (0.3), if *worker randomization* took place *(2.4.2),* if *additional examined persons interested in a positive sector reputation (2.6.2)* and where the *data collection took place (2.7).* The results are summarized in *Figure 10*. The internal validity is promoted by a high number of studies addressing the upper five criteria in *Figure 10*. Randomization, peer-reviewing and a sound research design with fitting methods are already approached by several studies. Only a small number of studies address the criteria regarding the hypothesis (criterion 1.3.1 and 1.3.2), but this is also caused by the approach of the reviewer of this thesis and might be biased. It is discussed in *chapter 5.1.2*. The internal validity might be inhibited by latter criteria. The criteria *financier linked to fairtrade*, and *interested in a positive sector reputation* (*Figure 10*, red bars in the chart) are criteria that would inhibit the internal validity. As to see, criterion 2.6.2 is a matter in a great share of studies. All in all, the criteria in Figure 10 seem to be not very convincing in order to draw distinct conclusions on the internal validity of studies in the fairtrade research field. For assessing the internal validity of those studies, it would be more important to understand the data collection and analysis in detail. This was impossible due to poor reporting in most of the studies. This critique was already stated, in fairtrade research for example by Oya et al. (2018) and Cramer et al. (2014), but also for development studies and impact evaluations in general by Deaton (2010) and Thomas and Harden (2008).

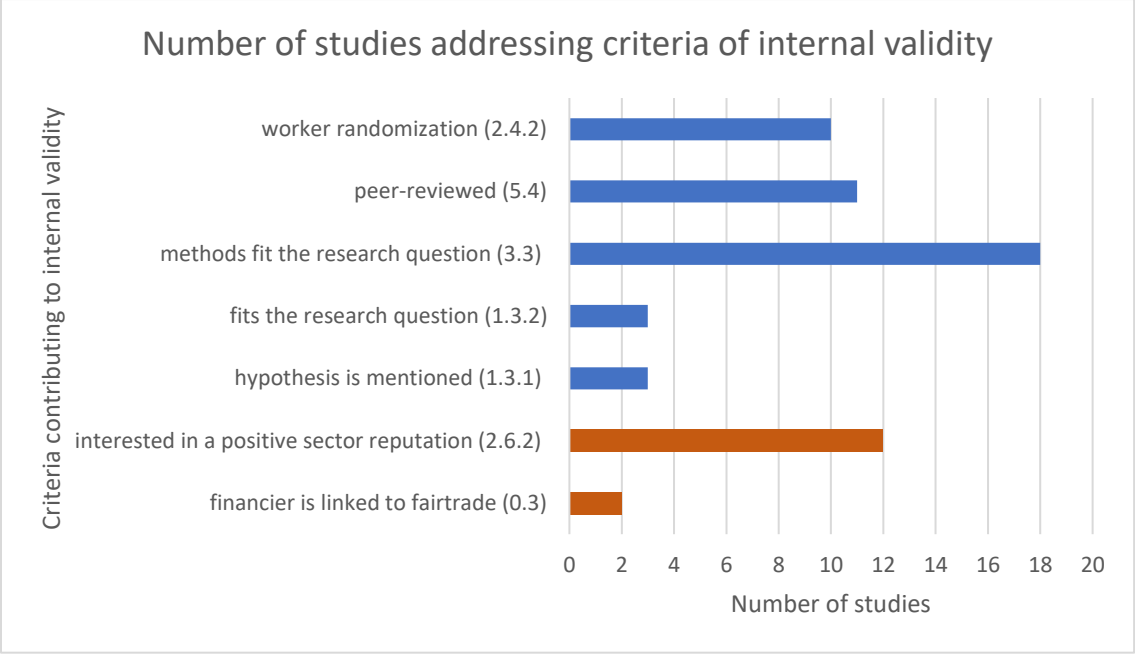

Figure 10 - Criteria contributing to the assessment of internal validity, and the respective number of studies addressing the criteria (own figure).
The chart represents all studies on which the criteria were assessed with yes.

External validity contributes to the question, if the results of a study can be transferred to other situations. As discussed in *chapter 3.4.4*, external validity is strengthened by a solid theoretical foundation and when context variables are considered. The theoretical basis was addressed by the criteria *theory mentioned (1.2)* and *explanation on results (5.2).* Seven studies addressed both of the

criteria, and three studies addressed none of them. All in all, the theoretical bases were not convincing in most of the studies. Furthermore, on the level of meta-research in the research field, there are no theory building publications so far. Therefore, a stronger focus on theory in the research field is suggested. Context analyses were not examined in this thesis, but are revealed to be important. Therefore, as already emphasized when discussing the criteria correlations (*chapter 6.1*), mixed methods approaches are recommended for further research.

6.3 Generalizability

This sub-chapter refers to the research question *RQ4*: *What limitations constrain the generalizability?* In fairtrade research, there is a demand for generalizable study results, because it helps stakeholders to take good decisions, for example on the development of the fairtrade system, and to justify them, for example when promoting fairtrade products on the western food market. The generalizability of a study's results is supported by its internal and external validity. But additionally, it also needs several empirical studies with different approaches, as well as theory building research, to draw reliable conclusions and give general policy recommendations.

The fairtrade system is a complex research field. Several approaches were conducted already. This thesis aims to describe the characteristics of the studies but it showed to be difficult to conclude on the generalizability of the research approaches. The examination of the internal and external validity is difficult. Several of the studies offer weak reporting on their exact research approaches. And it was not always possible to distinguish between the quality of the study and the quality of the reporting of the study. That is why the catalogue of criteria, which was made to unify the studies best possible, is rather superficial in the criteria that meet internal and external validity. The analysis therefore forbids a deeper understanding of the details. Thus, the generalizability is constrained by weak research quality and reporting. Moreover, theory building research is missing in the research field.

Besides the initiation of theory building research, further improvement of the research body might be achieved by varying the research design and context. Two possibilities shall be discussed exemplarily. At first, it could be attempted to conduct more (quasi-)experimental designs. There are supporters and opponents of experimental approaches in field research, which was already mentioned in *chapter 3.1*. However, there is only a small number of experimental designs and the research field might be enriched by some more. Secondly, regarding the place of conducting a study, the findings of this review fit those of Nelson and Pound (2009), as well as Oya et al. (2018). Both discovered a bias of studies placed in Latin American and African countries, since fairtrade was introduced and is still mainly practiced in those regions. Nevertheless, the examinations in Asian countries exist and there are several more, for example summarized in the appendix of Siegmann et al. (2019). They were not included into this review out of time reasons, but represent a growing research body in this region.

This is of high importance in order to transfer the insights on the impacts of fairtrade on less examined regions of the world.

The impulse to draw on causal relations in the research field of fairtrade is high, since policy decisions have to be made and consumers expect reliable insights on positive effects of a fairtrade certification. However, in this thesis a weak quality and reporting in the research field is confirmed, which was already criticised by other authors before (Cramer et al. 2014). Therefore, to generalize the results might be wrong or should at least be done with great caution.

6.4 Research Constraints

The approach of this research was accompanied by constraints, and some of them just got revealed during the research process.

In the literature research, all kind of papers were included, namely working papers as well as book chapters, peer-reviewed papers and reports. As the occurrence of significant results over all studies is quite even distributed, a publication bias can be excluded. Though this implicates a variety in publications, it is not a quality indicator for itself. Since the aim of this thesis is to represent the state of research it was reasonable to include all kinds of paper, although the quality of the collected data decreased by it. An additional constraint is that only papers that could be accessed freely got included. There are a few studies that might had fit the topic but were excluded due to missing access, which weakens the aim to provide a comprehensive overview on the research field.

Furthermore, poor reporting of some of the studies lead to a compromise of the details that got examined in the catalogue. To approximately fit all studies, the criteria had to be rather superficial, so that poor reported studies could be reflected as well. This was, for example, crucial for the reporting of the study results (Part 4: The *Data Editing, Analysis and Results)*. Some studies laid out on the non-monetary results quite comprehensively, distinguishing between effects on health, education, or mobility, among others. In contrary, some studies did not mention non-monetary results at all. It was therefore reasonable to distinguish monetary and non-monetary effects, but to not examine them in greater detail.

The different wording of the studies made it challenging to unify the characteristics. In some cases, it was the subjective assessment of the reviewer how to categorize a certain characteristic of a study exactly. When interpreting the results, this subjective influence is to keep in mind. This constraint was also challenging regarding the understanding of the term *workers*. Often workers were referred to with certain terms, like *hired labourers, employees,* or *farm workers*. Details were provided rarely, like kinds of employment (full, part-time or casual, for example) or precise working tasks.

This thesis refers to the fairtrade system only. However, it would also be interesting to examine the impacts of other private certification standards on hired farm workers. Literature on other standards is for example provided by Schuster and Maertens (2016).

With the catalogue criteria *theory mentioned (1.2)*, *hypothesis mentioned (1.3.1)*, and the *structure of questioning (1.5)* it was attempted to discover the logical reasoning beyond the research approach. The criteria are not suitable to draw conclusions on this aim, other than a draft suggestion on the reasonability of the approach (*reasonability 5.3*). A detailed analysis of the mechanisms would be required, which exceeded the time resources for this thesis.

7 Conclusion

This thesis reveals that examinations of the performance of the fairtrade certification system demonstrate ambiguous efficacy of fairtrade in improving worker well-being. The review provides a deeper understanding of the existing studies and how they are carried out. *Chapter 1* includes the introduction to the thesis and the research motivation; the research objectives and research questions are described.

Chapter 2 provides insights into the problem background. The practical details of the fairtrade system are described, as well as the employment situations in rural areas in developing countries and the contribution of the fairtrade system to those employment situations. It is shown that fairtrade developed a certification system for farms in developing countries with certain incentives, but also requirements that have to be met. Fairtrade encounters a labour market in which the employment situations are not always clear or reliable for the workers and written contracts are rare. Fairtrade tries to meet the needs of the workers, e. g. by requiring such written contracts. However, some challenges of the fairtrade system are identified, demonstrating that an easy answer to the question of the impact of fairtrade is not be found.

In *chapter 3,* a literature review provides an overview of the research field and addresses *RQ1*; the studies are briefly described and it becomes obvious that approaches used in different studies, and their results, vary widely. Additionally, a sub-chapter sheds light on the general performance of fairtrade, which again reveals a multifaceted picture of the overall impacts of fairtrade. The current state of meta-research in the field is used to distinguish between independent research and research commissioned by *Fairtrade International*. This shows that hired farm workers are neither the focus of fairtrade nor the focus of the assessments provided within the existing meta-research. Furthermore, the scientific standards of the existing fairtrade research are discussed. Qualitative and quantitative primary and meta-research are addressed, as well as the topic of impact evaluations. All approaches are discussed with a focus on the question whether generalization of studies´ results are possible.

While impact evaluations often seem to lack scientific evidence, qualitative and quantitative research are confronted with the challenge of taking the right amount of context variables into consideration. It is concluded that theory building plays an important role when insights from separate studies are to be generalized. For the topic of generalization, the concepts *internal and external validity* are introduced. It is the theoretical basis for *RQ4*, later addressed in *chapter 6*. Eventually, *chapter 3* provides a discussion about the approach to systematically assess the quality of the studies examined in this thesis. However, the process of analysis revealed that it is not suitable to assess all studies by all of the quality standards, since not all of the information required for the various categorization is provided in each study. The methodology, therefore, is a combination of a quality assessment of each study's approach and also considers the reporting quality of the authors of the studies. A table with the respective quality criteria is provided and prepares for the quality discussion in *chapter 6*.

In *chapter 4*, the methods used in this thesis are discussed. The descriptive review as research frame is derivated. It includes the literature research and points to the systematic coding of the studies. This is done with a catalogue of criteria, which was developed by the author. The catalogue contains the criteria, which are specified by indicators, which consist of categories, which in turn reflect the characteristics of the studies in a unified way (see *Figure 2*). The criteria are explained in detail, providing the basis for the analysis in *chapter 5*. Additionally, the editing of the data is explained. An important adjustment was made, since the studies mostly revealed no strong effects of fairtrade certification on hired farm workers, neither positive nor negative. Therefore, a category was implemented which presents the *positive or negative tendency* of the results.

In *chapter 5* and *chapter 6*, the results of the analysis are presented and discussed, with the help of the *results tendencies*, introduces in *chapter 4*. The author distinguishes between the descriptive results, the examination of possible correlations and the discussion of those criteria that stood out during the analysis. A broad variety of study approaches were noted in the research. It is discovered that mainly workers on plantations were examined, and all those studies reveal results with a positive tendency. In comparison, the studies examining employees working on small farms have a negative tendency (all but one), in demonstrating the efficacy of fairtrade certification to improve well-being of workers. This suggests that scale is required in order for fairtrade to be effective for improving working conditions. However, this could also be a consequence of the fairtrade system which tends to overlook workers hired by small-farmers (a considerable deficiency of the system).

Further correlations between design characteristics and results with a positive tendency are related to cross-sectional studies, a sample size less than or equal to 100, a descriptive or explorative phrasing of the research questions and qualitative methods. Testing approaches, longitudinal studies and a sample size greater than 100, generally indicate a negative tendency of the impact of fairtrade. However, it is

not possible to conclude on those findings with certainty. Whether these positive results are based on the advantages of plantations, or whether they are caused by a certain design is not inferred. Whether testing designs are indeed better suited to reveal grievances or whether they were applied on cases with lower fairtrade performance by accident, is guesswork. However, some patterns seem to exist. In turn, mixed methods proved to be the least likely to involve patterns that would lead to an imbalance between certain design characteristics and certain results. Therefore, for further research, mixed method approaches are recommended.

Besides the analysis of the studies, other main parts of the thesis are, firstly, the discussion whether a unified quality assessment of the studies is actually possible and, secondly, the possibilities and limitations of generalizability of study results. The quality assessment was dependent on the reporting of the respective researchers, which is revealed to be rather incomplete, and on the setup of the catalogue of criteria, which is revealed to be biased, since it is orientated on the studies themselves. However, it is concluded that the studies to some extent meet quality criteria but still have capacities to be improved. The generalizability of the studies is revealed to be dependent on the quality of the studies, especially with regards to the internal and external validity. Furthermore, the generalizability does not emerge from primary research alone, but additionally requires the examination of underlying mechanisms of the impacts of fairtrade on hired farm workers, and the development of evidence-based theories. Such publications are not available in the research field yet. It is therefore highly recommended to contribute with theory building approaches to the research field in the future.

In conclusion, this thesis provides a deeper understanding of the different research approaches in the field of fairtrade's impact on hired farm workers. This implies various characteristics of studies that might bias the results, e. g. the sample size, the time frame of the study, the methodology, and the comprehensiveness of reporting. Further research should focus on the weaknesses of the research field, for example, by improving the reporting quality of studies and using mixed-method approaches.

Fairtrade appears to be performing well in providing decent prices for farmers, though it has its weaknesses when it comes to decent wages for workers. How those weaknesses can be overcome, and what circumstances lead to positive or negative impacts of fairtrade, especially on small farms, must be further examined in future research.

8 Publication bibliography

Arnot, Chris; Boxall, Peter C.; Cash, Sean B. (2006): Do Ethical Consumers Care About Price? A Revealed Preference Analysis of Fair Trade Coffee Purchases. In *Canadian J Agri Economics* 54 (4), pp. 555–565. DOI: 10.1111/j.1744-7976.2006.00066.x.

Basu, Kaushik (2013): The Method of Randomization and the Role of Reasoned Intuition (World Bank Policy Research Working Paper, No. 6722). Available online at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2367082.

Baur, Nina; Blasius, Jörg (Eds.) (2022): Handbuch Methoden der empirischen Sozialforschung. Springer Fachmedien Wiesbaden. 3., vollständig überarbeitete und erweiterte Auflage. Wiesbaden: Springer VS.

Bennett, Elizabeth A. (2018): Voluntary Sustainability Standards: A Squandered Opportunity to Improve Workers' Wages. In *Sustainable Development* 26 (1), pp. 65–82. DOI: 10.1002/sd.1691.

Besky, Sarah (2014): The Darjeeling distinction. Labor and justice on fair-trade tea plantations in India. Berkeley, Los Angeles, London: University of California Press (California studies in food and culture, 47).

Bissinger, Katharina (2019): Essays on the Economics of Fairtrade: Analyzing Food Labels with Sustainability Characteristics. Dissertation. Justus-Liebig-Universität, Gießen. Institut für Agrarpolitik und Marktforschung.

Borenstein, Michael; Hedges, Larry V.; Higgins, Julian P. T.; Rothstein, Hannah R. (2013): Introduction to meta-analysis. [Nachdr.]. Chichester: Wiley.

Bortz, Jürgen; Döring, Nicola (2006): Forschungsmethoden und Evaluation. Für Human- und Sozialwissenschaftler. 4., überarb. Aufl., [Nachdr.]. Heidelberg: Springer-Medizin-Verl. (Springer-Lehrbuch Bachelor, Master).

Brot für die Welt (2024): Zeitstrahl. Available online at https://www.brot-fuer-die-welt.de/themen/dossier-50-jahre-fairer-handel/zeitstrahl/, updated on 5/7/2024, checked on 6/14/2024.

CEval (2012): Final Report Fairtrade Impact Study. Assessing the Impact of Fairtrade on Poverty Reduction through Rural Development. Commissioned by TransFair Germany and Max Havelaar Foundation Switzerland. With assistance of Sarah Klier. Centre for Evaluation. Saarbrücken. Available online at https://www.fairtrade-deutschland.de/fileadmin/DE/mediathek/pdf/fairtrade_impact_study_evaluation_report_en.pdf, checked on 17.06.24.

Chiputwa, Brian; Spielman, David J.; Qaim, Matin (2015): Food Standards, Certification, and Poverty among Coffee Farmers in Uganda. In *World Development* 66, pp. 400–412. DOI: 10.1016/j.worlddev.2014.09.006.

Cooper, Harris M.; Hedges, Larry V.; Valentine, Jeffrey C. (2009): The handbook of research synthesis and meta-analysis. 2. ed. New York: Russell Sage Foundation.

Cramer, Christopher; Johnston, Deborah; Mueller, Bernd; Oya, Carlos; Sender, John (2017): Fairtrade and Labour Markets in Ethiopia and Uganda. In *The Journal of Development Studies* 53 (6), pp. 841–856. DOI: 10.1080/00220388.2016.1208175.

Cramer, Christopher; Johnston, Deborah; Oya, Carlos; Sender, John (2014): Fairtrade, employment and poverty reduction in Ethiopia and Uganda. Available online at https://eprints.soas.ac.uk/20156/1/ftepr%20final%20report%2019%20may%202014%20final.pdf.

Darko, Emily; Lynch, Alainna; Smith, William (2017): The impact of Fairtrade: a review of research evidence 2009–2015. Edited by Overseas Development Institute. Available online at https://odi.org/en/publications/the-impact-of-fairtrade-a-review-of-research-evidence-20092015/.

Davenport, Eileen; Low, William (2015): Fair Trade, Peace and Development in Conflict Zones. In Laura T. Raynolds, Elizabeth A. Bennett (Eds.): Handbook of research on fair trade. Cheltenham: Elgar, pp. 355–371.

Deaton, Angus (2010): Instruments, Randomization, and Learning about Development. In *Journal of Economic Literature* 48 (2), pp. 424–455. DOI: 10.1257/jel.48.2.424.

Denzin, Norman K.; Lincoln, Yvonna S. (Eds.) (2018): The SAGE handbook of qualitative research. Fifth edition. Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: SAGE.

Dragusanu, Raluca; Giovannucci, Daniele; Nunn, Nathan (2014): The Economics of Fair Trade. In *Journal of Economic Perspectives* 28 (3), pp. 217–236. DOI: 10.1257/jep.28.3.217.

Dragusanu, Raluca; Montero, Eduardo; Nunn, Nathan (2022): The Effects of Fair Trade Certification: Evidence from Coffee Producers in Costa Rica. In *Journal of the European Economic Association* 20 (4), pp. 1743–1790. DOI: 10.1093/jeea/jvac026.

Fairtrade Deutschland (2024): Produkte: Fairtrade Deutschland. Available online at https://www.fairtrade-deutschland.de/produkte, updated on 9/18/2024, checked on 9/18/2024.

Fairtrade International (2016): Journeys to Change. Fairtrade Theory of Change. Bonn.

Fairtrade International (2019): Fairtrade Standard for Small-scale Producer Organizations. 03.04.2019_v2.7.

Fairtrade International (2024a): Fairtrade Standard for Hired Labour. 16.04.2024_v 2.1.

Fairtrade International (2024b): About Fairtrade. Available online at https://www.fairtrade.net/about, updated on 6/14/2024, checked on 6/14/2024.

Fairtrade International (2024c): Pricing Table. Available online at https://www.fairtrade.net/standard/minimum-price-info, updated on 6/14/2024, checked on 6/14/2024.

Fairtrade International (2024d): Producer networks and producer organizations. Available online at https://www.fairtrade.net/about/producer-networks, updated on 6/18/2024, checked on 6/18/2024.

Fairtrade International (2024e): How Fairtrade certification works. Available online at https://www.fairtrade.net/about/certification, updated on 6/19/2024, checked on 6/19/2024.

Flick, Uwe (2022): Gütekriterien qualitativer Sozialforschung. In Nina Baur, Jörg Blasius (Eds.): Handbuch Methoden der empirischen Sozialforschung. 3., vollständig überarbeitete und erweiterte Auflage. Wiesbaden: Springer VS, pp. 533–547. Available online at https://link-springer-com.ezproxy.uni-giessen.de/chapter/10.1007/978-3-658-37985-8_34.

Forum Fairer Handel (Ed.) (2023): Aktuelle Entwicklungen im Fairen Handel. 2023. With assistance of Sandra Bäthge, Matthias Fiedler, Katrin Frank, Jette Ladiges, Steffen Weber. Berlin.

Forum Fairer Handel (2024): Materialdatenbank. Available online at https://www.forum-fairer-handel.de/materialien/uebersicht, updated on 2/28/2024, checked on 2/28/2024.

Frundt, Henry J. (2009): Fair bananas! Farmers, workers, and consumers strive to change an industry. Tucson, Ariz.: Univ. of Arizona Press.

Glewwe, Paul (2020): Impact Evaluation in International Development. Theory, Methods, and Practice. With assistance of Petra Todd. 1st ed. Erscheinungsort nicht ermittelbar: World Bank Publications.

Granville, Brigitte; Telford, Steven (2012): Evidence from South Africa. In Brigitte Granville, Janet Dine (Eds.): The Processes and Practices of Fair Trade : Trust, Ethics and Governance. London, UNITED KINGDOM: Taylor & Francis Group, pp. 286–340.

Gröne, Katharina; Braun, Boris; Kloß, Sinah Theres; Schüller, Martin; Bollig, Michael (Eds.) (2020): Fairer Handel. Chancen, Grenzen, Herausforderungen. Gesellschaft für Ökologische Kommunikation mbH. München: oekom verlag.

Guzzo, Richard A.; Jackson, Susan E.; Katzell, Raymond A. (1987): Meta-analysis analysis. In *Research in Organizational Behavior* 9, pp. 407–442. Available online at https://www.researchgate.net/publication/232496069_Meta-analysis_analysis.

Hoebink, Paul; Ruben, Ruerd; Elbers, Willem; van Rijsbergen, Bart (2014): The Impact of Coffee Certification on Smallholder Farmers in Kenya, Uganda and Ethiopia. Edited by Radboud University Nijmegen. Centre for International Development Issues Nijmegen. Available online at https://www.researchgate.net/profile/Urbanus-Mutwiwa/publication/271768026_The_effects_of_coffee_certification_in_Kenya/links/5cac3313a6fdcca26d08e97e/The-effects-of-coffee-certification-in-Kenya.pdf, checked on 13.03.24.

Howse, Robert; Trebilcock, Michael J. (1996): The fair trade-free trade debate: Trade, labor, and the environment. In *International Review of Law and Economics* 16 (1), pp. 61–79. DOI: 10.1016/0144-8188(95)00057-7.

Hurst, Peter (2007): Agricultural Workers and their contribution to sustainable agriculture and rural development. With assistance of Paola Termine, Marilee Karl. Available online at https://www.ilo.org/wcmsp5/groups/public/---ed_dialogue/---actrav/documents/publication/wcms_113732.pdf, checked on 13.03.24.

International Labour Organization (2023): Policy guidelines for the promotion of decent work in the agri-food sector. Meeting of Experts on Decent Work in the Agri-food Sector: An Essential Part of Sustainable Food Systems. Geneva.

International Labour Organization (2024): Agriculture; plantations; other rural sectors. Available online at https://www.ilo.org/industries-and-sectors/agriculture-plantations-other-rural-sectors#intro, updated on 5/29/2024, checked on 6/18/2024.

Jaffee, Daniel (2009): 'Better, but not great': the social and environmental benefits and limitations of Fair Trade for indigenous coffee producers in Oaxaca, Mexico. In Ruerd Ruben (Ed.): The impact of fair trade. Wageningen: Wageningen Acad. Publ, pp. 195–222.

Jodrell, David; Kaoukji, Dwan (2020): Exploring Fairtrade´s Impact. A Review of Research on Fairtrade from 2015-2020. Edited by Fairtrade International. Available online at https://files.fairtrade.net/publications/Fairtrade_Evidence-mapping-full-report.pdf, checked on 2.3.24.

Johannessen, Silje; Wilhite, Harold (2010): Who Really Benefits from Fairtrade? An Analysis of Value Distribution in Fairtrade Coffee. In *Globalizations* 7 (4), pp. 525–544. DOI: 10.1080/14747731.2010.505018.

King, William R.; He, Jun (2006): Understanding the Role and Methods of Meta-Analysis in IS Research. In *Communications of the Association for Information Systems* 16 (16). Available online at https://www.researchgate.net/publication/251874270_Understanding_the_Role_and_Methods_of_Meta-Analysis_in_IS_Research.

Krebs, Dagmar; Menold, Natalja (2022): Gütekriterien quantitativer Sozialforschung. In Nina Baur, Jörg Blasius (Eds.): Handbuch Methoden der empirischen Sozialforschung. 3., vollständig überarbeitete und erweiterte Auflage. Wiesbaden: Springer VS, pp. 549–565. Available online at https://link-springer-com.ezproxy.uni-giessen.de/chapter/10.1007/978-3-658-37985-8_35.

Krumbiegel, Katharina; Maertens, Miet; Wollni, Meike (2018): The Role of Fairtrade Certification for Wages and Job Satisfaction of Plantation Workers. In *World Development* 102, pp. 195–212. DOI: 10.1016/j.worlddev.2017.09.020.

Kuckartz, Udo; Rädiker, Stefan (2022): Datenaufbereitung und Datenbereinigung in der qualitativen Sozialforschung. In Nina Baur, Jörg Blasius (Eds.): Handbuch Methoden der empirischen Sozialforschung. 3., vollständig überarbeitete und erweiterte Auflage. Wiesbaden: Springer VS, pp. 501–516. Available online at https://link-springer-com.ezproxy.uni-giessen.de/chapter/10.1007/978-3-658-37985-8_32.

Kunz, Martin (1987): Dritte-Welt-Läden. Einordnung und Überprüfung eeines entwicklungspolitischen Bildungsmodells anhand der Fallbeispiele der Leonberger und Ludwigsburger Ladeninitiativen. Zugl.: Tübingen, Eberhards-Karls-Univ., Diss., 1987. Wiesbaden, Darmstadt: World University Service; Verl. f. wiss. Publ (Bildung, Repression, Befreiung).

LeBaron, Genevieve (2018): The Global Business of Forced Labour: Report of Findings. Available online at https://www.researchgate.net/publication/325472644_The_Global_Business_of_Forced_Labour_Report_of_Findings.

Lück, Detlev; Landrock, Uta (2022): Datenaufbereitung und Datenbereinigung in der quantitativen Sozialforschun. In Nina Baur, Jörg Blasius (Eds.): Handbuch Methoden der empirischen Sozialforschung. 3., vollständig überarbeitete und erweiterte Auflage. Wiesbaden: Springer VS, pp. 517–532. Available online at https://link-springer-com.ezproxy.uni-giessen.de/chapter/10.1007/978-3-658-37985-8_33.

Maertens, Miet (2019): Fairtrade does not walk the talk. In *Nat Sustain* 2 (7), pp. 549–550. DOI: 10.1038/s41893-019-0332-0.

Makita, Rie (2012): Fair Trade Certification: The Case of Tea Plantation Workers in India. In *Development Policy Review* 30 (1), pp. 87–107. DOI: 10.1111/j.1467-7679.2012.00561.x.

Makita, Rie (2015): Fair trade and plantation workers in Asia. In Laura T. Raynolds, Elizabeth A. Bennett (Eds.): Handbook of research on fair trade. Cheltenham: Elgar, pp. 491–508. Available online at https://www.researchgate.net/publication/282845928_Fair_trade_and_plantation_workers_in_Asia.

Meemken, Eva-Marie; Sellare, Jorge; Kouame, Christophe N.; Qaim, Matin (2019): Effects of Fairtrade on the livelihoods of poor rural workers. In *Nat Sustain* 2 (7), pp. 635–642. DOI: 10.1038/s41893-019-0311-5.

Méndez, V. Ernesto; Bacon, Christopher M.; Olson, Meryl; Petchers, Seth; Herrador, Doribel; Carranza, Cecilia et al. (2010): Effects of Fair Trade and organic certifications on small-scale coffee farmer households in Central America and Mexico. In *Renew. Agric. Food Syst.* 25 (3), pp. 236–251. DOI: 10.1017/S1742170510000268.

Mitiku, Fikadu; Mey, Yann de; Nyssen, Jan; Maertens, Miet (2017): Do Private Sustainability Standards Contribute to Income Growth and Poverty Alleviation? A Comparison of Different Coffee Certification Schemes in Ethiopia. In *Sustainability* 9 (2), p. 246. DOI: 10.3390/su9020246.

Murray, Douglas; Raynolds, Laura T.; Taylor, Peter L. (2006): The future of Fair Trade coffee: dilemmas facing Latin America's small-scale producers. In *Development in Practice*. Available online at https://www.semanticscholar.org/paper/The-future-of-Fair-Trade-coffee%3A-dilemmas-facing-Murray-Raynolds/07981161965d8c7a3110908fe181a01dfba68787.

Nelson, Valerie; Martin, Adrienne (2013): Assesing the poverty Impacts of sustainability Standards. University of Greenwich. Greenwich.

Nelson, Valerie; Pound, Barry (2009): The last ten years: A comprehensive review of the literature on the impact of Fairtrade. Natural Resources Institute. Available online at https://www.nri.org/images/documents/development-programmes/equitable_trade/2010_03_nri_full_literature_review.pdf.

Oi, Walter Y.; Idson, Todd L. (1999): Firm size and wages. In : Handbook of Labor Economics, vol. 3: Elsevier, pp. 2165–2214. Available online at https://www.sciencedirect.com/science/article/pii/S1573446399300195.

Oya, Carlos; Schaefer, Florian; Skalidou, Dafni (2018): The effectiveness of agricultural certification in developing countries: A systematic review. In *World Development* 112, pp. 282–312. DOI: 10.1016/j.worlddev.2018.08.001.

Paré, Guy; Trudel, Marie-Claude; Jaana, Mirou; Kitsiou, Spyros (2015): Synthesizing information systems knowledge: A typology of literature reviews. In *Information & Management* 52 (2), pp. 183–199. DOI: 10.1016/j.im.2014.08.008.

Pawson, Ray (2002a): Evidence-based Policy: In Search of a Method. In *Evaluation* 8 (2), pp. 157–181. DOI: 10.1177/1358902002008002512.

Pawson, Ray (2002b): Evidence-based Policy: The Promise of `Realist Synthesis'. In *Evaluation* 8 (3), pp. 340–358. DOI: 10.1177/135638902401462448.

Przyborski, Aglaja; Wohlrab-Sahr, Monika (2022): Forschungsdesigns für die qualitative Sozialforschung. In Nina Baur, Jörg Blasius (Eds.): Handbuch Methoden der empirischen Sozialforschung. 3., vollständig überarbeitete und erweiterte Auflage. Wiesbaden: Springer VS, pp. 123–142. Available online at https://link-springer-com.ezproxy.uni-giessen.de/chapter/10.1007/978-3-658-37985-8_7.

Raynolds, Laura T. (2012): Fair Trade Flowers: Global Certification, Environmental Sustainability, and Labor Standards. In *Rural Sociology* 77 (4), pp. 493–519. DOI: 10.1111/j.1549-0831.2012.00090.x.

Raynolds, Laura T. (2014): Fairtrade, certification, and labor: global and local tensions in improving conditions for agricultural workers. In *Agric Hum Values* 31 (3), pp. 499–511. DOI: 10.1007/s10460-014-9506-6.

Raynolds, Laura T. (2022): Can certification increase trade fairness and worker empowerment? Lessons from Fairtrade International-certified plantations in Ecuador. In *International Sociology* 37 (6), pp. 716–739. DOI: 10.1177/02685809221103598.

Raynolds, Laura T.; Bennett, Elizabeth A. (Eds.) (2015): Handbook of research on fair trade. Cheltenham: Elgar.

Ruben, Ruerd; van Schendel, Lucie (2009): The impact of Fair Trade in banana plantations in Ghana: income, ownership and livelihoods of banana workers. In Ruerd Ruben (Ed.): The impact of fair trade. Wageningen: Wageningen Acad. Publ, pp. 137–153. Available online at https://brill-com.ezproxy.uni-giessen.de/edcollchap/book/9789086866472/BP000010.xml.

Rubio-Jovel, Karla; Sellare, Jorge; Damm, Yannic; Dietz, Thomas (2024): SDGs trade-offs associated with voluntary sustainability standards: A case study from the coffee sector in Costa Rica. In *Sustainable Development* 32 (1), pp. 917–939. DOI: 10.1002/sd.2701.

Rural21 (2024): Publications. The International Journal for Rural Development. Available online at https://www.rural21.com/english/publications.html, updated on 2/28/2024, checked on 2/28/2024.

Said-Allsopp, Muhaimina; Tallontire, Anne (2014): Enhancing Fairtrade for women workers on plantations: insights from Kenyan agriculture. In *Food Chain* 4 (1), pp. 66–77. DOI: 10.3362/2046-1887.2014.006.

Schuster, Monica; Maertens, Miet (2016): Do private standards benefit workers in horticultural export chains in Peru? In *Journal of Cleaner Production* 112, pp. 2392–2406. DOI: 10.1016/j.jclepro.2015.10.038.

ScienceOpen (2024): Journal of Fair Trade. Edited by Pluto Journals. Fair Trade Society. Available online at https://www.scienceopen.com/collection/689e7113-9888-45e3-ada4-603e59fbedb7, updated on 2/28/2024, checked on 2/28/2024.

Seale, Clive (1999): Quality in Qualitative Research. In *Qualitative Inquiry* 5 (4), pp. 465–478. DOI: 10.1177/107780049900500402.

Sender, John; Cramer, Christopher; Oya, Carlos (2005): Unequal prospects : Disparities in the quantity and quality of labour supply in sub-Saharan Africa. Available online at https://www.semanticscholar.org/paper/Unequal-prospects-%3A-Disparities-in-the-quantity-and-Sender-Cramer/dbfecab5c3bc4384a51f165044c7eb41e9df52d9.

Siegmann, Karin A.; Ananthakrishnan, Sajitha; Fernando, Karin; Joseph, K. J.; Kulasabanathan, Romeshun; Kurian, Rachel et al. (2019): Fairtrade certified Tea in the Hired Labour Sector in India and Sri Lanka: Impact Study and Baseline Data

Collection: Amrita Vishwa Vidyapeetham. Available online at https://www.researchgate.net/publication/346489083_Fairtrade_certified_Tea_in_the_Hired_Labour_Sector_in_India_and_Sri_Lanka_Impact_Study_and_Baseline_Data_Collection.

Snilstveit, Birte; Oliver, Sandy; Vojtkova, Martina (2012): Narrative approaches to systematic review and synthesis of evidence for international development policy and practice. In *Journal of Development Effectiveness* 4 (3), pp. 409–429. DOI: 10.1080/19439342.2012.710641.

Stein, Petra (2022): Forschungsdesigns für die quantitative Sozialforschung. In Nina Baur, Jörg Blasius (Eds.): Handbuch Methoden der empirischen Sozialforschung. 3., vollständig überarbeitete und erweiterte Auflage. Wiesbaden: Springer VS, pp. 143–162. Available online at https://link-springer-com.ezproxy.uni-giessen.de/chapter/10.1007/978-3-658-37985-8_8.

Stockmann, Reinhard (2019): Systematic Evaluation Analysis (SEA) als Instrument der Wirkungsevaluation. Edited by CEval, -Centrum für Evaluation. Universität des Saarlandes. Saarbrücken (Arbeitspapier, 22). Available online at https://www.ssoar.info/ssoar/bitstream/handle/document/62896/ssoar-2019-stockmann-Systematic_Evaluation_Analysis_SEA_als.pdf?sequence=1&isAllowed=y&lnkname=ssoar-2019-stockmann-Systematic_Evaluation_Analysis_SEA_als, checked on 23.07.24.

Stockmann, Reinhard; Meyer, Wolfgang (2014): Evaluation. Eine Einführung. 2., überarb. und aktualisierte Aufl. Opladen, Stuttgart: Budrich; UTB (utb-studi-e-book, 8337).

Terstappen, Vincent; Hanson, Lori; McLaughlin, Darrell (2013): Gender, health, labor, and inequities: a review of the fair and alternative trade literature. In *Agric Hum Values* 30 (1), pp. 21–39. DOI: 10.1007/s10460-012-9377-7.

The Fair Trade Institute (2017): Publications. Available online at https://www.fairtrade-institute.org/publications/?fwp_year_of_publication=1987.00%2C2003.00, updated on 12/29/2017, checked on 2/28/2024.

Thomas, James; Harden, Angela (2008): Methods for the thematic synthesis of qualitative research in systematic reviews. In *BMC Med Res Methodol* 8 (45). DOI: 10.1186/1471-2288-8-45.

Towner, Elizabeth; Dowswell, Therese; Jarvis, S. (1996): Reducing Childhood Accidents - The Effectiveness of Health Promotion Interventions: A Literature Review. Health Education Authority. London.

Valkila, Joni; Haaparanta, Pertti; Niemi, Niina (2010): Empowering Coffee Traders? The Coffee Value Chain from Nicaraguan Fair Trade Farmers to Finnish Consumers. In *J Bus Ethics* 97 (2), pp. 257–270. DOI: 10.1007/s10551-010-0508-z.

Valkila, Joni; Nygren, Anja (2010): Impacts of Fair Trade certification on coffee farmers, cooperatives, and laborers in Nicaragua. In *Agric Hum Values* 27 (3), pp. 321–333. DOI: 10.1007/s10460-009-9208-7.

van Rijn, Fédes; Fort, Ricardo; Ruben, Ruerd; Koster, Tinka; Beekman, Gonne (2020): Does certification improve hired labour conditions and wageworker conditions at banana plantations? In *Agric Hum Values* 37 (2), pp. 353–370. DOI: 10.1007/s10460-019-09990-7.

Wagner, Michael; Weiß, Bernd (2022): Meta-Analyse. In Nina Baur, Jörg Blasius (Eds.): Handbuch Methoden der empirischen Sozialforschung. 3., vollständig überarbeitete und erweiterte Auflage. Wiesbaden: Springer VS, pp. 1117–1126.

Weischer, Christoph (2007): Sozialforschung. 1. Aufl. Konstanz, Stuttgart: UVK; utb GmbH (utb-studi-e-book, 2924). Available online at https://elibrary.utb.de/doi/book/10.36198/9783838529240.

Wikipedia (Ed.) (2023): Stichting Max Havelaar. Available online at https://en.wikipedia.org/w/index.php?title=Stichting_Max_Havelaar&oldid=1165723147, updated on 7/17/2023, checked on 6/14/2024.

9 Appendix

9.1 Appendix 1 – Overview on the Catalogue Criteria with the respective Categories, Excel Abbreviations and Results

Criteria	Indicator/ Indicator question	Categories	Excel Abbreviations	Results
		Frame		
0.1.1 Main author (Pawson 2002)	Who is the main author?	open text	name (one word)	Bennett; Besky; Cramer; Cramer; Dragusanu; Granville; Krumbiegel; Makita; Meemken; Oya; Raynolds; Raynolds; Raynolds; van Rijn; Ruben; Rubio-Jovel; Siegmann; Valkila
0.1.2 Joint authors	What other authors were involved?	open text	open text	none; none; Oya, Johnston, Sender; Johnston, Oya, Sender; Montero, Nunn; Telford; Maertens, Wollni; none; Sellare, Kouame, Qaim; Schaefer, Skalidou; none; none; none; Fort, Ruben, Koster, Beekman; van Schendel; Sellare, Damm, Dietz; Ananthakishnan, Fernando, Joseph, Kulasabanathan, Kurian, Samuel, Viswanathan; Nygren
0.1.3 Publication year (Pawson 2002a)	When was the study published?	year	year (one word)	2018; 2013; 2017; 2014; 2022; 2012; 2018; 2012; 2019; 2018; 2022; 2014; 2012; 2019; 2009; 2023; 2019; 2009
0.1.4 Title	What is the title of the study?	open text	open text	Voluntary Sustainability Standards: A Squandered Opportunity to Improve Workers' Wages; The Darjeeling Distinction: Labor and Justice on Fair-Trade Tea Plantations in India; Fairtrade and Labour Markets in Ethiopia and Uganda; Fairtrade, Employment and Poverty Reduction in Ethiopia and Uganda; The Effects of Fair Trade Certification: Evidence from Coffee Producers in Costa Rica; Empirical Evidence from South Africa; The Role of Fairtrade Certification for Wages and Job Satisfaction of Plantation Workers; Fair Trade Certification: The Case of Tea Plantation Workers in India; Effects of Fairtrade on the livelihoods of poor rural workers; The effectiveness of agricultural certification in developing countries: A systematic review;

Criteria	Indicator/ Indicator question	Categories	Excel Abbreviations	Results
				Can certification increase trade fairness and worker empowerment? Lessons from Fairtrade International-certified plantations in Ecuador; Fairtrade, certification and labour: global and local tensions in improving conditions for agricultural workers; Fair Trade Flower: global Certification, Environmental Sustainability, and Labor Standards; Does certification improve hired labour conditions and wageworker conditions at banana plantations?; The impact of Fair Trade in banana plantations in Ghana: income, ownership and livelihoods of banana workers; SDGs trade-offs associated with voluntary sustainability standards: A case study from the coffee sector in Costa Rica; Fairtrade certified tea in the hired labour sector in India and Sri Lanka: Impact Study and baseline data collection; Impacts of Fair Trade certification on coffee farmers, cooperatives, and laborers in Nicaragua
0.2 Scientific background	Current occupation or department description		(single choice)	
		Economics	Economics	5
		Sociology	Sociology	9
		Interdisciplinary	Interdisciplinary	4
0.3 Financier linked to ft	Is the financier directly linked to fairtrade?		(single choice)	
		yes	1	2
		no	0	12
		n/a	n/a	4
0.4.1 Country (Pawson 2002a)	In what country was the study conducted?	open text	country	all; India; Ethiopia, Uganda; Ethiopia, Uganda; Costa Rica; South Africa; Ghana; India; Cote d'Ivoire; all; Ecuador; Ecuador; Ecuador; Dominican Republic; Ghana; Costa Rica; India, Sri Lanka; Nicaragua
0.4.2 Continent	-	open text	(single choice)	
			Africa	6
			Asia	3
			Central America	3
			Latin America	4
			all	2

Criteria	Indicator/ Indicator question	Categories	Excel Abbreviations	Results
0.5 Product	What product is in focus?	open text	(multiple choice)	
			coffee	3
			cocoa	1
			tea	3
			flowers	3
			bananas	2
			pineapples	1
			wine	1
			tea, coffee, flowers	2
0.6 Main research object (Pawson 2002a)	What is the main research object?	open text	1 keyword	Voluntary sustainability standards; plantations; sites; workers; workers; workers; worker households; small-farm sector; CS-studies; stakeholder, workers; plantations; plantations; workers; workers; producers; workers; workers
Logical Derivation				
1.1 Problem background		open text	keywords	impact of ft; patron-client relations; mechanisms poverty reduction; impact of ft; impact of ft; impact of ft; impact of ft; patron-client relations; small-farm sector; first systematic research; new application; impact of ft; first systematic research; need for empirical evidence; impact of ft; new application; impact of ft; impact of ft
1.2 Theory mentioned (Deaton 2010)	Is a basis theory mentioned?	yes	1	7
		no	0	11
1.3.1 Hypothesis mentioned	Is a hypothesis for the research mentioned?	yes	1	3
		no	0	15
1.3.2 Hypothesis fits questioning	Does the hypothesis fit the questioning?	yes	1	3
		no	0	15
1.4 Research questions (Jodrell and Kaoukji 2020)	-	open text	-	-
1.5 Structure of questioning (Stein 2022; Bortz and Döring 2006)	How are the questions structured?		(single choice)	
		explorative	explorative	3
		descriptive	descriptive	4
		testing	testing	9

Criteria	Indicator/ Indicator question	Categories	Excel Abbreviations	Results
		prognosis	prognosis	0
		evaluation	evaluation	2
Design				
2.1 Study scope	What is the study's scope?		(single choice)	
		case study	case	2
		comparative study	comp	15
		discourse analysis	discourse	1
2.2 Level of experimental design	Is the design (quasi)experimental or an ex-post-facto design?		(single choice)	
		ex-post-facto design	post-facto	17
		(quasi)-experimental design	exp	1
2.3 Time dimension	What is the time structure of the study?		(single choice)	
		cross-sectional study	cross	12
		longitudinal study	long	5
			none	1
2.3.1 Longitudinal Sampling	What kind of longitudinal design is used?		(single choice)	
		cohort study	cohort	2
		panel study	panel	2
		trend study	trend	0
			none	14
2.3.2 Number of Surveys	How many surveys were conducted?	open text	one number or n/a	1; n/a; 3; 2; 9; 1; 2; 1; 1; 1; 1; n/a; n/a; 1; 1; 2; 1; n/a
2.3.3 Time span	How long was the research period?	open text	months, years	4, 2016; 31, 2006-2012; "-", 2010-2013; 3,2011-2012; "-", 2001-2009; 2, 2008; 9, 2014-2015; 2, 2009; "-", 2018; "-", 2018; n/a; n/a; "-", 2010-2011; 4, 2015; 3, 2007; n/a , 2016-2019; 3, 2016; 12, 2005-2006
2.4.1 General randomization	Had a randomization of took place?		(single choice)	
		yes	1	2
		no	0	6
		partly	partly	6
		n/a	n/a	3
			none	1
2.4.2 Worker randomization	Had a randomization took place on the selection of examined workers?		(single choice)	
		yes	1	10
		no	0	6

Criteria	Indicator/ Indicator question	Categories	Excel Abbreviations	Results
		n/a	n/a	1
			none	1
2.4.3 Sample Size	How many persons were examined?	open text	1 number or n/a	16; n/a; 1700; 1700; n/a; 381; 387; 62; 1050; 138; 36; n/a; 36; 369; 100; 434; 4-100% of workforce/ per estate; 62
2.5 Labour Kind	What kind of employment is associated with the examined workers?	open text	(multiple choice)	
			Plantation workers	11
			Small farm employees	3
			Organizational level	2
			Casual wage workers	1
			Unskilled workers	1
			Landless labourers	1
			workers	1
			factory workers	1
2.6.1 Persons examined additionally	What other persons, except workers, were examined in the study?	open text	(multiple choice)	
			Management	7
			NGO	1
			Government	2
			Owner	1
			Stakeholder	2
			Farmer	3
			FT Network	1
			Trader	1
			Industry	1
			Worker representatives	3
			Organizational	2
			producers	1
2.6.2 interested in a positive sector reputation	Can they be suggested to be interested in a positive reputation of their production sector?		(single choice)	
		yes	1	12
		no	0	0
			none	6
2.7 Place of data collection	Where did the surveys take place?		(single choice)	
		private	private	6
		work	work	(2)
				10

Criteria	Indicator/ Indicator question	Categories	Excel Abbreviations	Results
		n/a	n/a none	2
Methods				
3.1 Methods	What methods were used?	open text	keywords	reviews, coding; interviews, observation; interviews, focus groups, questionnaire; questionnaire; questionnaires, data analysis; questionnaires, RCTs; questionnaires; interviews, semi-structured; interviews, questionnaires; meta-analysis, thematic synthesis; document review, interviews, semi-structured; interviews; document analysis, interviews, semi-structured, follow-up conversations; interviews, surveys; "-"; Interviews; focus groups, interviews semi-structured, questionnaire, observation; interviews, semi-structured
3.1.1 Methods Kind	How can the methods be categorized?	qualitative quantitative mixed	(single choice) qualitative quantitative mixed n/a	6 2 9 1
3.1.2 structured/ open regarding workers	Are the workers examined in a structured or in an open form?	structured open	(single choice) structured open both none	8 7 1 2
3.2 Indicators explained	Are indicators of the methods explained?	yes no	1 0	11 7
3.3 Methods fit research question	Do the methods fit the research question?	yes no	1 0	18 0
3.4 Cost-Benefit-Analysis included	Is a cost-benefit-analysis included in the study report?	yes no	1 0	0 18
Data Editing, Analysis and Results				

Criteria	Indicator/ Indicator question	Categories	Excel Abbreviations	Results
4.1 Editing (Kuckartz and Rädiker 2022)	How was the data edited after the collection?	open text	keywords or n/a	n/a; n/a; n/a; n/a; n/a; n/a; n/a; n/a; Stata 14; n/a; n/a; n/a; n/a; n/a; n/a; n/a; software for qualitative data analysis; Atlas-Ti
4.2.1 Analysis (Lück and Landrock 2022)	How was the data analysed?	open text	keywords	descriptive; descriptive; descriptive, regression analysis; regression analysis; estimating hazard models, event study analysis, robustness and sensitivity checks, OLS; pair-wise independent test, chi-sq, Spearman correlation, Kendall tau-b; linear regression model; descriptive; regression analysis; effect size estimates, inverse variance-weighted random effect meta-analysis, bias correction by Hedges' method; n/a; literature analysis; n/a; OSL, Propensity Score Matching, Entropy balancing method; QTA; ANCOVA, IPW, OSL, MHT; code list derived from research questions; n/a
4.2.2 Used Estimators	What estimators were used in the analysis?	open text	key-abbreviations or n/a	keywords; n/a; SE, r^2, b; n/a; SE; correlation coefficients; n/a; n/a; n/a; SMD, CI, p-value; n/a; n/a; n/a; means, SD; means, SD; SD, SE, R^2; n/a; n/a
4.3 Tendency of results	Do the general results of the study have a rather positive or rather negative tendency?		(single choice)	
		positive tendency	pos	10
		negative tendency	neg	7
		negligible	negligible	1
4.3.1 Monetary Results	What results does the study reveal regarding monetary effects?		(single choice)	
		positive	pos	4
		negative	neg	4
		negligible	negligible	8
		no answer	n/a	2
		answer not possible	none	0
4.3.2 non-monetary results	What results does the study reveal regarding non-monetary effects?	positive	pos	9
		negative	neg	5
		negligible	negligible	1
		no answer	n/a	2
		answer not possible	none	1
4.3.3 Premium mentioned	Is the fairtrade premium mentioned, and how?	positively mentioned	pos	10
		negatively mentioned	neg	1
		not mentioned	no	7

Criteria	Indicator/ Indicator question	Categories	Excel Abbreviations	Results
4.4 Significance	Are the results stated to be significant?	yes	1	6
		no	0	6
		partly	partly	2
		none	none	4
4.5 Implications fit questioning	Do the implications fit the kind of questioning?	yes	1	11
		no	0	5
			n/a	2
4.6 Useful in terms of impact evaluation	Could the study be treated as Impact Evaluation?	yes	1	9
		no	0	9
Quality Criteria				
5.1.1 Quality criteria addressed	Are quality criteria addressed in the report?	yes	1	12
		no	0	6
5.1.2 Confounding variables (Jodrell and Kaoukji 2020)	Are confounding variables addressed in the report?	yes	1	15
		no	0	3
5.1.3 Divergent results handling	Is it mentioned, how divergent results were handled?	yes	1	5
		no	0	13
5.1.4 Reflection on internal and external validity (Krebs and Menold 2022)	Is a reflection on internal and external validity included in the report of the study?	yes	1	4
		no	0	14
5.2 Explanation on results	Is an explanation on the results provided?	yes	1	11
		no	0	7
5.3 Reasonability	Are the logical linkages of the study reasonable?	reasonable	reasonable	11
		partly reasonable	partly	6
		not reasonable	not reasonable	1
5.4 Peer-reviewing	Is the study peer-reviewed?	yes	1	11
		no	0	5

Criteria	Indicator/ Indicator question	Categories	Excel Abbreviations		Results
(Jodrell and Kaoukji 2020)				n/a	2
5.5 Practical Relevance stated	Did the authors state their findings to be practical relevant?	yes no		1 0	12 6
5.6 Transparency	How much information is provided in the study report?	counting number of n/a, in relation to the number of possible answers	=100%-ZÄHLENWENN(B3:BB3; "n/a")/SPALTEN(B3:BB3)		98%; 93%; 93%; 94%; 87%; 93%; 96%; 91%; 96%; 96%; 87%; 85%; 89%; 96%; 94%; 94%; 96%; 94%

9.2 Appendix 2 – Template to be filled out for each Study

No.	Characteristic	Indicator	Categories
0	**Frame**		
0.1	Researchers:	Scientific background:	○ Economics 0.2 ○ Sociology ○ Interdisciplinary
0.3	Financier/ Commissioner:	Is the financier directly linked to fairtrade?	O yes O no O n/a 0.3
0.4	Country	In what country was the study conducted?	0.4
0.5	Product	What product is in focus?	0.5
0.6	Main Research Object	What is the main research object?	0.6
1	**Logical Derivation**		
1.1	Problem background (QTA):		Keywords: 1.1
1.2	Theoretical basis (QTA):	Is a basis theory mentioned?	O yes O no 1.2
1.3	Hypothesis (QTA):	Is a hypothesis for the research mentioned?	O yes O no 1.3.1 O yes O no 1.3.2

		If yes, do the research questions fit the hypothesis?			
1.4	Research questions (QTA):	1.5	Research questioning How are the questions structured?	○ Explorative 1.5 ○ Descriptive ○ Testing hypothesis ○ Prognosis ○ Evaluation	
2	**Framework (Stein 2022)**				
2.1	Scope	What is the study´s scope?		○ Case study ○ Comparative study ○ Discourse analysis	2.1
2.2	Level of experimental design	Is the design (quasi-)experimental or with a ex-post-facto design?		○ (Quasi-)experimental ○ Ex-post-facto design	2.2
2.3	Dimension of time	What is the time structure of the study?		○ Longitudinal study ○ Cross-sectional study	2.3
		What kind of longitudinal design?		○ Cohort study ○ Panel study ○ Trend study	2.3.1
		Number of surveys:			2.3.2
		Timespan:			2.3.3
2.4	Sampling	General randomization		O yes O no O partly	2.4.1
		Randomization of workers		O yes O no Sample size:	2.4.2 2.4.3

2.5	Kind of labour	What kind of workers/ labour situation got examined?	☐ Plantation workers ☐ Small-farm employees ☐ Organizational level ☐ ☐	2.5
2.6	Persons examined additionally	What persons got examined additionally?	☐ Management ☐ NGO ☐ Government ☐	2.6.1
		At least one person interested in positive sector reputation?	O yes O no	2.6.2
2.7	Place	Where was the data collected?	Exact place of data collection:	2.7
3	**Method(s) (Baur and Blasius 2022)**			
3.1	Methods: 3.1		O Qualitative methods (QTA) O Quantitative methods (QTA) O Mixed methods (QTA)	3.1.1

3.2	Indicators	Had indicators been mentioned?	O yes O no O partly	3.2
3.3	Methods fit the research question		O yes O no	(3.3)
3.4	Cost-Benefit-Analysis	Is a Cost-Benefit-Analysis included?	O yes O no	3.4
4	**Data Editing and Analysis (Lück and Landrock 2022; Kuckartz and Rädiker 2022)**			
4.1	Editing	What kind of software was used?	○ ○ No answer	4.1
4.2	Analysis, explanation of the model (QTA):	4.2.1	Used estimators:	4.2.2
4.3	Results of the study: How is the impact of fairtrade on hired farm workers?	Monetary impact	○ positive ○ negative ○ negligible	4.3.1
		Non-monetary impact	○ positive ○ negative ○ negligible	4.3.2
		Premium explicitly mentioned?	O positively O negatively O no	4.3.3
4.4	Significance	Are the results statistically significant?	O yes O no O partly	4.4
4.5	Implications (QTA):	Do the implications fit the questioning?	O yes O no O yes O no	4.5 4.6

		Is the study useful in terms of impact evaluation?		
5	**Quality criteria**			
5.1	Design and methods	Have quality criteria been addressed? Are possible confounding variables addressed? Is it mentioned how divergent and inconsistent results had been handled? Had a critical reflection on internal and external validity took place?	O yes O no O yes O no O yes O no O yes O no	5.1.1 5.1.2 5.1.3 5.1.4
5.2	Theoretical explanation (cf. No. 1.2)	Is there an explanation on why the results exist?	O yes O no	5.2
5.3	Logical deduction: The linkages have to be logically consistent 1.1 -> 1.4 -> 2 -> 3 -> 4.2 -> 4.3 ○ reasonable: with 4 or 5 logical linkages ○ partly reasonable: with 2 or 3 logical linkages ○ not reasonable: with 0 or 1 logical linkage		○ reasonable ○ partly reasonable ○ not reasonable	5.3
5.4	Peer-reviewing	Was the study peer-reviewed?	O yes O no	5.4
5.5	Practical relevance: Do the researchers state their findings to be practically relevant?		O yes O no	5.5
5.6	Transparency: How many numbers of this template can be answered?		Numbers answered:	5.6

9.3 Appendix 3 – Results summarized in Excel Table

➔ Please refer to separate excel document "Triebel_masterthesis_appendix_3.xlsx"