Giel Ton
Sam Desiere
Wyste Vellema
Sophia Weituschat
Marijke D'Haese

The effectiveness of contract farming in improving smallholder income and food security in low- and middle-income countries

A mixed-method systematic review

August 2017

Systematic Review 38

Agriculture

3ie International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) is an international grant-making NGO promoting evidence-informed development policies and programmes. We are the global leader in funding, producing and synthesising high-quality evidence of what works, for whom, how, why and at what cost. We believe that using better and policy-relevant evidence helps to make development more effective and improve people's lives.

3ie systematic reviews

3ie systematic reviews appraise and synthesise the available high-quality evidence on the effectiveness of social and economic development interventions in low- and middle-income countries. These reviews follow scientifically recognised review methods, and are peer-reviewed and quality assured according to internationally accepted standards. 3ie is providing leadership in demonstrating rigorous and innovative review methodologies, such as using theory-based approaches suited to inform policy and programming in the dynamic contexts and challenges of low- and middle-income countries.

About this review

The effectiveness of contract farming in improving smallholder income and food security in low- and middle-income countries: a mixed-method systematic review, was submitted in partial fulfilment of the requirements of grant SR6.1088 awarded under Systematic Review Window 6. This review is available on the 3ie website. 3ie is publishing this technical report as received from the authors; it has been formatted to 3ie style, however the tables and figures have not been reformatted. 3ie will also publish a brief of this review, designed for use by decision makers, which is forthcoming. This review will also be published in the Campbell Collaboration Library and will be available here.

All content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors are also the sole responsibility of the authors. Comments or queries should be directed to the corresponding author, Giel Ton, g.ton@ids.ac.uk

Funding for this systematic review was provided by 3ie's donors, which include UK aid, the Bill & Melinda Gates Foundation, Hewlett Foundation and 16 other 3ie members that provide institutional support.

Suggested citation: Ton, G, Desiere,S, Vellema, W, Weituschat, S and D'Haese, M 2017. *The effectiveness of contract farming in improving smallholder income and food security in low- and middle-income countries: a mixed-method systematic review. 3ie Systematic Review 38.* London: International Initiative for Impact Evaluation (3ie).

3ie systematic review executive editors: Edoardo Masset and Beryl Leach
Production manager: Angel Kharya
Assistant production manager: Akarsh Gupta
Cover design: John F McGill and Akarsh Gupta

The effectiveness of contract farming in improving smallholder income and food security in low- and middle-income countries: a mixed-method systematic review

Giel Ton
Wageningen Economic Research

Sam Desiere
Ghent University

Wytse Vellema
Ghent University

Sophia Weituschat
Wageningen University

Marijke D'Haese
Ghent University

3ie Systematic Review 38

August 2017

Plain language summary

Contract farming is a sales arrangement between a farmer and a firm, agreed before production begins, which provides the farmer with resources or services. Many governments and donors promote contract farming as part of agricultural development policies. This systematic review analysed the evidence in the literature on income effects for smallholders. The review included all studies with an econometric design to reduce selection bias in effect estimates. The meta-analysis covered 26 empirical instances of contract farming in 13 developing countries. The contracts varied widely, with varying service packages provided by the firm to the farmers.

When we consider these studies representative for enduring contract farming arrangements, the average income effect would be in the confidence interval of 23 to 54%. However, we show that this estimate is upward biased. We show that non-significant effects are systematically underreported in the literature. Moreover, all studies assessed the effectiveness of the contractual arrangement when these had already survived the start-up problems.

Nevertheless, the findings point to the need for substantial income effects for contract farming arrangements to survive over time. Firms need to offer smallholders above-local market prices, especially in annual crops, and when no cooperative is involved as the intermediary between the firm and the farmers. The poorest farmers are rarely participating in contract farming arrangements; we show that, in 62% of the cases covered by the review, the contract farmers had significantly larger landholdings or more assets than the average farmers in the region.

Executive summary

Background

Contract farming is used by an increasing number of firms as a preferred modality to source products from smallholder farmers in low and middle-income countries. Quality requirements of consumers, economies of scale in production or land ownership rights are common incentives for firms to offer contractual arrangements to farmers. Prices and access to key technology, key inputs or support services are the main incentives for farmers to enter into these contracts. There is great heterogeneity in contract farming, with differences in contracts, farmers, products, buyers, and institutional environments. The focus of this review lies on contract farming, defined as:

> a contractual arrangement for a fixed term between a farmer and a firm, agreed verbally or in writing before production begins, which provides material or financial resources to the farmer and specifies one or more product or process requirements for agricultural production on land owned or controlled by the farmer, which gives the firm legal title to (most of) the crop or livestock — Adapted from Prowse, 2012:12

The last decade shows a rapid increase in studies that use quasi-experimental research designs to assess the effects of specific empirical instances of contract farming on smallholders. The objective of this systematic review was to distill generalised inferences from this rapidly growing body of evidence.

Objectives

The review synthesised the studies in order to answer two questions:

- Question 1: *What is known about the effect size of contract farming on income and food security of smallholder farmers in low- and middle-income countries?*
- Question 2: *Under which enabling or limiting conditions are contract farming arrangements effective for improving income and food security of smallholders*

Search methods

A comprehensive electronic search was applied to Scopus, CAB Abstracts, Econlit, Web of Science, Tropag & Rural, and Agricola between 30 September and 21 October 2015. Snowballing the reference list in review articles and other repositories of research (e.g. worldwidescience.org, FAO, World Bank, Google Scholar) added more studies to the review. The search results were uploaded in EPPI Reviewer 4 and screened for relevance and the rigour of analysis of the effect estimates, in order to combine these results in a meta-analysis of effectiveness. The main terms used to identify the pool of studies within which we expected to find studies that covered the effectiveness of contract farming arrangements were: contract farming, nucleus estate, cooperative, producer organisation, pre-harvest agreement, value chain, farm-firm, outgrower, and vertical integration.

Selection criteria

Each study selected for the meta-analysis was required to resolve the counterfactual, that is, to use a comparison group to mimic the expected situation of farmers not having a contract. When assessing net-effects, the characteristics of groups with or without a contract needed to be fairly similar. Ideally, the only difference was the condition of having a contract or not. Because firms tend to offer contracts to farmers having certain characteristics and farmers self-select when they accept or reject the offer, econometric methods are required to credibly assess the net-effects of contract farming.

To be included in the review, studies needed to analyse the impact of the intervention on income or food security of smallholder farmers. However, only one study was found with food security as an outcome variable (Bellemare and Novak 2016); all other studies included focused on income effects. The review, therefore, has a focus on the income effects of contract farming and a meta-analysis explored this outcome.

Data collection and analysis

The electronic search retrieved 8,529 unique studies. After the full-text screening, 195 studies were found to present research on contract farming. We excluded all papers that did not study the effectiveness of contract farming. The remaining set of papers was referred to as the core set and consisted of 75 studies that presented quantitative outcomes on smallholder farmers. Of the 75 studies in the core set, most did not meet the criteria for methodological and econometric rigour and had to be excluded from the meta-analysis. The meta-analysis was based on data from 22 studies, covering 28 empirical instances of contract farming, two of which had insufficient data to use in the meta-analysis. The studies covered 7,471 respondents.

Results

We applied meta-analysis on the studies that reported income effects. Based on the significance levels and effect sizes, we showed that the set of studies selected for meta-analysis suffered from publication bias. All studies reported at least one empirical instance with a statistically significant positive income effect. Test results suggested that studies with non-significant effects of contract farming are likely to exist but are not reported in the academic literature. Studies also suffer from survivor bias. All studies are cross-sectional studies that assess the effectiveness of the contractual arrangement at one moment in time, but only after the contractual arrangements had been in place for some years when these had already survived the start-up problems. This implies that contractual arrangements that had ceased to function are absent in the literature. The publication and survivor bias detected in this review preclude strong conclusions on the income effects of contract farming arrangements.

The one study that analysed effects on food security reported a positive result (the duration of the hungry season was 8% lower for farmers having a contract). The meta-analysis of the 22 studies showed that the average income effect of participation in these contractual arrangement is highly heterogeneous. The (uncorrected) pooled average effect-size on the proxy for income used in each study, computed in the meta-analysis, indicated a 62 percent increase (95% confidence interval=40%, 87%) in income for contract farmers over incomes of non-contract farmers. However, strong evidence for

publication bias suggests that the true effect of contract farming is likely to be much lower, although still substantially higher than non-contract farming. When we consider these studies representative for enduring contract farming arrangements in general, the pooled average income effect is estimated in 38% (95% confidence interval=23%, 54%).

In almost two-thirds of the studies, the contracted farmers proved to have significantly larger holdings or to be richer than the average farmers in the area. A plausible explanation for this phenomenon, as mentioned in the studies, is that there are lower transaction costs with increasing farm scale and the capacity of the better-endowed farmers to bear the production and post-harvest quality risks inherent to contract farming arrangements. In the four studies in which smaller farmers dominate, the income effects are relatively low.

Authors' conclusions

Contract farming is a container concept that covers a wide range of contractual arrangements. This heterogeneity makes it difficult to draw general conclusions from the literature published on this topic. The studies have a marked publication bias. All studies report at least one case of contract farming that has a positive and statistical significant income effect. Moreover, due to limits inherent to the (cross-sectional) study designs used in these investigations, the estimated effect size is upward biased. The lack of studies on 'failed treatments' leads to an overestimation of the effectiveness of contract farming.

Nevertheless, the results of the meta-analysis suggest that contract farming arrangements need to offer clear incentives to farmers in order to survive over time in the context of free entry and exit of farmers. We generated the hypothesis that relatively large positive effects on income may be a precondition for farmers to continue the contractual arrangements with the firm and give up their autonomy in marketing, production and quality control. High benefits are needed to keep an arrangement attractive and to prevent farmers from dropping out.

Modest expectations and careful planning are needed for contract farming to be effective and sustainable. The practitioner-oriented literature indicated the high risk of failure in the first years and stressed the need for adaptive management and mechanisms to settle disputes. Whereas it is unlikely that contract farming arrangements will on average result in the income effects that we derived from the meta-analysis, it shows the need for substantial income effects for contract farming arrangements to survive over time. If farmers may opt out of a contract - which was the case in all empirical instances covered by the studies except oil palm in Indonesia - those contractual arrangements having low effects are likely to disappear or be amended, and negative effects on smallholder well-being are unlikely.

Contract farming is an institutional arrangement that may be attractive for farmers who want to get access to services or inputs that they cannot obtain in the traditional (spot) market, or reach markets that are more remunerative. Farmers who are able to enter a contract farming arrangement tend not to be the poorest farmers in their region. Both firms and farmers face risks of non-compliance. Relatively larger or richer farmers can cope better with these risks and are, therefore, more likely to take part in a contractual

arrangement. This implies that contract farming is more suited to the relatively better-off segment of the farming population. For annual crops, a price premium seems to be a necessary component of the service package in order to result in high income effects for farmers, especially in situations where no cooperative is involved as an intermediary between the firm and the farmers.

Contents

Plain language summary ..... **i**
Executive summary ..... **ii**
List of figures and tables ..... **vii**
1. Background **1**
1.1 The issue 1
1.2 The intervention 2
1.3 How the intervention might work 3
1.4 Why it was important to do the review 7
2. Objectives **8**
3. Methods **8**
3.1 Criteria for considering studies for this review 8
3.2 Search methods for identification of studies 14
3.3 Data collection and analysis 15
4. Results **22**
4.1 Description of Studies 22
4.2 Risk of bias in effect estimates 33
4.3 Risk of bias in research synthesis 42
4.4 Synthesis of effects 46
5. Discussion **67**
5.1 Summary of main results 67
5.2 Overall completeness and applicability of evidence 69
5.3 Quality of the evidence 69
5.4 Limitations and potential biases in the review process 70
5.5 Agreements and disagreements with other studies or reviews 71
6. Authors' conclusions **72**
6.1 Implications for practice and policy 72
6.2 Implications for research 74
Appendix **76**
References **90**

List of figures and tables

Figure 1: Process of establishing a contract farming arrangement 3
Figure 2: Contractual arrangements as a result of a negotiation process 4
Figure 3: Outcome areas that result from the contractual arrangement 6
Figure 4: Sources of included studies 23
Figure 5: Funnel plot 43
Figure 6: Funnel plot: replacing four effect sizes by its synthetic effect size 44
Figure 7: Funnel plot of a trim-and-fill analysis 45
Figure 8: Forest plot of the effect of contract farming on yields. 47
Figure 9: Forest plot of the effect of contract farming on household labour. 48
Figure 10: Forest plot of the effect of contract farming on prices. 49
Figure 11: Forest plot of the effect of contract farming on income: all empirical instances 50
Figure 12: Forest plot of the effect of contract farming on income: synthetic effect for Narayanan, 2014. 51
Figure 13: Forest plot of the effect of contract farming on income per crop type. 55
Figure 14: Forest plot by level of service provisioning 57
Figure 15: Average landholding in selected studies 60

Table 1: Overview of the studies included in the research synthesis 28
Table 2: Risk of bias assessment 32
Table 3: Details of risk assessment of 'selection mechanism' and 'group equivalence' for PSM studies 35
Table 4: Details of risk assessment of 'selection mechanism' and 'group equivalence' for IV/Heckman studies 37
Table 5: Details of risk assessment of 'selection mechanism' and 'group equivalence' for endogenous switching regression models 38
Table 6: Sampling strategy used in the studies 41
Table 7: Meta-regression: examining publication bias based on Egger’s test 43
Table 8: Timing of impact evaluation and start of contract farming 46
Table 9: Sensitivity analysis 53
Table 10: Moderator analysis 54
Table 11: Overview of the heterogeneity of service provisioning in the empirical instances covered by the meta-analysis 56
Table 12: Meta-regression on moderator and sensitivity variables for income effects of contract farming 58
Table 13: Assessment of differences in scale and wealth between participants and non-participants 61
Table 14: Dataset of cases and conditions for the qualitative comparative analysis 62
Table 15: Combination of services and incentives embedded in the contracts 64
Table 16: Truth-table with all contract farming arrangements in the sample 65
Table 17: Truth-table of contract farming related to perennial crops 66
Table 18: Truth-table of contract farming related to annual crops 66
Table 19: Truth-table with contract farming arrangements related to animal husbandry 67

1. Background

1.1 The issue

Contract farming is a commercial relationship between a firm and a group of farmers. It is a business model in which farm products are bought in advance by a firm in exchange for certain services and other benefits. Although principally a commercial initiative, contract farming is considered to be a way to overcome the challenges that small farmers face when linking to remunerative markets. It assists farmers in connecting to output markets and often provides inputs, credit, or agricultural extension (Da Silva and Rankin 2013, Eaton and Shepherd 2001, World Bank 2007). These services can be provided not only by private firms, but can also come from, or be facilitated by, multi-actor partnerships between companies, governments and NGOs (Prowse 2012). Estimates of the incidence of contract farming in developing countries are unreliable and differ markedly between countries; they are generally below 10% of total area under production (Minot and Ronchi 2015). Unfortunately, there are no reliable data at national levels on the incidence of contract farming to assess its relevance in agrarian change (Oya 2012).

Nevertheless, a rapidly growing number of firms – at least in modern market channels – are relying on contracts for the procurement of products from preferred suppliers (Da Silva and Rankin 2013, Reardon and Berdegue 2002). Modern market channels pose higher demands to value chain coordination and traceability than traditional markets. Contract farming is one of the institutional arrangements available to organise this chain coordination. It is an alternative to centralised governance systems with complete control by the firm, such as plantation production (Bijman 2008), especially in countries in which firms face constraints in access to land, or face high risks in production.

Generally, companies offer a contract only to those farmers who comply with some minimum requirements (for example land ownership, irrigated lands, minimal plot sizes). Even if these arrangements are beneficial to farmers directly or indirectly through spill-over effects, there will be heterogeneity in impacts, with certain farmers benefitting more than others, with some even losing out. It is clear that contracts will not be randomly distributed within a farming community, and contracted farmers will always have special characteristics; a situation referred to in the literature as firm-selection and self-selection biases (Barrett, Bachke, Bellemare et al. 2012, Minot and Ronchi 2015). Studies that infer quantitative effects of contract farming on income and food security need a proper research design to control for these biases.

Contracting firms are almost always relatively large processors, exporters, or supermarket chains. Rarely do small-scale traders, or even wholesalers offer farmers pre-planting contracts. This is not surprising, given the large fixed costs associated with contracting (Minot and Ronchi 2015). Contract farming is induced by a firm’s need to source products with specific qualities and in sufficient quantities and is more likely to be established in the presence of appropriate geographical and political-economic conditions as well as an enabling business environment (Jia and Bijman 2013). Relevant geographical conditions are, for example, road infrastructure, access to water, soil types and climatic conditions. Public policies and institutions influence the forms of transaction used by farmers and firms. Relevant political-economic conditions are land-rights

policies, market regulation, trade policies and a low risk of socio-economic shocks. Whether a firm chooses to start offering a contract farming arrangement is also highly influenced by the local business environment, such as financial services, conflict resolution systems, investment subsidies, business development services, brokering services, and farmer organisations. Political decision-making influences many of these conditions. Therefore, policy makers can enable or constrain the opportunities for contract farming, influencing its attractiveness to firms and farmers as a way of organising transactions and embedded services. Since 2007, coinciding with the investor rush for land in sub-Saharan Africa, international development agencies have increasingly presented contract farming as an alternative or complementary development opportunity for smallholder inclusion (Lindholm 2014).

Contract farming is considered by most authors to be a positive development for agricultural innovation in developing countries, improving the inclusion of farmers in markets (Eaton and Shepherd 2001, Minot 1986). Yet, there is serious concern whether smaller farmers benefit from these arrangements, because the relative size of buyers may result in an unequal power relationship, which influences the terms of the arrangements (Sivramkrishna and Jyotishi 2008, von Hagen and Alvarez 2011). In the earlier literature on contract farming, the issue of power imbalances was especially prominent (Glover and Kusterer 1990, Little and Watts 1994), and the discussion was rather polarised between proponents and critics of contract farming (Oya 2012).

Most of these earlier studies compared the incomes of participating farmers with those not participating, or compared incomes before and after the contract was signed. Such direct comparisons of averages have a high risk of suffering from selection bias or other confounding factors (e.g. weather or prices), rather than reflect the results of the contract farming arrangement itself. The last decade shows a rapid increase of studies on contract farming that assess the effects of contract farming using stronger econometric research designs, which provide more reliable estimates of net-effects. These econometric impact evaluations report mixed effects. Therefore, a systematic review of this rapidly growing body of evidence is timely and may help to distill generalised inferences from these specific instances.

This report presents the results of a systematic review of the effectiveness studies that present primary data and were published up to the summer of 2015. To do so, our review followed a two-stage process. First, it identified impact studies that applied a research design able to reduce selection bias and assess the counterfactual situation (What would have happened to the smallholder farmers had there been no contract farming arrangement?). Secondly, it placed these contractual arrangements in context. More specifically, the qualitative synthesis mapped the relevant contextual conditions for each empirical instance of contract farming covered in these rigorous effectiveness studies and made a case-based comparative analysis to identify enablers and barriers of effectiveness, to distill recommendations for policy makers and practitioners.

1.2 The intervention

Contractual arrangements in agriculture are extremely diverse, with varying embedded services, credit arrangements, payment systems and price-setting mechanisms. Contract farming is defined very broadly as ‘agricultural production carried out according to an

agreement between a buyer and farmers, which establishes conditions for the production and marketing of a farm product or products' (FAO 2008). There are many different types of contracts, going from full resource provisioning contracts with detailed production and marketing conditions to mere verbal agreements to buy whatever quantity is produced at the going market price. In this review, we only consider studies on contractual arrangements that have a service provisioning component, next to a marketing agreement. We use a definition of contract farming partly based on Prowse (2012), which covers contracts in which the farmer is provided with seeds/breeds, inputs and/or credit. Contract farming is defined as *"a contractual arrangement for a fixed term between a farmer and a firm, agreed verbally or in writing, before production begins, which provides material or financial resources to the farmer and specifies one or more product or process requirements, for agricultural production on land owned or controlled by the farmer, which gives the firm legal title to (most of) the crop or livestock"* (adapted from Prowse [2012:12])

Input-providing contracts are believed to be beneficial for the poorest farmers, as they lack the financial capacity to invest in these inputs themselves and they do not have access to finances that would enable them to buy these inputs on credit (Key and Runsten 1999, Minten, Randrianarison and Swinnen 2009, Schipmann and Qaim 2010, Vorley and Proctor 2008).

1.3 How the intervention might work

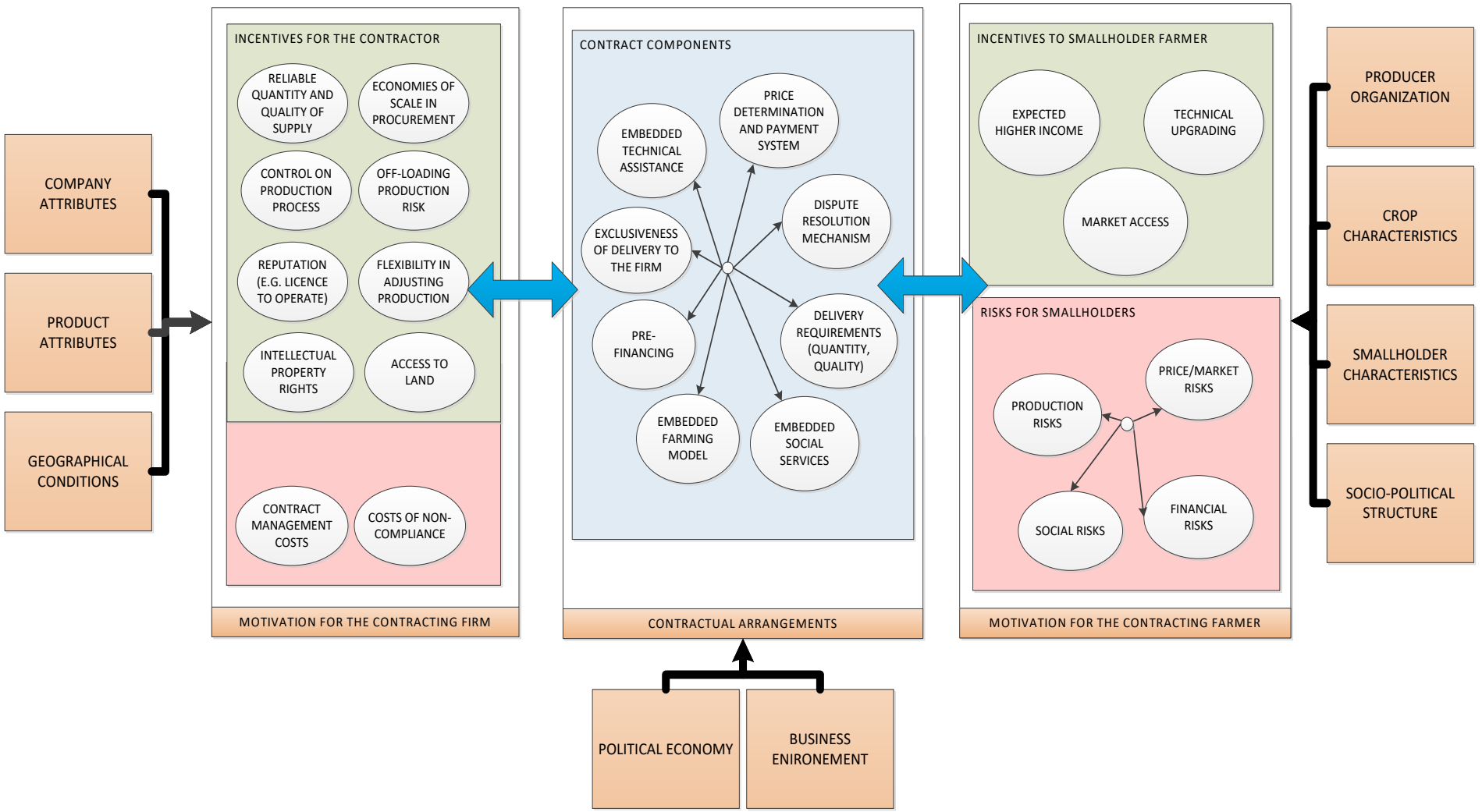
1.3.1 Contract farming arrangements as transactions

Following Barrett et al. (2012), we identify four stages of decision-making by the parties that are involved in contracting (Figure 1).

Figure 1: Process of establishing a contract farming arrangement

In stage 1, a firm chooses its procurement location; next, in stage 2, the firm proposes a contract to a (specific group) of farmers, which these farmers may or may not accept. The firm may deliberately restrict the contract only to those farmers complying with specific characteristics. Stage 3 is the acceptance of the contract details by the farmer(s). Finally, in stage 4, the firm and smallholder decide to honour the contract for its duration. A farmer organisation, (local) government, financial institution or NGO may play a role in any of the steps in this process. After each agricultural cycle or contract period, the parties may decide to either stop or renew the arrangements. The possibility to opt-out is dependent on the fixed investments made by each of the contracting parties and the alternative use of these investments outside the contractual arrangement (asset specificity). For example, discontinuation of a contract with a firm after having planted perennial crops or installed highly specialised buildings and equipment will imply higher costs than when stopping production of a vegetable crop requiring only variable inputs. Furthermore, the availability of alternative output markets (e.g. spot markets) and the availability of good quality inputs outside the contract (seeds/breeds, agrochemicals, credit) reduces the switching costs for smallholders in case the contractual arrangements does not yield the expected benefits (Mitchell, Keane and Coles 2009, Warning and Key 2002).

Figure 2: Contractual arrangements as a result of a negotiation process



1.3.2 Immediate outcomes

The most obvious immediate outcomes of contract farming arrangements are related to the uptake and renewal of contracts. The continuation of a contractual arrangement is an indicator of the overall satisfaction of the firm and the farmers with their contractual arrangement. Figure 2 illustrates the incentive structure for firms and farmers to decide to enter and renew a contractual arrangement. Effects on smallholders will be greater for contract farming arrangements that are in place for a longer time because investments in productive assets and knowledge take time to bear fruit. Moreover, it is likely that unsuccessful farmers will have left the arrangement.

Research by Narayanan (2013) in India shows that farmers may move in and out of contracts quite regularly, thus indicating a potential difference between observed short-term dynamics for individual farmers and a more structural positive effect on the (local) institutional environment. In time, contractual arrangements between firms and farmers may become a normal business option in a region, and part of the palette of options available to farmers for organising their farms.

1.3.3 Intermediate outcomes

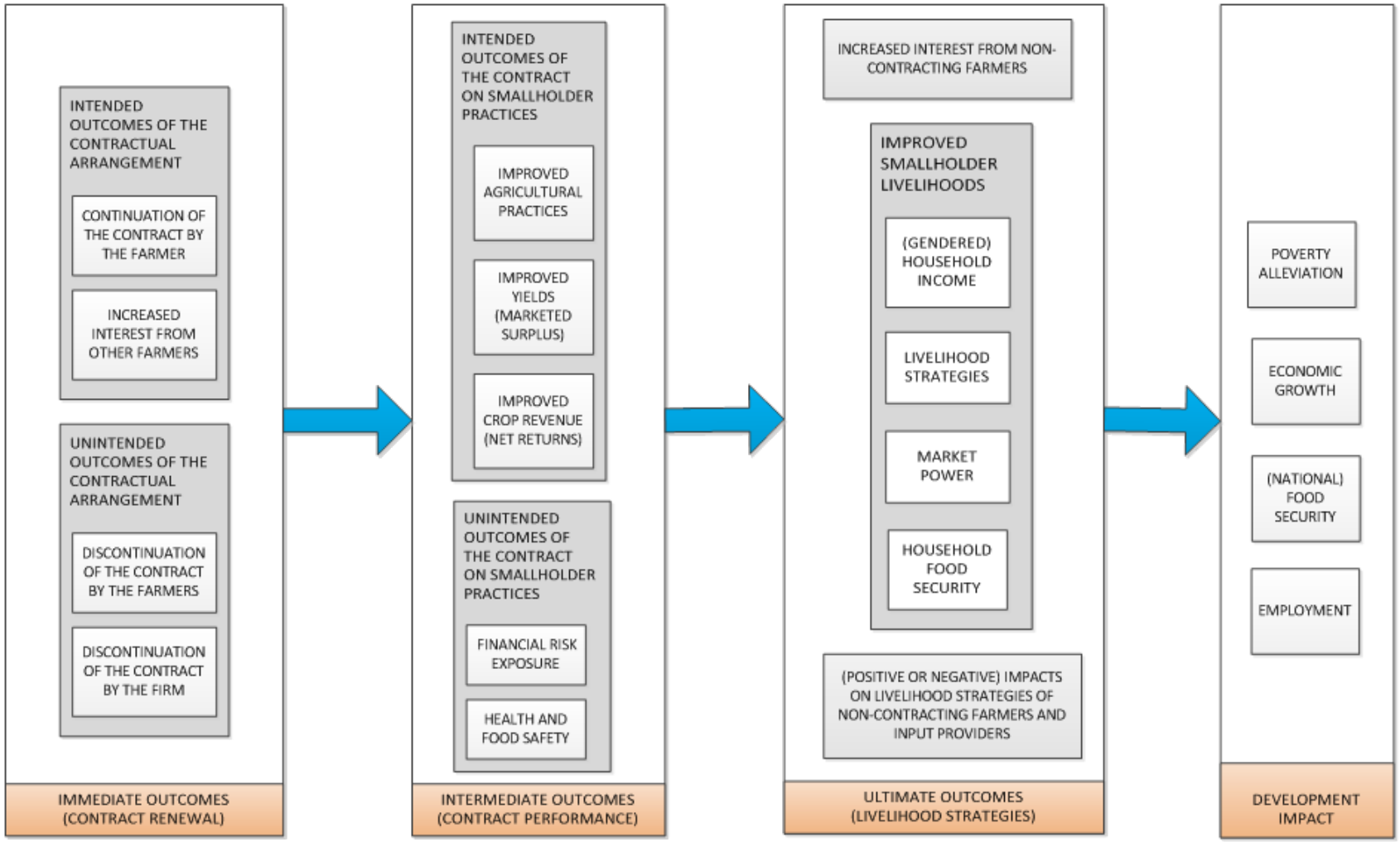
Intermediate outcomes are indicators of contract performance and productivity. These intended effects of the contractual arrangements are directly related to the contracted production, for example, improved agricultural practices, improved yields, better quality, and improved crop revenue. Non-contracted farmers may experience spill-over effects (e.g. information and technology obtained from neighbours). The technology spill-over effects can be expected to be larger if the contractual arrangements in a certain location are continued for several years.

Effects on intermediate outcomes can be visible over a shorter time-span than changes in ultimate outcomes (income and food security). It is nevertheless more challenging to compare the intermediate outcomes between different studies because the indicators used to measure these outcomes are more likely to be context- and product-specific and thus rely on different constructs and ways of measurement.

1.3.4 Ultimate outcomes and development impact

The ultimate outcomes of contract farming arrangements considered here were income and food security of smallholder farmers. The causal link between changes in these ultimate outcomes and the contract farming arrangements is not always clear-cut, especially when contract farming covers only a small part of a farmer’s agricultural activities. Few studies provided empirical evidence on development impact, e.g. poverty alleviation, economic growth, national food security and employment. Even though no study quantified the effects at this outcome level, most studies did claim narratively, in their introductions and conclusions, that these impacts exist.

Figure 3: Outcome areas that result from the contractual arrangement



1.4 Why it was important to do the review

The body of literature on contract farming is growing rapidly. However, direct comparisons between studies are complex because of the large heterogeneity in contract farming, with differences in contracts, farmers, products, buyers, and institutional environments. Paradoxically, it is this heterogeneity which makes a systematic review of these studies particularly useful.

Our review synthesised the findings of research that measured the net-effects of contract farming on smallholder farmers, identified commonalities and differences between these contractual arrangements, and explored the evidence for lessons learned that might improve current and future contract farming arrangements. Policymakers are increasingly interested in working with the private sector in development initiatives, using co-investment or developing joint projects. Therefore, this review not only aimed to map the evidence about the average effects that contract farming has had on income and food security of smallholder farmers, according to studies with a strong research design, but it also aimed to show what seems to be working where, when, and under which conditions. In all the instances of contract farming included in the meta-analysis, we searched for information on the service package provided to farmers, the role of social capital in the start-up of the arrangement, and if a price-premium was paid vis-a-vis production for the local market. This part of the study was explorative in nature, in order to identify potential (combinations of) enabling or limiting conditions that influence the effectiveness of contract farming. This review was more systematic and detailed than earlier reviews (Minot and Ronchi 2015, Prowse 2012, Wang, Wang and Delgado 2014b), which used document counting and variable-based analysis to generate generalised inferences on effectiveness. We report the results of a meta-analysis, weighing the studies according to their effect-size, sample size, and variance, following the process defined by the Campbell and Cochrane Collaboration (Higgins and Green 2011), and we added configurational comparative analyses (Rihoux and Ragin 2009) to find configurations of conditions that predict effectiveness.

2. Objectives

The systematic review had two primary research questions to guide the review and synthesis process, with six sub-questions.

Question 1: What is known about the effectiveness of contract farming on smallholder farmers in low- and middle-income countries?

1.1. What are the effects of contract farming on income and food security of smallholder farmers?

1.2. What are the effects of the contract farming arrangement on intermediate outcomes such as yields and/or net-returns derived from producing the contracted crop or livestock?

1.3. What are the effects on intermediate and ultimate outcomes for non-participating neighbouring farmers living in the same communities as the contracted farmers?

1.4. What are the drop-out rates and side-selling associated with the contract farming arrangement (immediate outcomes)?

Question 2: Under which enabling or limiting conditions are contract farming arrangements more effective?

2.1. What are the enabling and limiting conditions associated with the effectiveness and continuity of the contract farming arrangements for smallholder farmers?

2.2. Are there configurations of conditions that may enable success or failure of contract farming arrangements?

3. Methods

3.1 Criteria for considering studies for this review

The review followed the search and screening process as defined in the protocol, published as Ton et al. (2015). The criteria used to select studies are described below following the PICOS format (Participants, Interventions, Comparisons, Outcomes, Study designs). Selected studies were used to identify the empirical instances on which the quantitative and qualitative analyses were based. The same studies were eligible to answer review question 1 (summative evidence on effectiveness) and review question 2 (enabling and limiting conditions for effectiveness).

3.1.1 Types of participants

The review focused on the effects of contract farming on smallholder farmers. Generally, smallholders have at least two of the following criteria proposed by Stewart et al. (2014):

1. Limited size of farm (as compared to other farms in the sector).
2. Mostly dependent on family labour.
3. Subsistence farming or mix of subsistence and market-oriented farming.
4. Reportedly limited resources in terms of land, technical and technological support, and/or capital for maintenance and investment.

The land needed to qualify as smallholder depends per location and crop-type, with large variation even within one country resulting from agro-climatic conditions. Therefore, during the screening, we did not exclude studies based on their definition of smallholders. All selected studies relate to contracts with such farmers in low-income economies, lower-middle income economies and upper-middle income economies. We used the classification according to the World Bank.

3.1.2 Types of interventions

We only included studies that covered contractual arrangements where a firm provided services to farmers next to a price for their products. In the screening procedure, studies were excluded if related to:

- Contractual arrangements without service-providing clauses. This excluded commercial contracts such as forward sales and price hedging. These price stabilisation strategies are common practice in agro-export crops, such as coffee, cocoa, soybeans, and corn. Though they use contracts, it is not relate to a contractual relation between a procuring firm and supplying farmers.
- Contractual arrangements only concerning marketing, such as collective marketing, marketing boards, and preferred suppliers to supermarkets. We only included these in the review when additional services were provided.
- Traditional sharecropping arrangements in which a tenant farmer was provided with inputs and works the land for the owner in exchange for an agreed share of the value of the crop minus charges. We only included studies where farmers had land that was not owned by the firm with which they agreed the contract.
- Certification schemes such as UTZ Certified, Fair Trade, and Rainforest Alliance unless a non–transferable fixed-term forward sales contract with a specific firm was offered.
- Hybrid situations, such as when a contract between a cooperative and a fixed buyer specified service-delivery of the cooperative to its members, have only been included when the decision to enter into the contract was optional for individual members.
- Contractual arrangements outside the realm of ‘smallholder agriculture’, such as timber exploitation and marine fishery.

3.1.3 Types of comparisons

Contract farming is a complex intervention, consisting of a package of services provided to a farmer in exchange for a pre-planting sales agreement. Arrangements may provide farmers with various packages: e.g. higher than local prices, inputs, credit, agricultural extension, and transport. If any of these services in the package were also available to the comparison group, we considered the net- effect of this differential service package, comparing the outcome in the treatment and comparison group.

In many instances, comparison-group farmers were located in the same region as treatment-group farmers, creating a potential for spill-over effects. Studies which did not explicitly controlled for spill-over effects were accepted in the meta-analysis but mapped as having a higher risk of bias (see section 4.3).

3.1.4 Types of outcomes

Primary outcomes included in the review are income and food security. Income may be measured by farm household reporting of crop income, farm income, household income

or household expenditure, for example. No studies were excluded based on the way in which they measured income, although there were substantial differences between the proxy-indicators used by various authors. Definitions of income differed in three ways: whether expenditure or income was measured, if the costs of inputs or hired labour were discounted or not, and whether total household income, farming income, or only the net-returns from the contracted crop or livestock was considered.

Multiple proxy-indicators exist for food security, such as length of the hungry season, food consumption recall, diet composition or anthropometric measures (Masset, Haddad, Cornelius et al. 2011, Webb 2013). No studies were excluded based on the definition of food security outcome indicators used.

Secondary outcomes used in the studies were yield, price and labour use. No study was excluded based on their definition of secondary outcomes. However, we only included secondary outcomes included in studies that also reported primary outcomes. As the aim of the review was to analyse whether contract farming may benefit smallholder farmers, the ultimate outcomes considered were income and food security. The immediate outcome indicators, side-selling and drop-out rates were not available in any of the selected studies.

3.1.5 Types of study design

In the search strategy and the initial title-abstract screening, no studies were excluded based on the study design, in order to get a fairly complete set of quantitative and qualitative contract farming studies. From this set of substantive-relevant studies, a selection was made based on study design, to identify the 23 studies that contained a quantitative assessment of the effectiveness of one or more empirical instance of contract farming and met the criteria for methodological rigour described in the next section.

We included all studies with a credible design to reduce the risk of selection bias. This includes experimental studies (Randomised Controlled Trials - RCTs) and quasi-experimental design. We include quasi-experimental designs that used statistical matching (e.g., propensity score matching or PSM, or covariate matching), regression adjustment (e.g., difference-in-differences or DID, and single difference regression analysis, instrumental variables or IV, estimation and Heckman selection models), as well as similar cross-sectional or longitudinal designs.

Selection can be the result of decisions by either the firm or the farmer. It is common practice for the company to prefer those farmers which it believes can offer the right quality at the right price. Large farms in easy-to-reach locations are more likely to be selected. Often, the criteria used by firms to select farmers are a combination of observable and non-observable characteristics (Bellemare 2012). Farmers who are offered a contract can decide whether or not to take it. Those farmers who expect to benefit most from the contract are more likely to accept the offer. This decision is at least partly based on unobservable characteristics such as ability, dedication, and knowledge. To be selected for the meta-analysis, effectiveness studies need to have a design that addresses the issue of selection bias. A simple comparison of outcomes between contracted and non-contracted will therefore almost never retrieve the causal impact of contract farming. More complicated strategies are needed and have been proposed in

the vast literature on impact evaluations (Imbens and Wooldridge 2008). These research designs have strengths and weaknesses, and their validity and appropriateness will depend on the context and dynamics of each empirical instance. Also, their validity depends on the way that the research methods were implemented in these designs (e.g. the quality of the Instrumental Variables or matching algorithms that are used).

Analysing the validity of the statistical conclusions involved a critical scrutiny of the econometric methods used. In general, the net-effects of contract farming on income can be conceptualised as a system of three equations that guide the estimation procedures:

$$y_{0i} = \alpha_0 + \beta_0 X_i + \varepsilon_{0i}$$

$$y_{1i} = \alpha_1 + \beta_1 X_i + \varepsilon_{1i} \quad (1)$$

$$D_i = 1\,(\gamma Z_i + v_i > 0)$$

Farmer i can choose to operate under two different regimes: the farmer can participate in contract farming (regime 1) or the farmer can sell his or her produce on local markets (regime 0). With each regime, the farmer obtains a different outcome (y_i), which is a function of observable household characteristics (X_i) and unobservable characteristics captured by the error term ε. The choice between the two regimes depends on observable and unobservable characteristics of the farmer and the contract (Z_i). The indicator function, D, equals 1 if the farmer participates in contract farming and 0 otherwise.

This set of equations can be summarised by the following switching regression:

$$y_i = \alpha_0 + D_i(\alpha_1 - \alpha_0) + \beta_0 X_i + D_i(\beta_1 - \beta_0)X_i + \varepsilon_{0i} + D_i(\varepsilon_{1i} - \varepsilon_{0i}) \quad (2)$$

Most papers implicitly assume that contract farming only has a ‘level’ effect, but that the return on the observable characteristics, X_i, does not depend on the regime. In other words, it is assumed that $\beta_1 = \beta_0 = \beta$. This assumption simplifies equation 2 substantially and facilitates the empirical estimation. Whether this assumption is justified depends on the details of the contractual arrangements (Bolwig, Gibbon and Jones 2009, Narayanan 2014). Contractual arrangements that only change the channel through which farmers sell their produce, from spot markets to a firm, are unlikely to significantly alter marginal returns on land, labour and other inputs. Contractual arrangements that introduce a new crop that is not sold in local markets are, on the other hand, more likely to imply a regime switch. For instance, if contracted farmers switch from producing staple crops for local markets to cash crops for export such as cotton, production factors such as irrigation may become more important and profitable.

If the assumption of constant returns on inputs across regimes holds, the switching regression equation simplifies to:

$$y_i = \alpha_0 + D_i(\alpha_1 - \alpha_0) + \beta X_i + D_i(\varepsilon_{1i} - \varepsilon_{0i}) + \varepsilon_{0i} \quad (3)$$

This equation neatly summarises the selection problem that haunts causal identification of the impact of contract farming. As we cannot assume that farmers participate randomly in contract farming, the unobservable term $D_i(\varepsilon_{1i} - \varepsilon_{0i})$ will be correlated with the treatment. Hence, estimating this equation with OLS, while omitting the term $D_i(\varepsilon_{1i} - \varepsilon_{0i})$, will give biased estimates. In other words, selection into contract farming is endogenous. For example, a very able farmer may be more likely to participate in contract farming and always obtains, at the same time, higher outputs than a less able farmer, even if he would not participate in contract farming. In this case, OLS overestimates the impact of contract farming as it does not account for 'farming ability' which is unobserved by the researcher.

The simplest approach to estimate equation 3 consistently is to include all factors that determine participation in contract farming as explanatory variables. Conditional on these variables, participation in contract farming is then assumed to be random. In exceptional cases, this assumption is met. For instance, if the firm decides which farmers are eligible for contract farming based on clear criteria known and observable to the researcher and all farmers accept the offer, one can perfectly control for selection into the program and equation 3 can consistently be estimated with OLS. In general, however, some of the factors that determine participation in contract farming (such as risk aversion and farming ability) are unobservable. Hence, more advanced strategies are required to deal with bias due to self-selection.

The gold standard in net-effect studies, as codified by most systematic review boards such as the Campbell and Cochrane Collaboration –though contested by others (Cartwright 2007, Deaton and Cartwright 2016, Ravallion 2009)- is Randomized Controlled Trials (RCTs), in which farmers are randomly assigned to treatment and control groups. This random assignment assures, by design, that the error term $(\varepsilon_{1i} - \varepsilon_{0i})$ is zero and therefore allows the estimation of equation 3 with OLS. The implementation of a RCT for contract farming is rarely feasible. It would require a firm that accepts a research design for out-rolling its contract offer. For instance, the firm could offer contracts in a predefined number of villages (the treatment group), while not offering it in neighbouring villages (the control group). Data could then be collected before the implementation of the contract in all villages (the baseline) and after the implementation of the treatment (the end line).

Panel data are considered as the second best approach to establishing the causal impact of contract farming on welfare. Panel data allow us to control for time-invariant household characteristics, such as farming ability, that determine both the selection into contract farming and the resulting outcome (income). Hence, if the error term $(\varepsilon_{1i} - \varepsilon_{0i})$ in equation 3 is correlated with time-invariant household characteristics, household fixed effects will solve the endogeneity problem. Panel data need to be collected before the implementation of the treatment. If panel data are only collected once the treatment has been implemented, some households may have opted in and out of contract farming because of unobservable, idiosyncratic shocks. Unfortunately, no study was identified that collected data before and after a firm offered contractual arrangements to small-scale farmers. A study by Dedehouanou, Swinnen and Maertens (2013), which did not quantitatively estimate income effects, collected data at two points in time, but some

farmers were already participating in contract farming during the first wave of data collection.

Observational studies that rely on cross-sectional data to identify the causal impact of contract farming are more likely to be biased than RCTs or studies using panel data. Cross-sectional surveys take a ‘snapshot’ of income effects at one point in time. The variables (characteristics) that define participation in the contract may have changed due to this participation, and this may preclude credible net-effect estimates. Therefore, some studies attempted to create a panel dataset retrospectively by asking farmers to report on living conditions and assets before they participated in contract farming (e.g. Maertens and Swinnen [2009]). These lagged variables were then used to match households using propensity score matching or as instrumental variables.

A simple OLS regression with a set of explanatory variables for a certain outcome (e.g. income), comparing treatment and control group with a dummy variable, would yield unreliable results due to collinearity. It would be impossible to determine if the effect was the result of the intervention or a result of the other factors that define the propensity of the farmer to participate in the contract. This is solved econometrically by identifying one or more good instruments that can isolate the effect of the contract. With cross-sectional data, there are several approaches to addressing the selection problem: Propensity Score Matching (PSM), Instrumental Variable (IV)/Heckman models and endogenous switching regression models. PSM studies assume that the researcher has observed all relevant criteria for selection into contract farming, and they compute the likelihood of a farmer becoming involved, based on his/her observables, in the contractual arrangement. Using these propensity scores, comparable contracted and non-contracted farmers are selected (matched), and the difference in the outcome variable between both groups is considered as the causal impact of contract farming on income. However, this estimate may well remain biased if selection into contract farming is partially based on unobservable characteristics.

IV and Heckman studies use a unique variable or a vector of variables that has such a high correlation with the propensity to participate that it can be used to isolate the effects of the intervention. This instrumental variable should be strongly correlated with selection into contract farming, but uncorrelated with the outcome variable. In theory, an IV approach is preferable over PSM whenever it credibly controls for selection on unobservables. In empirical work, however, this distinction is less clear-cut, as many studies lack sufficiently strong instruments. Heckman studies use a vector of variables to do so. Based on a first-stage regression that estimates participation in the contract, a second-stage regression uses the estimated parameters of each farmer in the regression to estimate the net-income effects of the contractual arrangement. Switching Regressions take account of the different ‘production functions’ for the contracted and the traditional crops. For each ‘regime’, a separate regression is performed and the net-results are the subtraction of the (modelled) results in both regimes for a farmer with the same characteristics: the farmer modelled as producing the contracted crop versus producing the traditional crop.

All study designs may be vulnerable to the fact that they assess net-effects at one moment in time. Most longitudinal panel studies (e.g. RCT, Diff-in-diff) use two measurements to derive inference of impact. The years used for baseline or endline

may, however, refer to ‘unusual’ situations (e.g. weather, markets). These sources of variability are especially relevant when the contract farming arrangement covers new crops or livestock that are different than the traditional (counterfactual) farming in the comparison group. If conditions vary between years and affect these crops differently (e.g. terms of trade, drought resistance, etc.) a comparison between baseline and endline results in unreliable net-effect estimates.

In a context of high variability, baseline surveys are not always a cost-effective way to obtain impact data. McKenzie (2012) argues that more observations over time may improve the validity of any of the (quasi) experimental designs. He argues that for outcomes with a high variability over time (a low autocorrelation), multiple post-treatment measurements can dramatically increase the statistical power of the design, and, within budget constraints, even be preferable to a design with a baseline survey (McKenzie 2012:219-20). The time series that results from multiple follow-up surveys would better capture time-varying outcomes and register stepping-out (attrition), which are important validity threats to inferences based on cross-sectional surveys only.

This overview of the different strategies to identify the causal impact of contract farming on welfare shows that some study designs were more statistically robust than others. Studies reporting only comparative statistics without attempting to control for unobservable variables were excluded from the meta-analysis, and are listed in Appendix 9.2. All studies used were observational studies that used econometric methods to control for selection bias. Furthermore, some studies were excluded because they applied these methods incorrectly– for example, by using unconvincing instrumental variables. For all studies included, the risk of bias was assessed using 3ie’s Risk of Bias tool explained in section 3.5.

The 28 contract farming arrangements, covered by the 23 studies included in the meta-analysis, were subsequently used on the explorative analysis for enablers and barriers of effectiveness, using Qualitative Comparative Analysis.

3.2 Search methods for identification of studies

To find the studies that could be used in the meta-analysis, we used a comprehensive search strategy. The main terms used to identify the pool of studies within which we expected to find studies that covered the effectiveness of contract farming arrangements were: contract farming, nucleus estate, cooperative, producer organisation, pre-harvest agreement, value chain, farm-firm, outgrower, vertical integration. See the detailed search terms in Appendix 9.3 and the systematic research protocol (Ton et al., 2015).

We screened the retrieved studies to identify a core set of studies that give evidence on quantitative effectiveness (research question 1), to define the empirical instances of contract farming for meta-analysis. A wider pool of studies, related to these empirical instances was used for responding to research question 2, to identify relevant complementary information on the contextual conditions of these empirical instances.

3.2.1 Electronic searches

We used the results of a preliminary Scopus search in October 2014 and February 2015 to develop and fine-tune the search strategy with appropriate search terms. We applied the search terms between 30 September and 21 October 2015, as documented in Annex

2 of the protocol (Ton, Vellema, D'Haese et al. 2015). We searched the following electronic libraries:

- Scopus.com
- CAB Abstracts
- Web of Science
- Agricola
- Econlit
- Tropag & Rural

3.2.2 Searching other resources

In addition to the electronic search, hand-searching and snowballing added more studies to the review. First, we searched for complementary academic and non-academic literature in several databases in which electronic search results are impossible or cannot be exported in a useable format. Databases of organisations such as the FAO, World Bank, and IFAD were searched for additional grey literature. See Ton et al. (2015) for further details on this search process.

Second, we snowballed the references mentioned in review articles and books on the subject.

Third, citations to the studies used in the meta-analysis were identified through citation searches (Google Scholar, Scopus, Web of Science) and, if not yet included, screened on the inclusion/ exclusion criteria. Subsequently, the references of all newly included studies were screened for other relevant studies, using snowballing of references and forward citation tracking.

Fourth, the link and information on a website dedicated to this systematic review (http://contractfarming-systematicreview.wikispaces.com/) was used to present the results of the initial search process and used to contact key resource persons to suggest missing, unpublished or unfinished studies.

Fifth, to answer research question 2, we retrieved additional material referring to the same contractual arrangements and contexts which were the focus of the studies selected for meta-analysis. This allowed us to complement the information on context characteristics and contract modalities needed to reflect on the enablers and barriers in the empirical instances of contract farming covered by the studies selected for the meta-analysis. Furthermore, we did separate searches for each empirical instance: the name of the contract farming arrangement, the name of contracting firm, and the geographical location of the empirical instance. To do so, we used Google, www.google.com and Google Scholar (scholar.google.com). The retrieved documents were further checked for other useful references.

3.3 Data collection and analysis

3.3.1 Selection of studies

The search results used to identify a broad set of potentially relevant qualitative and quantitative empirical studies as well as conceptual papers on contract farming were stored in EPPI Reviewer 4. Each upload indicates the source and search terms used in search-specific RIS-files.

In the title-abstract screening, reviewers were deliberately over-inclusive and incorporated all studies that were related to agriculture and in which transactions in product or input markets were mentioned. After the first preliminary search in Scopus.com (February 2015), all retrieved references (around 2,500) were double-coded to align criteria between reviewers during the title-abstract screening. After the final search (October 2015), half of the new references were double coded and the other half single coded.

During full-text screening, all studies were assessed by two reviewers independently assessed for inclusion in the meta-analysis. The lead researcher reconciled the differences, consulting the full-text of the study and discussing his assessment with both reviewers. The data extraction of all studies included in the meta-analysis was done by two of the lead researchers. A third reviewer arbitrated any disagreements.

3.3.2 Title-abstract screening

In the title and abstract screening, studies were excluded if they did not relate to or comply with one or more of the following six exclusion criteria, mentioned in paragraph 3.1: no agricultural value chain; forestry (timber, wood); marine fishery (sea, coastal); not in low/middle-income country; no market-related contractual arrangement; not considering farmers; and relevant but no empirical instance.

3.3.3 Full text screening

After title-abstract screening, we retrieved the full text for the set of selected publications. The references were screened on the above-mentioned exclusion criteria about relevance. Some studies did not assess effectiveness but, for example, differences in characteristics of farmers with and without contracts, or explored differences in performance between contracting farmers. We used two additional exclusion criteria to focus further on the subset of studies potentially useful for answering the research questions: study not examining impact effects; and no pre-harvest service delivery. All studies that resulted from this full-text screening contained information on one or more empirical instances of contract farming.

3.3.4 Selection for meta-analysis

The studies that remained were differentiated based on the methodological rigour of the assessment of the effect size. To be selected for meta-analysis, data had to be collected at farm or household level in both intervention and comparison groups. Observational studies were included whenever they controlled for unobservable characteristics using statistical matching (propensity score or covariate matching) or regression adjustment (difference-in-differences, single difference regression analysis, instrumental variables estimation, and Heckman selection models). Both study designs that collect longitudinal data at baseline and end-line, and those that use cross-sectional end-line data only, were included.

The studies that did not have a counterfactual design were excluded from the meta-analysis. However, some of these studies were retrieved during synthesis if they related to one or more of the selected empirical instances of contract farming, and were used to obtain additional insights in processes or mechanisms related to the contract farming arrangement in that empirical instance, to answer our second research question.

3.3.5 Data extraction and management

All studies were entered in the specialist systematic review software – EPPI Reviewer 4 (Thomas, Brunton and Graziosi 2010). EPPI-Reviewer 4 has been developed and maintained by the EPPI-Centre at the Social Science Research Unit at the Institute of Education, University College London, UK. EPPI-Reviewer was used to screen the studies, to archive the studies used in the qualitative and qualitative analysis, and for the assessment of study quality and risk of bias.

The meta-analysis was done by using packages available in Stata. *Metan* is the main Stata meta-analysis command. Its latest version allows the user to input the cell frequencies from the 2 × 2 table for each study (for binary outcomes), the mean and standard deviation in each group (for numerical outcomes), or the effect estimate and standard error from each study. It provides a comprehensive range of methods for meta-analysis, including inverse-variance–weighted meta-analysis, and creates new variables containing the treatment effect estimate and its standard error for each study. These variables can then be used as inputs into a wide range of other Stata meta-analysis commands (Harris, Bradburn, Deeks et al. 2010).

The detailed analysis of evidence on enablers and barriers in each of the empirical instances used EPPI Reviewer 4 (Thomas et al. 2010) and Atlas.ti 7 (Friese 2014) to prepare for qualitative data analysis. Kirq 2.1.12 (Reichert and Rubinson 2014) was used to explore configurational patterns in the data.

3.3.6 Measures of treatment effect

Response ratios (RRs) were calculated to measure effect sizes. Response ratios have the advantage of being straightforward to calculate. Only two of the studies reported sufficient information to calculate SMDs directly (other methods would be needed to back-translate them from t-statistics and so on). To complete the information needed for calculating the RRs and confidence intervals, we contacted 15 authors, only seven of whom provided the missing information. One author indicated that they had no access to the data and nine did not react to the request and follow-up emails. These studies were therefore excluded from the meta-analysis (see Appendix 9.2).

Response ratios are expressed as the difference in mean outcome in the intervention group as a proportion of the outcome mean in the comparison group. RRs have the advantage of being easy to interpret. Values above and below one indicate proportionate changes in the outcome of the treatment group over the comparison. Thus, an RR of 0.90 indicates a 10 percent average decrease in the treatment group relative to the comparison, and an RR of 1.10 indicates a 10 percent average increase. All outcomes were measured on a continuous ratio scale with a natural zero point, which is a required condition for RR to be meaningful (Borenstein, Hedges, Higgins et al. 2011).

For matched-based studies, response ratios were calculated using the following formulae:

$$RR = \frac{Y_t}{Y_c} \qquad SE(RR) = Exp\left(\frac{Ln(RR)}{t}\right)$$

where Y_t is the outcome on the treated, Y_c the outcome for the comparison group, and t the value of the t-statistic. For regression-based studies, response ratios were calculated using the following formulae:

$$RR = \frac{Y_c + \beta}{Y_c} \qquad SE(RR) = Exp\left(\frac{Ln(RR)}{t}\right)$$

where β is the unstandardized regression coefficient. Effect sizes are only strictly comparable across studies using a common regression model (Keef and Roberts 2004:103). In the case of multivariate studies, this means that studies should analyse the same treatment, use the same way of measuring the outcome variable, the same method, and the same covariates. This was clearly not the case. Hence, the sensitivity of the results to these factors was checked using analyses described in section 4.3.2.

The data required to calculate the RRs depended on the estimation strategy used in the study. Often the studies did not report exact t-statistic, standard errors or p-values, but indicated significance at the 10, 5 and 1 percent levels. In those cases, we assumed that the p-value was exactly equal to the reported significance level.

The calculation of the RRs was most straightforward for studies using propensity score matching or endogenous switching regression models. These studies typically reported income of treatment (Y_t) and control group (Y_c) after matching along with the standard error or t-statistic of the difference between treatment and control group. RRs were then calculated as Y_c/Y_t and the standard error of the logarithm of RR as: $\ln(se_{RR}) = \ln(RR)/t$. Most of these studies matched treatment and control groups using nearest neighbour and/or kernel matching. Whenever the results of both techniques were reported, we used nearest neighbour matching. Some studies only reported the difference between control and treatment group after matching. In those case, the authors of the study were contacted and asked to report Y_t and Y_c. If the authors did not reply, we assumed that Y_t after matching equalled Y_t before matching.

The calculation of RR for studies using instrumental variables followed the guidelines of Waddington, Snilstveit, Hombrados et al. (2014). The response ratio was calculated as: $RR = (Y_c + \beta)/Y_c$, where Y_c is the average income of the control group and β identifies the impact of the contract. Standard errors of the logarithm of the RR were then calculated as $\ln(RR)\ /t$. Most studies reported β with and without correcting for self-selection. We always used β reported in the regression that addressed the endogeneity issue, even if a Hausman specification test rejected the endogeneity of participation in contract farming in the regression to estimate the income effects.

An additional complication occurred if the regression framework used the logarithm of the outcome variable as the dependent variable (five studies). In these studies, β represents the percentage increase in the outcome variable if the probability of participating in contract farming increases by 100%. To calculate RR, we followed the strategy proposed by Bellemare (2012) and assumed that the RR equal $\beta * n_c/(n_c+n_t)$, where n_c and n_t are the number of observations in control and treatment group, respectively. This is an upper limit of the impact of contract farming on the outcome variable, as a linear interpolation of the effect size is used, whereas the logarithmic function is concave.

3.3.7 Unit of analysis issues

For our meta-analyses of income effects, the unit of analysis was the empirical instance of contract farming. We did not find any paper that reported effects on higher cluster levels than the farm household. For each meta-analysis, only one effect size was used per empirical instance.

No studies reported more than one ultimate outcome. More problematic were studies containing several empirical instances of contract farming, covering distinct crops. There were four such studies: Bellemare (2012), Miyata et al. (2009), Narayanan (2014), and Simmons et al. (2005). Of these, the first two contained several empirical instances of contract farming but did not distinguish between them in the analysis. Hence, they had to be considered as consisting of a single case in the meta-analysis. The study of Narayanan (2014) did report separately on each of the four empirical instances. However, it relied on the same control group. Considering these instances as completely independent would put too high a weight on what is, in effect, a single study. We present the results of the meta-analysis using a synthetic effect. However, in the moderator analysis, the synthetic effect size would cancel out the existing heterogeneity between the instances. Hence, we reported both the separate effect sizes as well as the synthetic effect size, and showed how the difference affected overall results.

Simmons et al. (2005) reported on three empirical instances, two of which were used to calculate effect sizes. The effect size of the third instance was calculated based on data taken from another more detailed paper of the same authors on this particular contract farming scheme (Winters et al. 2005) in which they applied a two-staged regression using the same data as in Simmons et al. (2005). The analyses of these three instances relied on independent data and different model specifications and hence were considered to be independent.

When authors presented results of several models, the results of the methodology with the best control for selection bias were selected. For this decision, we follow the hierarchy of models explicated in section 3.1.5. When multiple regressions based on the same method were presented, we selected results from the most appropriate specification (based on use of appropriate covariates and whether covariates in less parsimonious models were statistically significant). Similarly, when the results of multiple PSM models were presented, we preferred nearest neighbour matching. The reason for preferring nearest neighbour matching over other matching algorithms was consistency: all papers that used PSM included the results based on nearest neighbour matching, which was not the case for other matching algorithms.

3.3.8 Dealing with missing data

Authors of studies with missing data were contacted, in some cases up to four times, with a request to provide for the specific data required to complete the effect size calculation. Five authors responded to the request. Of the six remaining studies, two had to be excluded. For the other four studies, effect size calculations were completed using assumptions or approximations. Additional assumptions were sometimes required because of missing data: t-statistic, standard errors or p-values.

3.3.9 Risk of bias assessment

For the risk of bias assessment of each study, we used a 3ie Risk of Bias tool (Hombrados and Waddington, 2012), which was developed to enable consistent assessment of internal validity of social experiments and quasi-experiments. Studies were assessed by each researcher individually, and resulting disagreements were discussed and resolved jointly. The 3ie Risk of Bias tool consists of eight evaluation criteria, each focusing on a different type of threat to validity (see section 4.2). The tool contains detailed descriptions of how to attribute a score to each criterion specified per type of study design.

3.3.10 Assessment of heterogeneity

The presence of heterogeneity was visually examined by making forest plots of the pooled results and tested using the chi-squared test. Given the small number of included studies, the level of significance at which we considered heterogeneity to be present was set at 0.1, to account for the low power of the test. Furthermore, heterogeneity was assessed with the I-squared statistic, which measures the percentage of variability that is due to real variability between studies. Values close to 100% indicate a large real variability and values close to zero indicate no observable real variability. Final evaluation of test results depended on the significance of the chi-square test described above and the sign and size of standardised effect sizes. As we used a random effects model for the meta-analysis, between-study heterogeneity could also be assessed with the tau-squared statistic. Potential sources of heterogeneity were formally investigated using subgroup analysis.

3.3.11 Assessment of reporting biases

Publication bias was assessed visually with funnel plots produced using the *metafunnel* and *metabias* command in Stata, and tested more formally with the Egger's meta-regression test (Egger et al., 1997).

3.3.12 Data synthesis

Quantitative information on the effectiveness of contract farming was synthesized using inverse-variance weighted statistical meta-analysis. Random effects rather than fixed effects meta-analysis was used, because the treatment was not uniform and its effects were likely to be context-specific. Effect sizes were also likely to have been affected by differences in study designs and control variables used, in the case of regression-based studies. Random effects meta-analysis allows for such differences, by assuming that there is some true variation between the included studies besides random variation due to sampling.

The results of the meta-analysis were checked for their sensitivity to outliers, synthetic effects, research design, and outcome measure (definition of income). These analyses were only done for the ultimate outcome variable income, as this was the only variable for which sufficient observations were available. It was not possible to compare studies based on the length of follow-up, because all included studies relied on cross-sectional data.

We complemented the impact studies with qualitative studies and other pieces of information that could shed light on the contextual conditions and content of the contractual arrangement in each empirical instance. Most studies included t-tests or

Probit regressions to identify differences between farmers with a contract and those without, but sometimes this information was not in the main study and had to be derived from other studies that reported on the same empirical instance. We used the results of these analyses to explore for plausible predictors of participation and exclusion in contract farming arrangements, and especially to answer the question if these arrangements tended to benefit larger or asset-richer, or smaller and asset-poorer households.

3.3.13 Investigation of heterogeneity

According to the practitioners in the Advisory Board of this review, the most important sources of heterogeneity between empirical instances were expected to be due to the type of production, region, and the service package embedded in the transaction. The effect of these sources on the impact of contract farming on income was explicitly tested for in the meta-analysis. Ex-ante, before analysis, the types of production were categorised as annual or perennial crops or animal husbandry, and depending on their perishability. Regions were broadly distinguished as Latin America, Africa, and Asia. Ex-post, during analysis, the contracts were characterised according to the market incentives, price-premium, the role of farmer organisations in the set-up, and the level of service provisioning, including credit, extension, and inputs.
In addition to the analysis of these conditions as single predictors in the meta-regression, we also used a Qualitative Comparative Analysis (QCA) (Ragin 2008, Reichert and Rubinson 2014) to explore configurations of conditions that could explain differences in effectiveness, using as conditons the incentives to farmers embedded in the contract: involvement of a farmer organization at the start, offering higher-than-local prices (price-premium), access to credit, and provision of key inputs (seeds/breeds, agro-chemicals).

QCA creates an overview (called a ‘truth table’) of all possible combinations of conditions and lists the cases that share the same combination. This matrix is called a ‘truth table’. The truth table shows which cases (empirical instances of contract farming) and associated combinations of conditions were consistent with a relatively high or relatively low effectiveness on income. Some cases will be ambiguous and cannot be classified as 0 (not having a condition) or 1 (having a condition). Therefore, fuzzy set scores are used, with 0.8 indicating most likely present, and 0.2 as most likely absent. For each combination of conditions, QCA presents the set-consistency score, which is the lowest value of the fuzzy-score of one of the cases that is member of the group of cases that shares the combination of conditions, be it in the set of conditions or the in the outcome.

To calibrate the outcome variable as high or low effectiveness, the pooled average effect size is used as the cross-over point. Cases with a response ratio below 25% are considered low effective, and cases above 50% are considered high effective. The intermediate values are converted to fuzzy set scores using a logistic regression function, as recommended by Ragin (2008). The pooled average effect-size is used as the cross-over point (fuzzy set score= 0.5)

4. Results

4.1 Description of Studies

4.1.1 Results of the search

A preliminary electronic search was performed in February 2015 on Scopus and retrieved 3,355 studies. These papers were title-abstract screened by a team of master students in Ghent and Wageningen University to select the agriculture related studies. Each paper was double-screened, using EPPI Reviewer 4, with a senior researcher reconciling differences. At this stage, papers not concerning the agricultural value chain, developing countries, market-related contracts, farmers, or empirical instances were excluded. When titles and abstracts provided insufficient information to justify exclusion, papers proceeded to the next screening stage. Twelve papers were excluded because a full-text version could not be retrieved.

We performed the final search for this review in all electronic libraries in October 2015, which retrieved a total of 8,529 studies (Figure 4). The new studies, compared with the Scopus results in February 2015, were single-screened by one of the reviewers that participated in the initial double-screening.

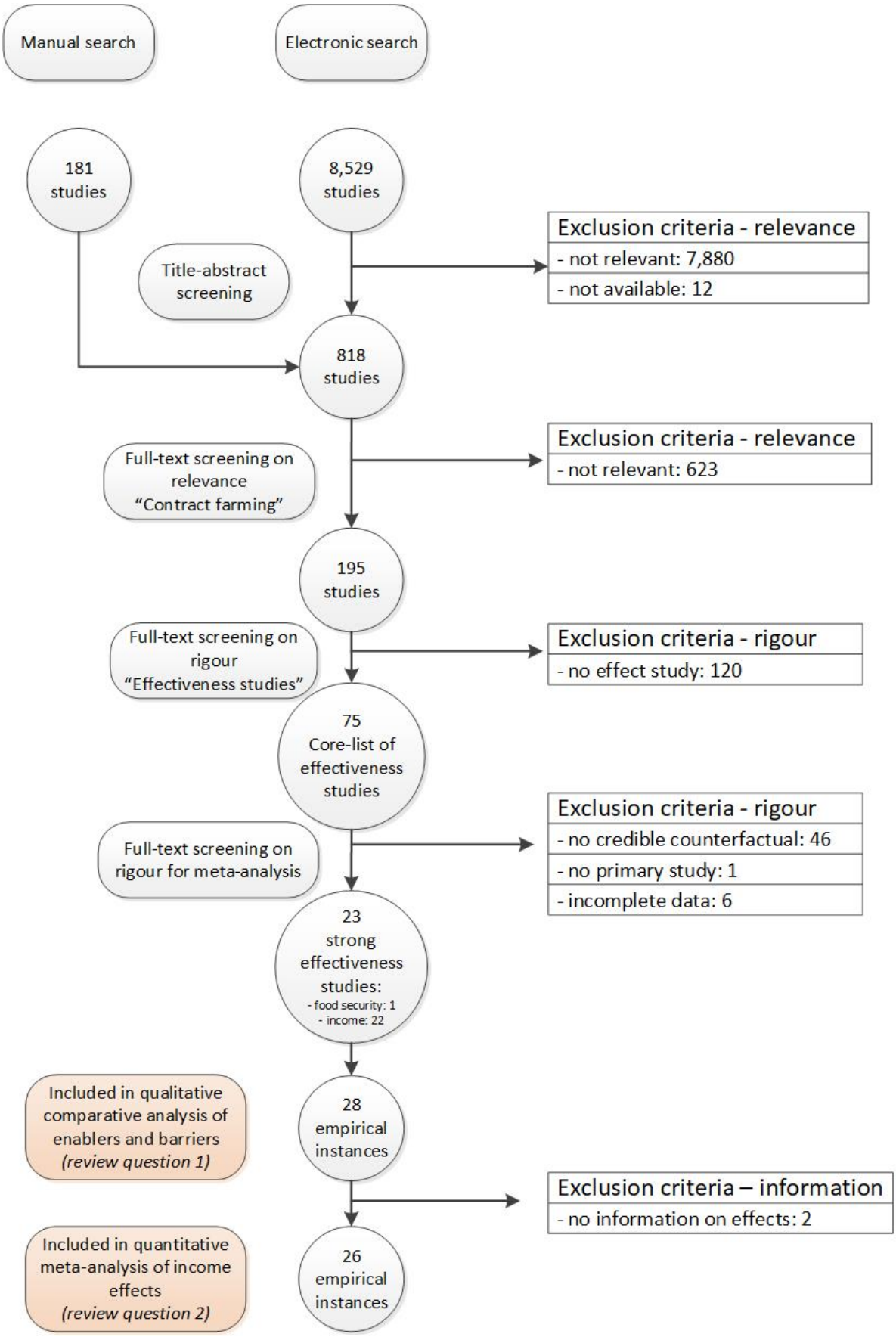
After title-abstract screening, we were left with 637 studies. In addition to the electronic search, a manual search was conducted on online databases and by snowballing references. This manual search retrieved another 181 studies.

The full-text screening was performed on all remaining studies. All these papers were double-screened, reconciling differences. All studies with no market-related contract, counterfactual design, report of smallholder effects, pre-harvest services, or farmers were excluded. After the full-text screening, 195 studies remained.

In the next step, all studies which were not effectiveness studies of contract farming were excluded. The definition given in section in section 3.1.2 was applied. The remaining set of papers was referred to as the core set and consisted of 75 studies.

Of the studies contained in the core set, 29 met the criteria for statistical rigour required for inclusion in the meta-analysis. Of these, six studies had to be excluded because effect sizes could not be calculated – after several attempts to contact the authors to obtain missing information. These studies met our definition of contract farming and the criteria on statistical rigour.

Figure 4: Sources of included studies



Only studies which reported on at least one of the ultimate outcomes of food security and income were included. Several papers were retrieved as working papers and subsequently published in 2015 and 2016. We used the data in the printed version of the study in the meta-analysis and synthesis. As only one study reported on food security (Bellemare and Novak 2016) it could not be included in a meta-analysis of food security effects for lack of comparable studies. This study relied on the same method and data as an earlier study by the one of the authors which was included (Bellemare, 2012). In the end, 22 studies could be included in the meta-analysis of income effects. These papers contained 28 empirical instances and provided sufficient data on 26 of these.

4.1.2 Included studies

Included studies dealt with contract farming arrangements in 13 countries. The majority of studies concerned African and Asian countries; only one study came from South America (Peru: Escobal and Cavero, 2012). Moreover, most empirical instances covered in the studies were highly regional in nature, and not necessarily representative for a country as a whole. As described in section 1.3, this geographical focus is understandable, because companies select regions with agro-ecological conditions suitable for a certain crop, having a sufficient number of farmers, and where some essential infrastructure is present (Barrett, 2010).

Although 22 included studies used income as the dependent variable in the regression, the definition and proxy-indicators varied between studies. Most studies reported crop income (41%), farm income (23%) or household income (27%). Two studies (9%) reported household expenditure. Only eight studies also reported some intermediate outcomes. Most of these studies reported more than one intermediate outcome, often part of a causal relation that leads to income effects (see Table 1). Yield (4), price (2), and labour use (2) were the most frequently reported intermediate outcomes. All other outcomes were only reported once.

Only one of the retrieved studies assessed the impact of contract farming on food security with an econometric design that controlled for selection bias. Bellemare and Novak (2016) asked the household about the ‘duration of the hungry season’ (in months), and reported the difference between contract farmers and non-contract farmers. Moreover, they used this variable to estimate a second effect-size, the likelihood that a household would exit the hungry season at any point in time.

Sample size in some studies was relatively low. The average sample size was 260 and ranged from 26 (Warning and Key 2002) to 1178 (Bellemare and Novak 2016, Bellemare 2012). In total, de studies covered 7,471 respondents in 28 contract farming arrangements.

In the following we describe the samples used in the included studies in more detail (see also Appendix 9.1 for a brief summary of the contractual arrangement). The study from Awotide et al. (2015) comprises a sample of 341 Nigerian rice farmers of which 150 participated in different contract farming schemes of several private firms providing seeds, agrochemicals and extension services. The study population was largely male (59%) with an average age of 43 years with five years of formal education, and owning an average of 2.9 ha of land.

The sample used for both Bellemare (2012) and Bellemare and Novak (2016) was larger and included 1178 farmers of green beans, leek, snow peas, barley and rice that were contracted by different private firms in Madagascar. The contracts offered differed between the firms with all offering agrochemicals and extension, with some adding seeds to the package. Just under 50% of the sample participated in one of the contracts. Similar to the study by Awotide et al. (2015), the household heads in the sample had an average age of 43 years with six years of formal education. Most of them were male (92%). The average size of land owned by the households was somewhat smaller with 1.7 ha.

The sample of Bolwig et al.'s study (2009) included 160 coffee farmers of which 112 participated in a contract. Participating household heads were 46 years of age on average and had seven years of education. The average land holding was 1.1 ha. Non-participant households owned 0.8 ha, and household heads were 47 years of age also with seven years of education. The contract farming arrangement was offered by one firm and included seeds and agrochemicals on a limited scale and extension services.

Cahyadi and Waibel (2011) looked at 245 palm oil producers in Indonesia contracted by a single private firm. The contract included seeds and extension. Of the sample, 126 farmers participated in the contracting scheme. The average age of the household head was 46 years, and households owned 2.9 ha of land on average.

The only study from South America included 360 potato farmers in Peru of which one fifth was contracted by a private firm in cooperation with an NGO (Escobal & Cavero, 2012). The participating farmers were provided with extension. Other services provided by a supporting NGO remained unclear. Descriptive statistics were reported separately for participants and non-participants. The age of the household head was 47 and 48 years and average years of education were 10 and 12, respectively. Participants owned 5.6 ha of land and non-participants 2.6 ha. Household heads were mostly male (84% for participants and 96% for non-participants).

Girma and Gardebroek's (2015) study looked at 195 honey producers in Ethiopia. The contract offered by a single private firm, in which 79 of the beekeepers participated, included credit, extension, inspections, honey containers and other necessary inputs. Again, summary statistics were reported separately. Participants had an average of 5 years of schooling and 23 years of experience in beekeeping. Similar to non-participants, who had four years of education and 20 years of experience in honey production.

Ito et al. (2012) considered 318 Chinese Watermelon farmers. In this case, the contract included seeds, agrochemicals, extension, subsidies on land improvement and access to the wholesale market and was offered by a farmer cooperative and not a private firm. Roughly half of the sample participated in the contract. The household heads were 53 and 55 years of age, participants and non-participants respectively, both with three years of education on average. The average land size between them was similar: 0.5 ha and 0.4 ha.

The study performed by Jones and Gibbon (2011) covered 222 cocoa farmers in Uganda of which 135 participated in a contract offered by a private firm. Initially, the firm offered seeds and agrochemicals. Later, only extension and support in the initiation of saving

societies were included in the contract. Farms had an average size of 1.7 ha, and household heads were 44 years of age.

Maertens and Swinnen (2009) considered contracts offered by several different exporting companies of French beans in Senegal. The services included in the contracts were credit (for some), seeds, agrochemicals, extension, and land preparation, and coordination and financing of planting and harvesting (on demand). The sample comprised 217 farmers of which 59 produced French beans on contract. Non-participants were from the same area but did not produce the beans. The average household head was 53 years old, 18% had primary education, and almost none of them were female. Average land holding was 5 ha.

The study performed by Miyata et al. (2009) dealt with both apples and green onions in China. The contract for apples included agrochemicals (for most), extension, and sometimes spraying services, while the green onion contract offered seeds in addition to these services. The total sample was 162 farmers with 98 participating in one of the two contracts. On average household heads were 45 years old, went to school for eight years and cultivated 0.8 ha of land.

Narayanan (2014) looks at four different empirical instances with sample sizes between 262 and 289 farmers per instance. For Marigold, farmers were 46 and 45 years and owned 2.4 ha and 2.1 ha on average, for the 208 non-participants and 59 participants, respectively. The contract included seeds and extension. The second instance is papaya for which there were 72 contract and 208 non-participants. The average age was 45 and 46 years and land size was 2.3 ha and 2.5 ha, respectively. The contract was similar to those for marigold farmers but added the organisation and training of hired labour. Broiler farmers were the third instance described in the study. However, all 81 broiler farmers produced on contract. They received the breeds, agrochemicals and extension; other services were not specified. The 208 farmers in the comparison group did not produce broilers. The former group had an average land size of 2.8 ha, the latter 2.5 ha. In both cases the average age was 46 years. The last empirical instance concerns gherkins. In this case, all 262 farmers produced on contract but only 77 produced gherkins. They had an average age of 38 years and owned 0.9 ha of land. The control group had an average age of 47 years and owned 2.7 ha of land. The gherkins contract offered seeds and extension; other services remained unspecified.

Ramaswami's study (2009) also looked at broiler production in India using a sample of 50 farmers of which 25 produced on contract and 25 were independent farmers. The contract included inputs and extension, credit (in kind), and price insurance. Contract farmers were on average 39 years old, had 11 years of schooling and owned 2.5 ha of land. Independent broiler farmers were 36 years old on average, had 12 years of education and owned 3.1 ha of land.

Rao and Qaim's (2011) sample of Kenyan vegetable farmers included 402 farmers of which 133 participated in a contract. However, the contracts were offered by multiple firms, and the provisions of the contracts were unclear. The average age of the farm operator was 47 and 49 years, years of education were 10 and 9, and operators were 93% and 88% male, for contract and non-contract farmers respectively. Land ownership was 1.1 ha for contract farmers and 0.8 ha for non-contract farmers.

Saigenji (2012) compared Vietnamese tea growers that were either contracted by private firms and farmer cooperatives, or they supplied to state-owned enterprises. The contracts included credit, seeds, agrochemicals, extension, and input application. The sample compared 40 farmers contracting with state-owned enterprises and 34 farmers contracting with private firms to 50 non-contract farmers. The average age was 34 years and 32 years for contract and non-contract farmers, respectively. Contract farmers cultivated 1.05 ha, while non-contract farmers cultivated 1.12 ha.

A sample of 585 rice farmers in Laos was studied by Setboonsang et al. (2008) of which 332 farmers produced on contract. The contract was offered by a private firm through a farmers' cooperative and included seeds, agrochemicals and extension services. Contract farmers planted 1.11 ha of land, while conventional farmers used 1.43 ha.

Simmonds et al. (2005) covered three empirical instances in Indonesia of which one is also covered by Winters et al. (2005) and will be described separately. The first empirical instance covers 300 seed rice farmers of which 150 were under contract. The related contract includes seeds and extension, with remaining services being unclear, and is offered by a state-owned enterprise.

Table 1: Overview of the studies included in the research synthesis

Author	Country	Product	Sample size	Study design	Outcomes[1]	Effect size (RR)	95% CI RR)[2, 3]
Awotide et al., 2015	Nigeria	Rice	341	PSM	Yield ➔ Crop income ➔ Poverty	1.55 1.71 0.83	[1.33, 1.77] [1.42, 2.00] [0.66, 1.00]
Bellemare, 2012	Madagascar	Green beans, leek, snow peas, rice, barley	1178	IV	Household income	1.55	[1.28, 1.81]
Bellemare and Novak, 2016	Madagascar	Green beans, leek, snow peas, rice, barley	1178	IV	Duration hungry season	0.92	[0.84, 1.00]
Bolwig et al., 2009	Uganda	Coffee	160	Heckman	Yield ➔ Practice adoption ➔ Crop income	1.30 1.53 1.92	[1.17, 1.44] [1.06, 2.23] [1.55, 2.29]
Cahyadi and Waibel, 2011	Indonesia	Oil palm	245	Heckman	Household income	1.24	[0.95, 1.53]
Escobal and Cavero, 2012	Peru	Potato	360	Switching regression	Crop income	1.76	[1.17, 2.35]
Girma and Gardebroek, 2015	Ethiopia	Honey	195	IV	Price ➔ Crop income	1.23 2.19	[1.20, 1.25] [1.66, 2.72]
Ito et al.,2012	China	Watermelon	318	PSM	Farming income	1.56	[1.22, 1.90]
Jones and Gibbon, 2011	Uganda	Cocoa	222	IV	Total output ➔ Price ➔ Crop income	1.25 1.08 1.52	[1.13, 1.36] [1.04, 1.12] [1.32, 1.72]
Maertens and Swinnen, 2009	Senegal	French beans	217	PSM	Household income	3.23	[1.41, 5.05]
Miyata et al.,2009	China	Apples, green onions	162	Heckman	Household income	1.27	[1.02, 1.51]
Narayanan, 2014	India	Marigold Papaya	262-289	Switching regression	Crop income	0.52 1.43	[0.18, 1.54][3]

		Broiler chicken Gherkins				17.64 1.27	[0.59, 3.45][3] [9.38, 33.15][3] [0.15, 10.92][3]
Ramaswami, 2009	India	Broiler chicken	50	IV	Crop income	1.85	[0.85, 2.84]
Rao and Qaim, 2011	Kenya	Various vegetables	402	Switching regression	Household income	1.48	[1.11, 1.85]
Saigenji, 2012	Vietnam	Tea	88, 90	PSM	Technical efficiency ➔ Household expenditure	Not possible 1.04	[1.01, 1.07]
Setboonsarng et al., 2008	Laos	Rice	585	PSM	Yield ➔ Price ➔ Crop income	1.26 1.18 1.80	[1.08, 1.44] [0.66, 1.70] [1.30, 2.18]
Simmons et al., 2005	Indonesia	Seed rice	124	IV	Farming income ➔ Household labour	0.94 1.10	[0.70, 1.18] [0.98, 1.22]
Simmons et al., 2005	Indonesia	Broiler chicken	200	IV	Farming income ➔ Household labour	4.91 0.67	[2.67, 7.15] [0.27, 1.70][3]
Sokchea and Culas, 2015	Cambodia	Rice	75	Heckman	Farming income	1.85	[1.03, 2.67]
Trifković, 2014	Vietnam	Catfish	191	Heckman	Household expenditure	1.29	[1.14, 1.45]
Wainaina et al., 2014	Kenya	Broiler chicken	180	PSM	Crop income	1.31	[1.03, 1.58]
Wang et al., 2014	Vietnam	Various vegetables	107	PSM	Household income	1.37	[1.06, 1.67]
Warning and Key, 2002	Senegal	Peanuts	26	Heckman	Farming income	1.29	[1.00, 1.58]
Winters et al., 2005	Indonesia	Seed corn	300	Heckman	Input use ➔ Labour use ➔ Farming income	2.13 1.15 2.83	[1.25, 3.01] [0.97, 1.33] [1.66, 4.01]

[1] Arrows reflect the causal logic as described in the study by the authors;

[2] The CI reported here differ from the CI reported in the meta-analysis (forest plots) because the CI is this table are symmetric around the RR, while the CI used in the meta-analysis are symmetric around ln(RR). The CI reported here equal $[\mathrm{RR} \pm 1.96\ \mathrm{se}(\mathrm{RR})]$, while the CI reported in the meta-analysis equal $[\exp(\ln(\mathrm{RR}) \pm 1.96\ \mathrm{se}(\mathrm{RR}))]$;

[3] CI interval calculated using $[\exp(\ln(\mathrm{RR}) \pm 1.96\ \mathrm{se}(\mathrm{RR}))]$ as CI would otherwise include negative RR values.

The average age in both groups is 48 years, and the average duration of schooling was six years. Contracted farmers owned 0.5 ha of land while non-contracted farmers owned 0.7 ha. The second instance covers 200 broiler producers. The 80 contracted farmers received credit, breeds, agrochemicals, extension and veterinary services. Their average age was 38 years, they had 12 years of education and owned 0.6 ha of land. The non-contract farmers owned 0.4 ha, were 43 years old and went to school for five years.

The study conducted by Sokchea and Culas (2015) has a sample of 75 rice farmers in Cambodia. A farmer cooperative contracts 39 of these farmers providing credit through collection savings, extension, transportation cost, packaging and annual dividends of the cooperative and the credit program. Contract farmers had an average age of 47 years and cultivated on average 2.4 ha of land. In contrast, non-contract farmers were 56 years old on average and cultivated 2.8 ha of land.

Trifković (2014) considered 191 catfish farmers in Vietnam. Here, private firms offered an outgrower contract providing fingerlings, fry, feed, and medicines to 17 percent of the 88 contract farmers or a simple marketing contract without further services to the other 83 percent.

Wainaina et al.'s study (2014) used a sample of 180 poultry growers in Kenya. Of these households, 69 were contracted. However, the contracts were offered by several firms, and no further information is given on the services provided by these contracts. In both groups, the average age was 47 years. Contract farmers had 13 years of education on average and owned 0.5 ha of land. Non-contract farmers, on the other hand, had 12 years of education and worked on 0.4 ha of land.

The sample used by Wang et al. (2014) included 41 farmers from Vietnam that supplied 'safe' vegetables to private firms through farmer cooperatives, but no additional services were provided. The total sample contained 107 observations. On average, farmers selling through contracts were 32 years of age, had eight years of education and owned 0.2 ha of land. Farmers selling to spot markets were 36 years old on average, had eight years of education and owned 0.26 ha.

The study published by Warning and Key (2002) had a small sample of 26 peanut farmer in Senegal. Of these, 15 were contracted by an investor-owned firm, in which the government owns shares. The contract included seeds, agrochemicals, and extension and collection services. Non-contract household heads were 52 years and owned 9.1 ha, while contract suppliers were 48 years of age and owned 9.4 ha, on average.

Winters et al. (2005) used a sample of 189 farmers in Indonesia producing seed corn. Half of them produced under contract. It included credit for land preparation, seeds, agrochemicals, extension, and risk insurance in the form of acceptance of all production regardless of quality. On average the farmers in the sample were 51 years old, had seven years of education and owned a total of 0.66 ha of land.

4.1.3 Excluded studies

Most studies in the core set – i.e. those considered to be effectiveness studies of contract farming – did not meet the criteria for statistical rigour and had to be excluded from the synthesis. A full list of these 53 excluded studies along with the reason for their exclusion is provided in Appendix 9.2.

Four of these studies fell short of inclusion by a very thin margin. Birthal et al. (2009), Roy and Thorat (2008), and Briones (2014) did meet the criteria for statistical rigour but the papers provided insufficient information to calculate effect sizes. These authors either did not respond to our repeated data requests or indicated they no longer had access to the data. Freguin-Gresh et al. (2012) used an appropriate counterfactual method but relied on instruments that were deemed too weak and whose choice was not motivated in the paper. Moreover, average farm size of contracted farmers in this study was around 50 hectares, which was deemed too large to be considered smallholder farming.

One study (Barrett et al. 2012) was excluded as it summarised results from other studies that were included in the review. Four more studies were excluded for having a unique outcome variable, making it impossible to include in the meta-analysis for lack of a relevant comparison group. These studies measured the effect of contract farming on happiness (Dedehouanou et al., 2013), growth rates (Herck et al., 2012), assets (Michelson, 2013), and marketed surplus (Tadesse and Guttormsen 2009). Another study was excluded for providing so little information that it was impossible to understand and interpret the tables and figures (Munungo, 2012).

Five studies were excluded for not being a primary effectiveness study of contract farming. Jabbar and Akter (2013) used contract farming as a control variable in a study of technical efficiency in poultry farming; Lee et al. (2014) did the same in a study of oil palm production in Indonesia; Bamiro et al. (2009) compared full with partial and non-integrated poultry systems, but did not specifically mention contracts nor control for self-selection; Rana et al. (2014) provided descriptive information on potato farming in India, including contract uptake; and Michelson et al. (2012) compared pricing between supermarket and traditional value chains.

All other excluded studies had to be excluded due to an inadequate counterfactual design, which meant they either had no comparison group or relied on simple comparative statistics such as t-tests, which did not control for self-selection bias.

Table 2: Risk of bias assessment

Short Title	Mechanism of assignment bias	Group equivalence bias	Motivation bias	Spill-over effect bias	Selective outcome reporting bias	Selective analysis reporting bias	Other sources of bias
Awotide et al., 2015	Unclear	Unclear	Low	Unclear	High	Low	Low
Bellemare, 2012	Unclear	Low	Low	Unclear	Low	Low	Low
Bellemare & Novak, 2016	Unclear	Unclear	Low	Unclear	Unclear	Low	Low
Bolwig et al., 2009	High	High	Low	Unclear	Low	Unclear	Low
Cahyadi and Waibel, 2011	High	High	Low	Unclear	Low	Unclear	Unclear
Escobal and Cavero, 2012	High	High	Low	Unclear	High	Low	Low
Girma and Gardebroek, 2015	Unclear	High	Low	Unclear	Unclear	Low	Low
Ito et al.,2012	Unclear	Low	Low	Unclear	High	Low	Low
Jones and Gibbon, 2011	Unclear	Unclear	Low	Unclear	High	Low	Unclear
Maertens and Swinnen, 2009	Unclear	High	Low	Unclear	Low	High	Low
Miyata et al.,2009	Unclear	Low	High	Unclear	Low	Low	Unclear
Narayanan, 2014	Unclear	Unclear	Low	Unclear	Low	Unclear	Low
Ramaswami, 2009	Unclear	High	Low	Unclear	Low	Low	Unclear
Rao and Qaim, 2011	High	High	Low	Unclear	High	Low	Low
Saigenji, 2012	Unclear	Unclear	Low	Unclear	High	Low	Unclear
Setboonsarng et al., 2008	High	Unclear	Low	Unclear	Low	Low	Unclear
Simmons et al., 2005	High	High	Low	Unclear	High	Low	Unclear
Sokchea and Culas, 2015	High	High	Low	Unclear	High	High	Unclear
Trifković, 2014	Unclear	High	Low	Unclear	Low	Low	Low
Wainaina et al., 2014	Unclear	Low	Low	Unclear	Low	Low	Unclear
Wang et al., 2014	Unclear	Unclear	Low	Unclear	Low	Low	Low
Warning and Key, 2002	High	High	Low	Unclear	High	Low	Low
Winters et al., 2005	High	Unclear	Low	Unclear	Low	Low	Low

Notes: 'Low' refers to low risk of bias in the relevant domain; 'high' refers to high risk of bias; 'unclear' means that information was not reported in order to assess bias.

4.2 Risk of bias in effect estimates

Although the studies differed considerably regarding context and contractual arrangements, they all faced similar challenges when attempting to identify the impact of contract farming. The validity of the approach to deal with these challenges – with respect to data quality, the way the assessment was conducted, and the analyses and outcomes reported – determined the risk of bias.

The 3ie Risk of Bias tool (Hombrados and Waddington, 2012) consists of eight evaluation criteria, each focusing on a different type of threat to validity. It assesses whether:

- the mechanism of assignment was able to control for confounding.
- the method of analysis was executed in such a way that it ensured comparability of groups throughout the study and prevented confounding.
- the method of being observed did not cause observation bias.
- the study design adequately controlled or corrected for performance bias, such as bias caused by spill-overs.
- there was no evidence of outcomes being selectively reported.
- authors used common methods of estimation and there was no evidence of biased exploratory research methods.
- there was no evidence of other sources of bias, such as bias in the sample of observations selected into the study; concerns about coherence of results; data on the baseline collected retrospectively; information was collected using an inappropriate instrument (or a different instrument/at different time/after different follow-up period in the comparison and treatment groups).
- the statistical significance of the effect was calculated correctly and appropriate statistical tests were used to check the appropriateness of the model.

Each criterion was scored “low risk of bias” if no such bias appeared present, “UNCLEAR” if the study provided insufficient details to assess whether or not the bias was present, and “high risk of bias” if coders determined that there was evidence of bias. The tool contains detailed descriptions of how to attribute a score to each criterion specified per type of study design, described in the coding tool published in the protocol (Ton et al., 2015).

The results of the risk of bias assessments are shown in Table 2. Five sources of bias were found to be particularly important: mechanism of assignment, group equivalence, spillover effects, selective outcome reporting, and other biases arising from sample selection into the study. These sources of bias are discussed in detail in the sections below. The other sources of bias were less important, and will not be discussed in depth. Motivation bias, which evaluates whether households change their behaviour (or answers to survey questions) because they participate in the study, was not mentioned in any of the studies, but are considered less important sources of bias in retrospective studies. Selective analysis reporting was determined to be not a major concern, nor statistical conclusion validity. Data quality and the coherence of the results, were considered as being unclear in only a single case (Cahyadi and Waibel 2011). However, we marked several of the studies as unclear, due to the fact that these studies collected data only after several years, while the studies mention drop-out dynamics in previous years.

4.2.1 Risk of bias due to mechanism of assignment and group equivalence

The risk of bias assessment evaluated the reliability of the statistical methods used to identify the causal impact of contract farming on income. As explained in 3.1.5., the decision to participate in contract farming is non-random, which creates a bias. Three different statistical methods were used to address this selection bias: Propensity Score Matching (PSM) (8 studies), Instrumental Variables (IV)/Heckman (12 studies) or endogenous switching regression models (3 models).

Table 3: Details of risk assessment of 'selection mechanism' and 'group equivalence' for PSM studies

		Awotide et al. 2015	Bellemare & Novak (2016)	Ito et al.2012	Setboonsarng et al. 2008	Maertens & Swinnen 2009	Saigeniji2012	Wainaina 2014	Wang et al. 2014
Mechanism of assignment bias	Authors match on all relevant characteristics	No	No	No	No	No	No	No	No
Group equivalence bias	Match on most relevant characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Matching is based on baseline data	No	No	Yes	No	Yes	No	No	No
	Matching based on time-invariant characteristics	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
	Rosenbaum bounds reported	No	No	Yes	No	Yes	No	Yes	No

PSM is a non-parametric approach which assumes that all household characteristics determining participation in contract farming are observable. Hence, conditional on these selection variables, differences regarding the income level between the matched comparison and treatment groups are assumed to occur randomly and can be attributed to the treatment. Several matching procedures exist, and most studies reported results for both nearest neighbour matching and/or Kernel matching. Ideally, matching relies on baseline data, that is, before the availability of contract farming in the region. None of the studies, however, collected baseline data. Maertens and Swinnen (2009) and Ito, Bao and Su (2012) used recall of baseline conditions to match treatment and control groups. Most PSM studies matched contracted and non-contracted farmers based on current household characteristics and considered these time-invariant or unlikely to be affected by participation in contract farming, such as total landholdings, household size and age of the head of the household. Bellemare and Novak (2016) used the results of the willingness to pay questions as controls in their regressions. Except the study of Sethboonsarng, Leung and Stefan (2008), which did not sufficiently discuss the first stage of the matching procedure, the risk of bias assessment concludes that all PSM studies used relevant and time-invariant variables to match the households (Table 3).

Because propensity score matching relies solely on observable characteristics, the estimates could still be biased if unobservable variables simultaneously determine participation in contract farming and income. In order to assess the sensitivity of the estimated impact to selection on unobservables, Rosenbaum bounds can be calculated (Caliendo and Kopeinig 2008). Although Rosenbaum bounds can provide evidence that the impact of contract farming is significantly positive (or negative), the estimate of the effect size will still be biased if unobservables play a role in the selection equation. As such, the risk of bias is considered to be relatively high for all PSM studies. Moreover, out of the seven PSM studies, only three studies reported (a variant of) Rosenbaum bounds. Most PSM studies were scored as medium or high risk of bias on the criterion on ‘group equivalence’, depending on whether the studies matched on time-invariant variables and reported Rosenbaum bounds (see Table 3 for details).

Four studies used an IV approach and eight studies used a Heckman selection model. Both approaches require an instrument that is correlated with the choice that farmers have to participate in contract farming, but which is uncorrelated with the error term in the regression with income as the independent variable. The quest for such an instrument is challenging. One instrument that was frequently used was distance of the household to a relevant location, such as the village leader (Miyata, Minot and Hu 2009), the rural bank (Birthal, Jha, Tiongco et al. 2009) or the forest (Girma and Gardebroek 2015). Other instruments included the social position of the household (Girma and Gardebroek 2015), the number of formal credit institutions in the village (Escobal and Cavero 2012) or eligibility for contract farming (Jones and Gibbon 2011). Two studies used more original instruments: a contingent valuation experiment (Bellemare and Novak 2016, Bellemare 2012) and a proxy for honesty (Warning and Key 2002). The instruments in five studies were classified as having a medium risk of being exogenous, whereas the instruments in seven studies were considered to have a high risk of bias.

Table 4: Details of risk assessment of 'selection mechanism' and 'group equivalence' for IV/Heckman studies

		Bellemare (2012)	Bolwig et al. (2008)	Cahyadi et al. (2011)	Girma & Gardebroek (2015)	Jones & Gibbon (2011)	Miyata et al. (2009)	Ramaswami (2009)	Simmons et al. (2005)	Sokchea & Culas (2015)	Trifkovic (2014)	Warning & Key (2002)	Winters et al. (2005)
		IV	Heckman	Heckman	IV	IV	Heckman	IV	IV	Heckman	Heckman	Heckman	Heckman
Mechanism of assignment bias	Instrument is credible	Unclear	No	No	Unclear	Unclear	Unclear	No	No	No	Unclear	No	No
Group equivalence bias	Instrumenting equation is significant F>10, R^2 is reported and assessed	Yes	No	No	No	No	No	No	No	No	No	No	No
	Instruments significant at 1%	Yes	No	No	No	Unclear	Yes	Yes	Yes	No	No	No	Yes
	Heckman: mill's ratio reported and significant at 5%	Not relevant	No	No	Not relevant	Not relevant	No	Not relevant	Not relevant	No	No	No	Yes
	Two instruments: over-identifying restrictions are reported	No	No	No	No	Yes	Not relevant	Yes	No	No	No	No	No

For IV or Heckman studies, the score on 'group equivalence' consisted of four elements (Table 4). Studies had to demonstrate that the correlation between the instrument and participation in contract farming was sufficiently strong for correct identification. It appeared that the correlation between the instrument and participation was weak ($p>0.01$) in some studies (e.g. Girma and Gardebroek [2015]), or that the Mill's ratio in the Heckman model was not significant at the 5% level (e.g. Bolwig et al. [2009]). When more than one instrument was used, the Hansen J-test was used to assess whether the over-identifying restrictions should be reported, which would support the validity of the instruments (Jones and Gibbon 2011). Unfortunately, this test was not consistently reported by all the studies that used more than one instrument. Depending on these three criteria, IV and Heckman studies scored low, medium or high risk of bias on 'group equivalence'.

Closely related to instrumental variable approaches are endogenous switching regression models, which were used in three studies. As with instrumental variable techniques, switching regression models require an instrument that explains participation in the program for proper identification. In contrast to PSM, the endogenous switching regression models do not assume that both regimes (spot market or contract farming) have a similar production function. Instead, they verify empirically whether the marginal return on observable characteristics differs between the regimes. This approach requires a larger sample size, as it needs more parameters to accurately estimate the impact of contract farming. Also, the estimate of effect sizes is more sensitive to specification errors. Out of the three endogenous switching regression models, two were considered to have a medium risk of bias because their instruments were found not strong enough (Narayanan 2014, Rao and Qaim 2011), whereas the instruments in the Escobal and Cavero (2012) study were considered to have a high risk of bias on 'mechanism of assignment'. With regards to 'group equivalence', the same criteria were used as for IV studies.

Table 5: Details of risk assessment of 'selection mechanism' and 'group equivalence' for endogenous switching regression models

		Escobal & Cavero (2012)	Narayanan (2014)	Rao & Qaim (2011)
Mechanism of assignment	Instrument is credible	No	Unclear	Unclear
Group equivalence	Instrumenting equation is significant F>10, R^2 is reported and assessed	No	No	No
	Instruments in selection equation are significant at 1%	Yes	No	No

4.2.2 Risk of bias due to spillovers and dynamics

Self-selection of farmers in a contractual arrangement is arguably the most important challenge in identifying the causal impact of contract farming in cross-sectional observational studies. Two other issues which complicate causal identification nevertheless deserve a brief discussion: spillovers and impact dynamics.

Spillovers challenge causal identification. They arise if participation in contract farming by some farmers has a direct or indirect effect on the income of farmers who do not participate in contract farming. An example is spillovers through labour markets. If

contract farming increases demand for labour or increases wages in the region, this may increase income for farmers who do not participate in contract farming but do engage in off-farm labour. Because this will increase the income of the farmers in the control group, the impact of contract farming will be underestimated. This is called the 'contamination effect'. Similar spillovers may arise through product markets (increased output of contracted farmers reduces local prices), input markets (side-selling of inputs provided by the firm to contracted farmers) and credit markets (better access to credit for all farmers in villages with contract farming).

Another spillover effect found in the literature on technology adoption is a learning effect: farmers in the control group may also adopt the technologies offered under the contract (De Janvry, Dustan and Sadoulet 2010). Learning and technology adoption are often an integral part of contract farming. An intriguing example is provided by Schipmann and Qaim (2010) – not included in the meta-analysis – who show that contract farming introduced a new crop, sweet pepper, in local agricultural markets in Thailand. Initially, this new crop was only cultivated by contracted farmers and not sold on local markets. Over time, however, many farmers adopted the new crop and started selling it in local markets. Although the authors did not find a direct positive effect of contract farming on income, they argued that the innovation by the contracting firms improve income for farmers who adopted the new crop and sold it on spot markets.

None of the studies included in the meta-analyses explicitly addressed spillovers or contamination of the control group. However, as most studies sampled the control group from the same village as the treatment group, contamination can be a concern. For this reason, we scored bias due to spillovers as 'unclear' in the risk of bias assessment. Arguably, spillovers are less of an issue for impact evaluations of contract farming than for many other impact evaluations, because in most regions only a minority of the farmers participates in contract farming, which limits general equilibrium effects.

The impact of contract farming is not static, but likely to evolve over time. For instance, the impact on income may increase over time as farmers and the firm learn and optimise production processes. These dynamic spillovers were not addressed by any of the studies. Assessing such dynamic effects would require panel or repeated cross-sectional data, whereas all selected studies used cross-sectional data.

4.2.3 Risk of bias due to selective reporting

Studies were considered guilty of selective outcome reporting when the authors did have data on a (more) relevant outcome variable but chose to base their main analysis on another outcome variable. An example of this was the study of Warning and Key (2002), who used agricultural income in their main analysis because households' income proved not to be significantly different between the groups. Bellemare and Novak (2016) measured the duration of the hungry season, but in their conclusions emphasize the higher effect-size of a derived outcome, 'the likelihood that the hungry season ends'. Another form of selective reporting was merging data. For example, Miyata et al. (2009) in their paper on contract farming in apples and green onions, report descriptive statistics separately for each crop. These statistics show clear differences between both groups. However, in the main analysis, both crops are grouped together, such that these differences are no longer visible. Both forms of selective reporting may be considered as

a kind of publication bias, with authors reporting only those outcomes which show a significant effect, as insignificant results are unlikely to be published.

4.2.4 Other sources of bias

Most studies took place some years after the contractual arrangement had been in place. The surveys use different methods to resolve this selection bias. However, there is a risk of bias to the treatment estimates of a contract farming arrangement because some farmers could have stopped the arrangement in these early years, for example, due to low income effects. These drop-out dynamics can only be measured and controlled for unambiguously with baseline sampling. Several studies mentioned drop-out dynamics in the years before the research took place (Jones and Gibbon 2011, Miyata et al. 2009, Ramaswami, Singh Birthal and Joshi 2009, Saigenji 2012, Sethboonsarng et al. 2008, Simmons, Winters and Patrick 2005, Sokchea and Culas 2015, Wainaina, Okello and Nzuma 2014). We consider endogenous switching and Heckman models as being relatively unaffected by this bias. The other methods use an arithmetic comparison between the outcomes of (matched) participants and non-participants, and, without baseline or follow-up measurements, are vulnerable for the bias of the treatment estimates.

4.2.5 External validity of net-effect estimate

In this section, we examine whether the results of the studies included in the systematic review can be expected to hold for the wider population from which these studies took their samples. This relates to the 'external validity' of the net-effect estimate of inferences made 'within' each study.

External validity depends on the sampling design adopted in the studies as well as how representative the sampled farmers are for the 'average' small-holder farmer. With regards to sampling design, most studies attempted to select households as randomly as possible. Out of the 22 studies, three did not select households randomly, and sampling design was not clearly discussed in another four studies (Table 6).

The remaining 15 studies adopted two distinct sampling strategies. A first strategy consisted of randomly drawing households from a complete list of contracted farmers provided by the contracting firm. Non-contracted farmers were subsequently randomly drawn from a complete list of households living nearby or in the same villages as the sampled contracted farmers. This list was often provided by village heads. In some cases, the control group was randomly selected from a list of households that grew the same crop as the contracted farmers. If a list of contracted farmers was available, this approach was straightforward, cost-effective, and ensured that contracted farmers were indeed randomly sampled.

When researchers did not have access to a complete list of contracted farmers, a different strategy was used. Regions where contract farming was common were purposively selected (based on qualitative information or on an agricultural census). Within these regions, villages were randomly selected and all households were listed and stratified according to participation in contract farming. From this list, contract and non-contracted farmers were randomly selected. In theory, this sampling approach does not exclude that farmers contract to different firms with different 'treatments', that is, different service packages provided by each firm as part of the contractual arrangement.

Table 6: Sampling strategy used in the studies

Author	*Year*	*Country*	*Sampling design*	*Remarks*
Awotide et al.	2015	Nigeria	Sampling villages	Purposive sampling of regions with high density of contract farming
Bellemare Bellemare & Novak	2012 2016	Madagascar	Complete list	Purposive sampling of regions with high density of contract farming
Bolwig et al.	2008	Uganda	Complete list	
Cahyadi & Waibel	2011	Indonesia	Sampling villages	
Escobal & Cavero	2012	Peru	Unclear	Discussed the representativeness of household in sample (compared to the census)
Girma & Gardebroek	2015	Ethiopia	Sampling villages	
Ito et al.	2012	China	Complete list	Purposive sampling of regions with high density of contract farming
Jones & Gibbon	2011	Uganda	Complete list	
Maertens & Swinnen	2009	Senegal	Unclear	Estimation of total contracted farmers, and unclear selection of households within 25 randomly selected villages
Miyata et al.	2009	China	Complete list of 4 contract firms	
Narayanan	2014	India	Combination of complete list and sampling villages	
Ramaswami	2009	India	Sampling villages	
Rao & Qaim	2011	Kenya	Complete list	
Saigenji	2012	Vietnam	Complete list	
Setboornsang et al.	2008	Laos	Unclear	List available, but unclear if randomly selected
Simmons et al.	2005	Indonesia	Sampling villages	
Sokchea & Culas	2015	Cambodia	No random sampling	
Trifkovic	2014	Vietnam	No random sampling	
Wainaina et al.	2014	Kenya	Complete list	
Wanget et al.	2014	Vietnam	Unclear	Sampling strategy not clearly discussed, farmers partly selected by cooperative leaders
Warning & Key	2002	Senegal	No random sampling	Data from a study about impact of a credit allocation program
Winters et al.	2005	Indonesia	Sampling villages	

In sum, all studies attempted to select contracted and non-contracted farmers randomly - although this at times required some creative thinking. However, all were selected from certain geographical locations in a country that are not necessarily representative for the whole country. Most likely, contractual arrangements were set up in regions which benefitted from external factors such as relatively good access to infrastructure or nearby processing facilities. Hence, one cannot conclude that contract farming will work in every

region in these countries. Most studies provided relatively little quantitative information to assess such external factors, which makes it challenging to determine the external validity of the findings.

4.3 Risk of bias in research synthesis

4.3.1 Construct validity

We already discussed extensively the heterogeneity of arrangements grouped under the construct ‘contract farming’. We have argued, already at the protocol stage of this review, that the meta-analyses on a widely heterogeneous sample of empirical instances of contract farming will have important construct validity threats. Therefore, in the forest plots and tables, we also provide information about the type of production and location of each empirical instance in order to facilitate interpretation and reduce the chances for misinterpretation of our findings. There are important construct validity threats to the results of the meta-analysis of contract farming, even when these are grouped in ‘similar types of studies’ to compute pooled average of effect-sizes. For example, when we group cases as being ‘animal production’, this may hide the fact that most of these are related to broiler chicken production, and that we did not include dairy production, which is, probably the most direct association that lay persons make with ‘animal production'.

We show that authors use different proxy-indicators to measure effects on income. For example, authors used crop income, farming income, or household income (see Table 1). We analyse these moderators and show that they affect the pooled average income effect estimate.

The same indicators may also be used to compute different effect-sizes. For example, to assess impact on food security, Bellemare and Novak (2016) measure the ‘length of the hungry season’, but also use the indicator to compute ‘the likelihood that a household's hungry season will end at any given time’, which illustrates the challenge to find unambiguous indicators for food security effects in future systematic reviews. Several other studies that study impacts on food security of production for supermarkets (Chege, Andersson and Qaim 2015) use 7-days food consumption recall. Masset et al. (2011) proposes the use of anthropometric measures to identify food security outcomes.

4.3.2 Publication bias

Scientific articles are more likely to be written and published when they find a significant effect of the program being evaluated (Ioannidis 2005, Ioannidis and Trikalinos 2007). This publication bias is apparent in the studies that we selected for meta-analysis, as only three of the 22 studies report insignificant or negative income effects of contract farming (Cahyadi and Waibel 2011, Narayanan 2014, Simmons et al. 2005). Notably, two of these three evaluated more than one empirical instance of contract farming in their paper, and report a positive effect in most of these other instances. This implies that the academic literature is biased towards studies that find significant effects. The pooled average effect sizes that result from the meta-analysis will inevitably over-estimate the ‘true’ effect of contract farming on income because (many) studies with insignificant effects could not be included in the meta-analysis.

Furthermore, publication bias was assessed formally using funnel plots and Egger’s statistical test (Egger, Smith, Schneider et al. 1997). A funnel plot is a graph that shows

the effect size (horizontal axis) against the precision of the study (vertical axis). In the absence of publication bias, the effect size of the different studies should be distributed symmetrically around the average effect size (the vertical line in the middle). Logically, the effect size of studies with low precision (plotted at the bottom of the graph) will on average deviate more from the pooled effect size than the effect size of studies with a high precision (plotted at the top of the graph), creating a funnel-shaped distribution. An asymmetric funnel plot suggests publication bias.

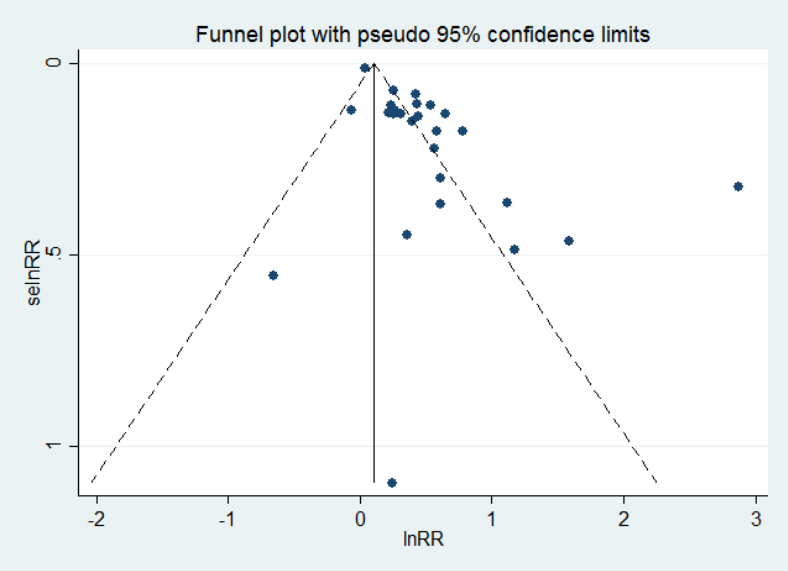
The funnel plot of the studies that measured income effects suggests substantial publication bias (Figure 5). The plot is clearly not symmetrical around the pooled effect estimate. Most studies are outside the 95 percent confidence interval (95%CI) that shows the expected distribution of effect sizes. This asymmetry indicates that there was a strong publication bias, which was confirmed by Egger's test (Table 7). The correlation between effect size and standard error was positive and significant at the 10 percent level, which is considerable given the small sample size (n=26).

Table 7: Meta-regression: examining publication bias based on Egger's test

	(1) Egger's test on all empirical instances	(2) Egger's test using a synthetic effect for Narayanan's instances
Log standard error	1.336*	2.414***
	(1.81)	(4.80)
Constant	0.257	0.0603
	(1.51)	(0.81)
Observations	26	23
R-squared adjusted	24.15%	65.74%
F-statistic	3.274	23.05
tau-squared	0.148	0.0137
I-squared	75.9%	41.9%

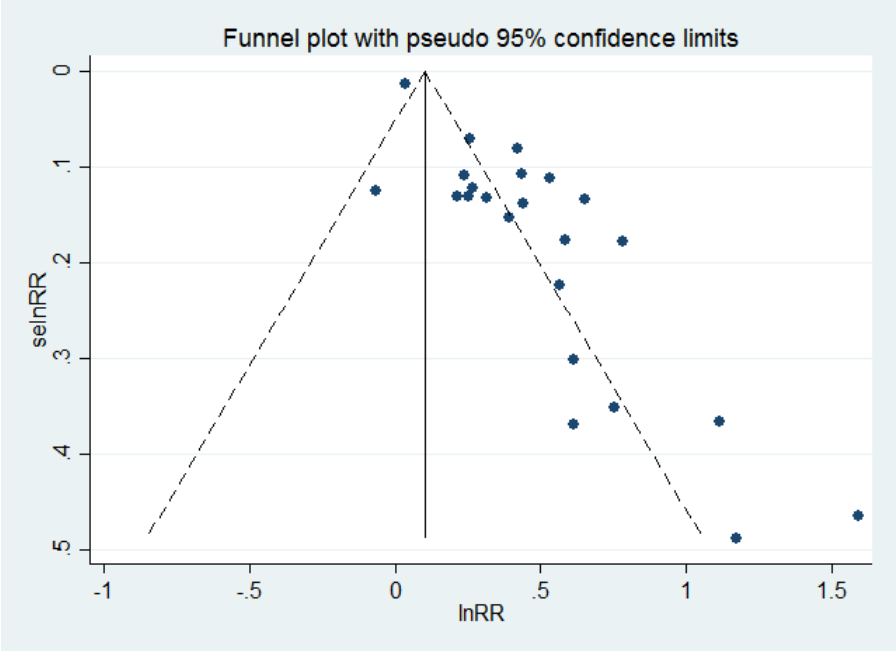
***,**,* indicate statistical significance at the 1, 5 and 10% level, respectively.

Figure 5: Funnel plot



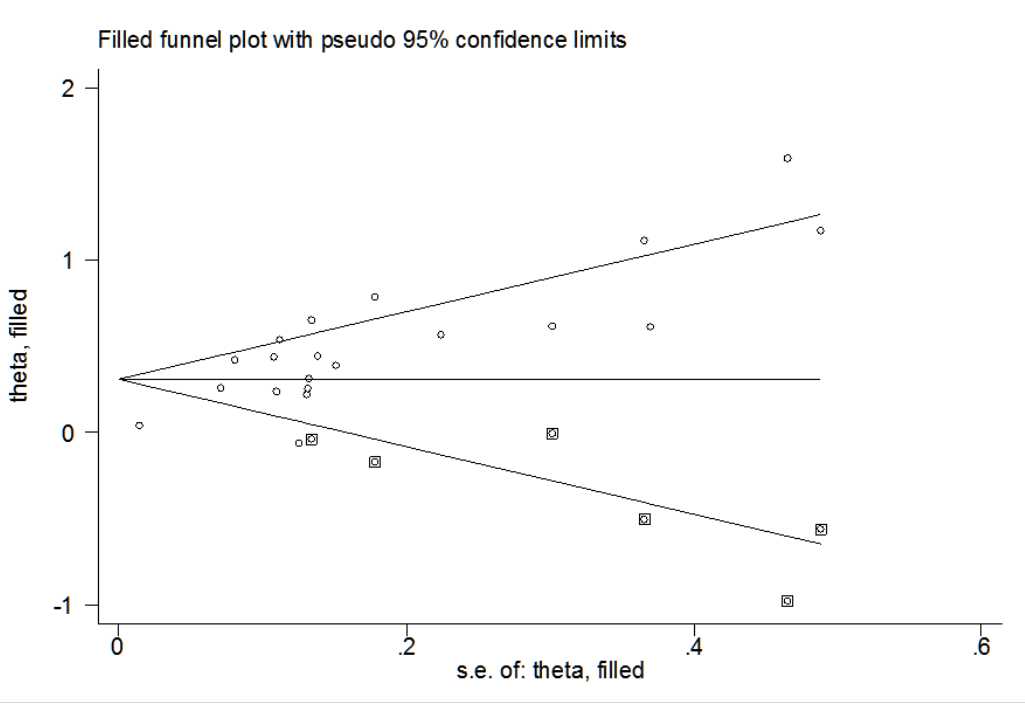
The funnel plot shows a highly imprecise estimated effect size reported by Narayanan (2014) for contract farming of gherkins (the dot at the bottom) and the extremely high effect size for broiler chickens (the dot at the right-hand side of the funnel plot). To test the sensitivity of the result to this single observation, we replaced the four effect sizes reported by Narayanan (2014) by its average synthetic effect size and redrew the funnel plot (Figure 6) and re-conducted the meta-regression (Table 7). This confirms that publication bias due to small study effects is a concern. It also provides evidence for the importance of appropriate analysis of dependency, since the funnel plot is even more asymmetrical and the correlation between effect size and its standard error is highly significant ($p<0.01$).

Figure 6: Funnel plot: replacing four effect sizes by its synthetic effect size



To assess the sensitivity of the results to publication bias, we conducted trim-and-fill analysis (Duval and Tweedie 2000) on the data used for the funnel plot with the synthetic effect for Narayanan (2014). This method exploits the fact that, without publication bias, a funnel plot can be expected to be symmetrical and also include the less favourable results (in our case these are the observations on the left-hand side of the funnel plot). To mimic this, the trim-and-fill exercise imputes missing studies to the funnel plot to achieve this symmetry. The pooled effect size is subsequently re-estimated including these non-existent studies. The trim-and-fill analyses added seven non-existent studies to the funnel plot (Figure 7) and estimates the pooled effect estimate at 1.38 (95%CI=1.23, 1.55). This is less than the pooled effect estimate without trim-and-fill error correction (RR=1.53; 95%CI=1.35, 1.74), reported below. In other words, the positive impact of contract farming on income decreased but remained significant once missing small studies are imputed using trim-and-fill methods. However, some caution is warranted when interpreting these results. The validity of trim-and-fill methods is questionable if there is substantial between-study heterogeneity (Peters, Sutton, Jones et al. 2007). This is clearly the case in our meta-analysis of contract farming, which varies in crop type or services provided.

Figure 7: Funnel plot of a trim-and-fill analysis



4.3.3 Survivor bias

In addition to publication bias, the meta-analysis is affected by survivor bias, which causes an overestimation of the pooled average effectiveness of contract farming. All studies used cross-sectional surveys to assess effectiveness after the contract farming arrangement had been in place for several years (see Table 8). Therefore, logically, all the empirical instances of contract farming covered by the studies needed to be operational at the time of the research. This implies that per definition the meta-analysis did not include studies on contract farming arrangements that had ceased. All studies covered empirical instances that already managed to survive the initial tensions between firm and farmers about prices, services and quality requirements, which are mentioned in the professional literature as being important challenges to contract farming (ActionAid 2015, Barrett et al. 2012, Bijman 2008, Da Silva and Rankin 2013, Eaton and Shepherd 2001, FAO 2008, Narayanan 2013, Oya 2012, Prowse 2012, Ton and Mheen-Sluijer 2009, Will 2013, Williamson 2003). Because failed attempts to establish a functional contractual relationship could not be studied with a cross-sectional design, their results could not be included in the meta-analysis. Therefore, the pooled average effect size is an upward-biased estimate of the effectiveness of contract farming. However, the response rates may well be indicative for the effect size that is required for contract farming arrangements to be maintained over time.

Table 8: Timing of impact evaluation and start of contract farming

Author	Country	Product	Start of the contractual arrangement	Year of data collection
Awotide et al., 2015	Nigeria	Rice	No unique contractual arrangement	2013
Bellemare, 2012; Bellemare and Novak, 2016	Madagascar	Green beans, leek, snow peas, rice, barley	No unique contractual arrangement. The main company operates since the early 1990s	2008
Bolwig et al., 2009	Uganda	Coffee	2000	2005
Cahyadi and Waibel, 2011	Indonesia	Oil palm	1989-1994 and 1995-2000.	2010
Escobal and Cavero, 2012	Peru	Potato	2000	2002/2003
Girma and Gardebroek, 2015	Ethiopia	Honey	2007	2009
Ito et al.,2012	China	Watermelon	2000	2009
Jones and Gibbon, 2011	Uganda	Cocoa	2001/2002	2005 & 2009
Maertens and Swinnen, 2009	Senegal	French beans	No unique contractual arrangement.	2005
Miyata et al.,2009	China	Apples, green onions	No details	2005
Narayanan, 2014	India	Marigold, papaya, broiler, gherkins	No details	2009/2010
Ramaswami, 2009	India	Broiler	No details	2002/2003
Rao and Qaim, 2011	Kenya	Various vegetables	No unique contractual arrangement	2008
Saigenji, 2012	Vietnam	Tea	1950s	2007
Setboonsarng et al., 2008	Laos	Rice	2002	2004
Simmons et al., 2005	Indonesia	Seed rice	1988	2002
Simmons et al., 2005	Indonesia	Broiler	1998	2002
Sokchea and Culas, 2015	Cambodia	Rice	2003	2010
Trifković, 2014	Vietnam	Catfish	No unique contractual arrangement	2010
Wainaina et al., 2014	Kenya	Broiler	No details	2010/2011
Wang et al., 2014	Vietnam	Various vegetables	Multiple firms involved, since 1995	2007/2008
Warning and Key, 2002	Senegal	Peanuts	1990	1992 & 1994
Winters et al., 2005	Indonesia	Seed corn	1986	2002

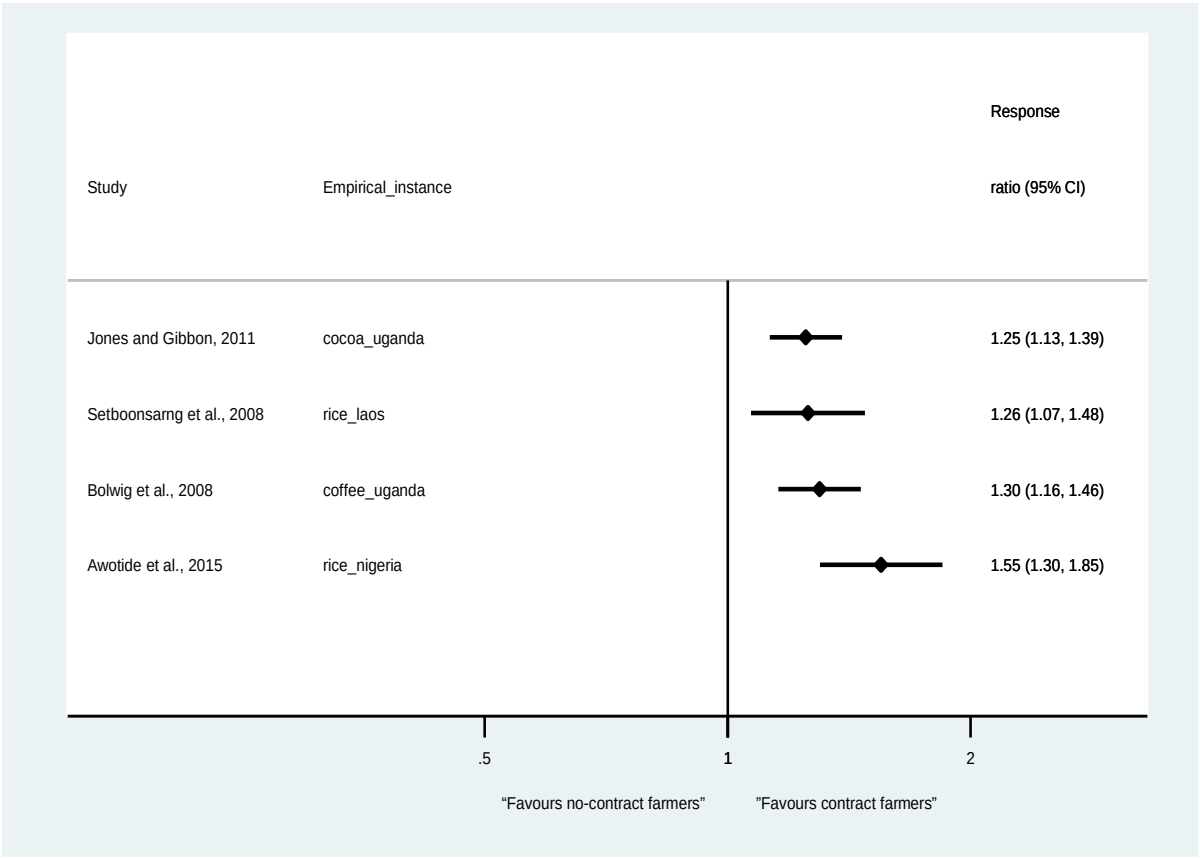
4.4 Synthesis of effects

This section reports the results of the sensitivity, moderator and meta-regressions that tested the impact of contract farming on the well-being of small-scale farmers. First, we present the pooled effect estimate of all studies included in the meta-analysis. We already pointed to the substantial heterogeneity across studies. To control for this observed heterogeneity, we conducted a sensitivity and moderator analysis. Sensitivity analysis examines whether the design of the study, including the estimation strategy and the reported outcome, explain the heterogeneity across studies. Moderator analysis tests whether crop or contract characteristics explain the between-study heterogeneity. Finally, the robustness of the bivariate sensitivity and moderator analysis was examined using meta-regressions.

4.4.1 Overall results

In Figure 3, we present the intervention logic of the development effect of contract farming. We depict as immediate outcomes the farmer's and firms' incentive to continue the contractual arrangement. Intermediate outcomes are the effects of the contract farming arrangement on production. The ultimate outcomes relate to income and food security.

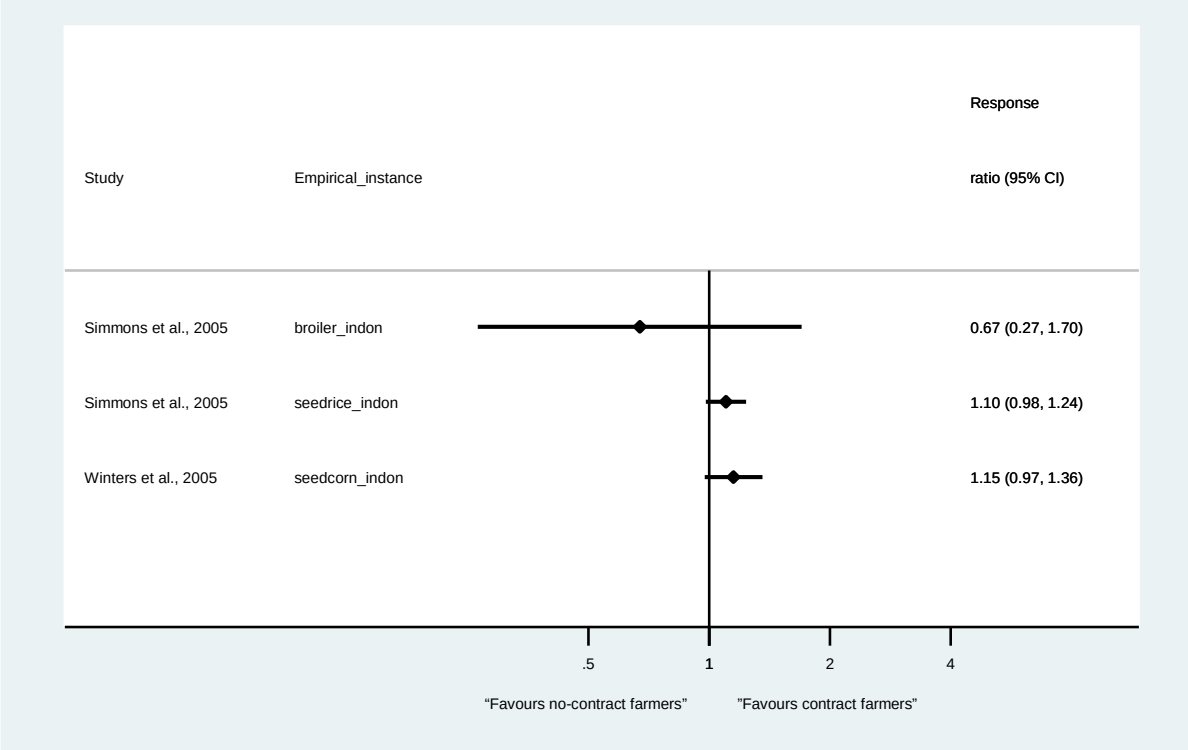
Figure 8: Forest plot of the effect of contract farming on yields.



Intermediate outcomes

From several studies, we could distill net-effect estimates of the intermediate outcomes (see Table 1). Because we only include intermediate outcomes reported in studies that also report ultimate outcomes, and the constructs used to measure these intermediate outcomes differ importantly between studies, reporting pooled average effect sizes is not appropriate. Some studies estimated yield effects of contract farming (see Figure 8). Awotide et al. (2015) found a yield effect of 55% for traditional rice. Setboonsarng et al. (2008) report a 26% yield increase for organic rice. Bolwig et al. found a 30% increase in coffee yields as a result of improved organic production practices. Jones and Gibbon (2011) found a yield effect of 25%. When the contracted crop is new, or a new variety, and only produced by the contract farmers, it is not possible to determine a yield effect. For example, some contract farming arrangements make it easier for farmers to use chemical inputs, such as fertiliser or pesticides, to grow special varieties. Winters et al. (2005) and Simmons et al. (2005) describe this mechanism for contract farming in seed corn production in Indonesia, where input use more than doubled as a result of contract farming and household labour use increased with 20%.

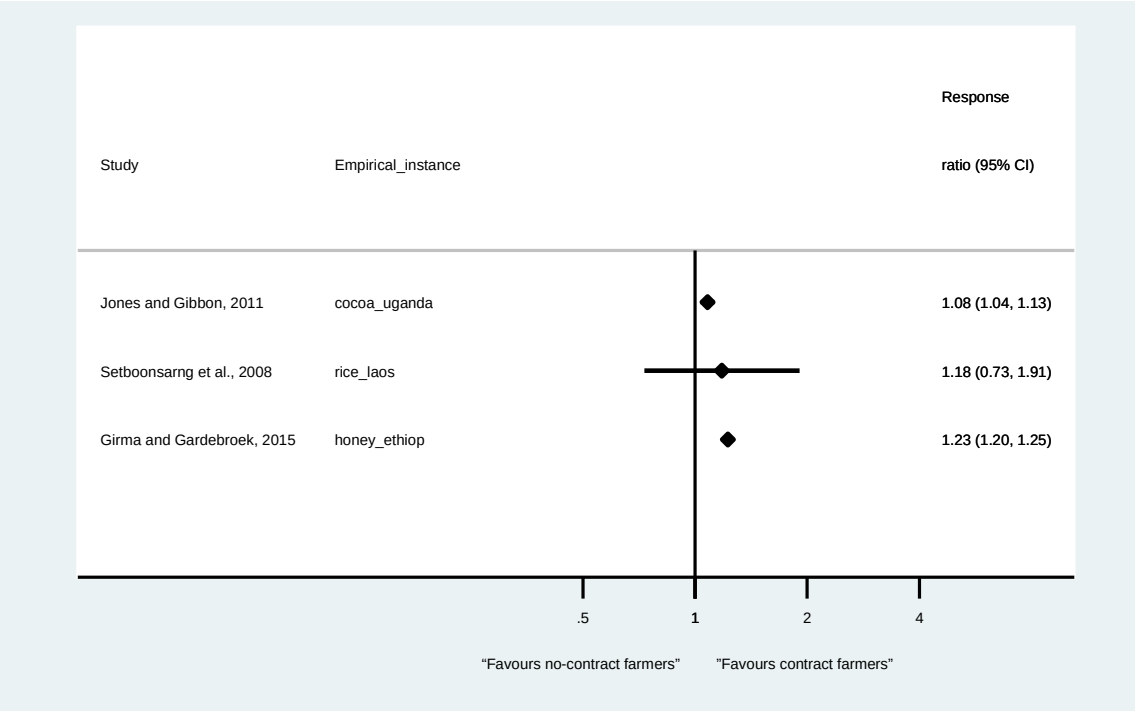
Figure 9: Forest plot of the effect of contract farming on household labour.



The effects on household labour were also reported in the other contractual arrangements (see Figure 9) covered by Simmons et al. (2005). They found an increase in the use of household labour for seed rice production (RR=1.10) but a decrease of household labour use in broiler chicken production (RR=0.67).

Jones and Gibbon (2011), for organic cocoa in Uganda, and Setboonsarng et al. (2008), for organic rice in Laos, estimated the net-effects on yields and prices separately, in order to analyse the main drivers of income effects. The reported price effects are 8% in organic cocoa and 18% in organic rice production. Girma and Gardeborek (2015) report a price increase of 23% due to the contractual arrangement in organic honey production. Prices (and price premiums) are crucially important in all other contract farming arrangements. However, in most of the studies prices were considered as a (fixed) component of the treatment (service package) provided by the firm, not as an outcome of the treatment, which explains the narrow confidence interval in two of the studies.

Figure 10: Forest plot of the effect of contract farming on prices.

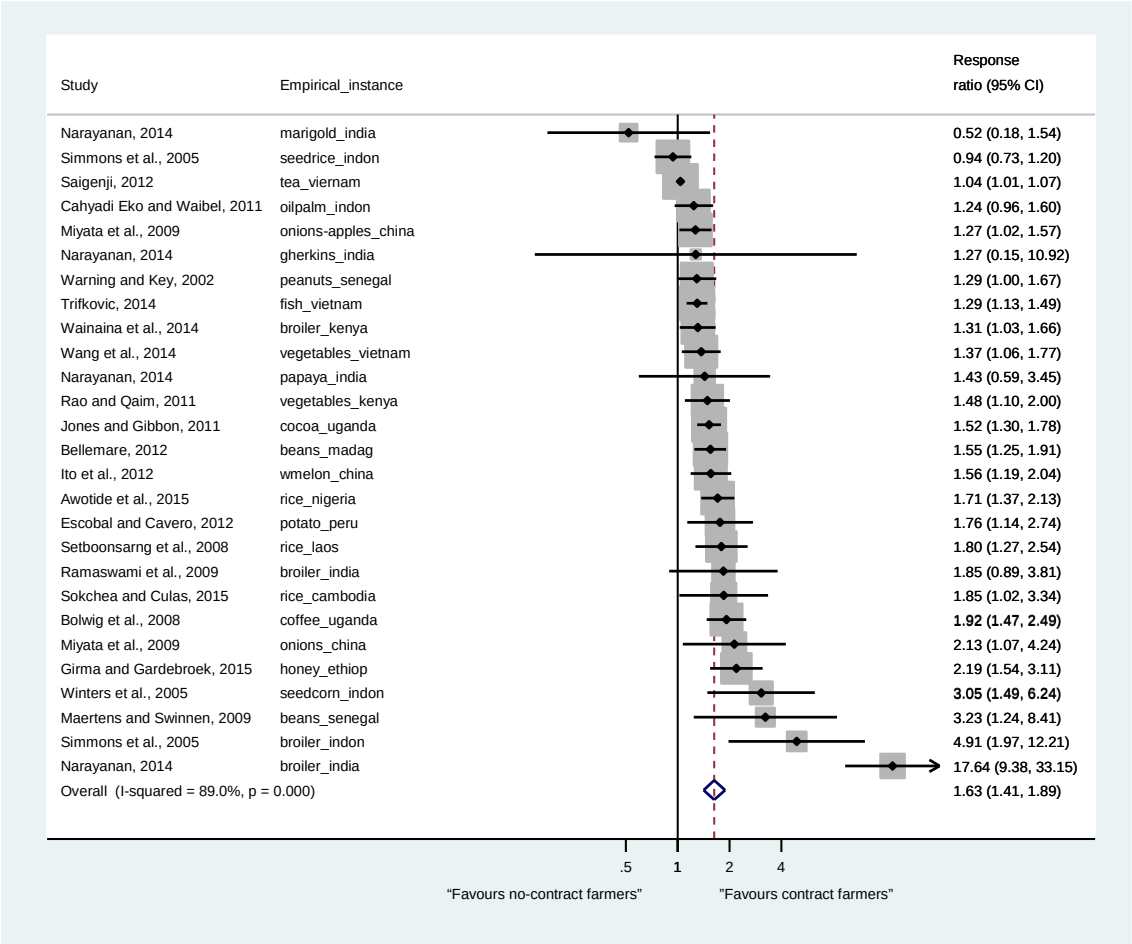


Ultimate outcomes

The ultimate outcomes considered in this review are household income, livelihood strategies, market power and household food security. We retrieved no study that reported net-effects of contract farming on market power and livelihood strategies. Within the set of 23 studies, only one reported net-effects on household food security. Bellemare and Novak (2016) reported an average effect size of 8% reduction in the length of the hungry season (RR=0.92; CI=0.85-1.00). In their conclusions, they also report ‘the likelihood that a household's hungry season will end at any given time’, with an effect size of 18% (RR=1.19; CI= 1.12-1.26).

Twenty-two studies estimated income effects. The forest plot (Figure 11) shows the effect size of contract farming on income for each empirical instance. Effect sizes are reported in response ratios. Results from the meta-analysis indicate that in the 26 empirical instances covered by the 22 studies, contract farming increased income on average by 62% (RR=1.63, CI=1.41–1.89, I-squared=89.0%, tau-square=0.0978)

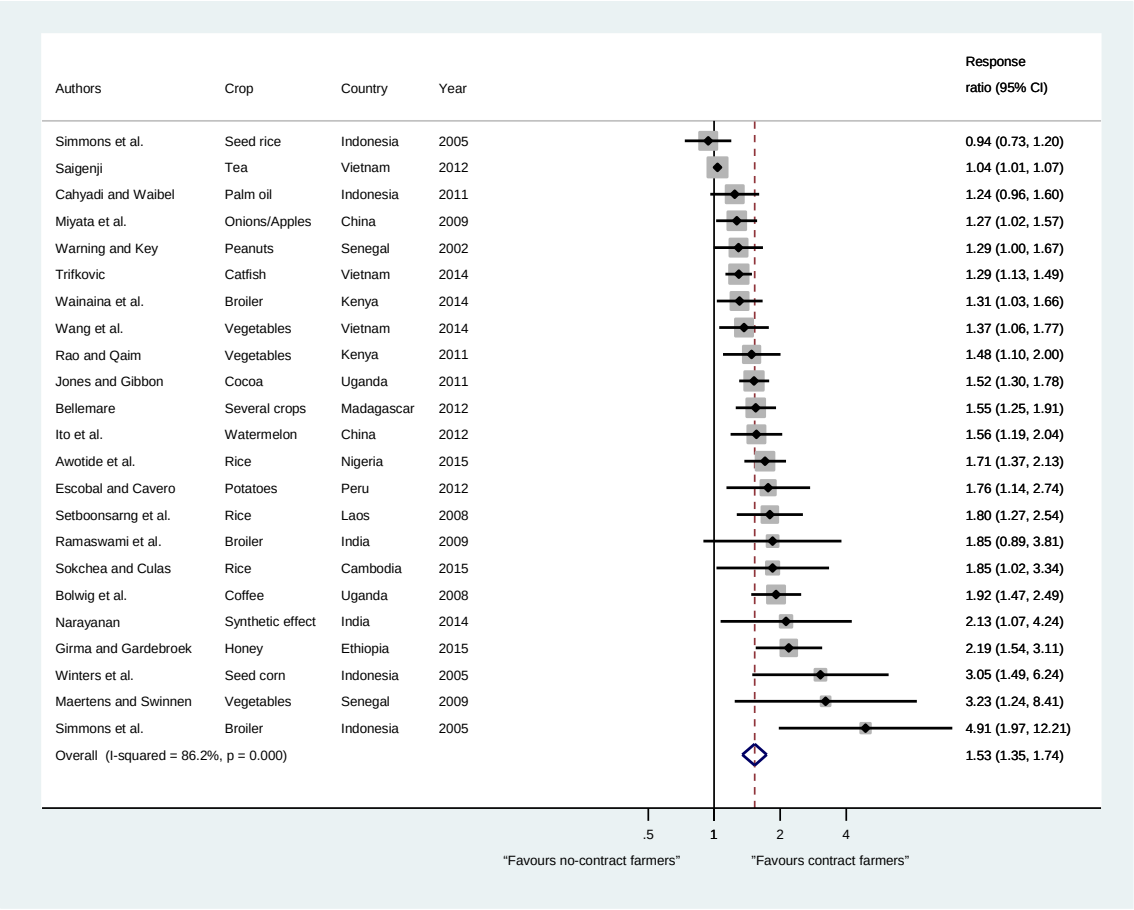
Figure 11: Forest plot of the effect of contract farming on income: all empirical instances



The pooled average effect reported in Figure 11 is an overestimation due to the inclusion of four empirical instances of contract farming which used the same control group. In Figure 12, we present the forest plot that uses the synthetic effect for Narayanan (2014). The pooled average effect size of the meta-analysis is 1.53 (CI=1.35-1.74, I-squares=86.2, tau-squared=0.065).

Out of the 26 empirical instances of contract farming covered by the studies, only two had a negative effect on income (see above), and even in those studies, the negative effect was not significantly different from zero. Five studies reported that contract farming more than doubled income. The largest effect was reported by Narayanan (2014), who found that contract farming increased income from broiler chicken more than 17-fold. The substantial heterogeneity between studies was confirmed by the low tau-squared and the large I-squared statistic. The I-squared statistic indicated that most of the variation (89.2% using all empirical instances; 86.2% using the synthetic effects for Narayanan) might be attributed to heterogeneity between studies rather than sampling variation within studies. The heterogeneity was (partly) due to differences in the contractual arrangements.

Figure 12: Forest plot of the effect of contract farming on income: synthetic effect for Narayanan, 2014.



In line with expectations, contract farming had a less pronounced effect when household income (RR=1.32, CI=1.13-1.54 was measured, compared with studies in which farming income (RR=1.65, CI=1.17-2.33) or income from the contracted crop (RR=1.92, CI=1.47-2.50) was measured. No study provided sufficient data to estimate net-effects at crop, farm and household level. In most of the studies that used crop income (in coffee, cocoa, broiler chicken) it provides a good estimate of overall household income effects. Also, when endogenous switching models were used, the main factor substitution effects are controlled for in the regression, and the results can be considered valid proxies of household income effects.

One study reported the effects of this income increase on the household’s poverty status. Awotide et al. (2015) reported 55% effect on rice yield (RR=1.55, CI=1.30-1.85), which contributed to a 71% increase of crop income (RR=1.71, CI=) and a 20% reduction in poverty status (RR=1.20,CI=1.00-1.44).

4.4.2 Sensitivity analysis

The sensitivity analysis examined if the estimated pooled effect size was sensitive to (1) the outlier of broiler farming in India, (2) study design, and (3) the reported outcome variable. We report the results of this bivariate analysis in Table 9.

The empirical instance of broiler production in India (Narayanan 2014) had a substantial effect on the pooled effect estimate. This large response ratio probably results from the very low alternative income measured for non-contracted farmers living in the same area, using the same plot size for other activities than broiler chicken production. Typically, broiler production takes place in sheds, on small plots, and located near urban areas. A comparison with other (non-agricultural) investment opportunities in the urban area would likely have provided additional insight into the possible counterfactual situation of the broiler producers. It is, in any case, evident that this empirical instance is a severe outlier. Such a severe outlier may also be expected to affect the results of the moderator and sensitivity analyses. Excluding this particular empirical instance reduced the pooled response ratio to 1.50 (CI=1.32-1.70). Replacing the four empirical instances reported in Narayanan's study by a synthetic effect reduced the response ratio to 1.53 (CI=1.35-1.74). The heterogeneity across studies, however, decreased only slightly (I-sq=86%) when it was excluded. This indicates that between-study heterogeneity is very high, even when we exclude these outliers.

Results suggested that the econometric design did not affect the estimated pooled effect estimates (Table 9). A meta-analysis by econometric strategy (IV, PSM or endogenous switching regression) showed no significant differences between the pooled effect estimates. The heterogeneity between studies was lower for IV-studies (I-sq=70%) than for studies using switching regressions (I-sq=91%) or PSM (I-sq=88%).

Pooled effect estimates differed with respect to the reported outcome variable (Table 9). In line with expectations, contract farming had a less pronounced effect when household income (RR=1.32, CI=1.13-1.54 was measured, compared with studies in which farming income (RR=1.65, CI=1.17-2.33) or income from the contracted crop (RR=1.92, CI=1.47-2.50) was measured. The variance was very high, which means that these differences were not statistically significant. However, these results do suggest that substitution effects may play a role – i.e. substituting land and labour from other activities towards the crop under contract. The two studies that measured income by consumption expenditure (RR=1.15, CI=0.92-1.42) reported lower response ratios than studies that measured income (RR=1.69, CI=1.45-1.96). By contrast, no difference in RRs was observed between studies reporting net income (RR=1.70, CI=1.42-2.04) or gross income (revenue) (RR=1.68, CI=1.33-2.11).

Table 9: Sensitivity analysis

	RR	*95% confidence interval*		*Q*	*Tau-sq*	*I-sq*	*Number of instances*
All studies	1.62	1.40	1.88	232.21	0.10	89.23	26
Outlier: broiler farming in India							
Synthetic effect size	1.53	1.35	1.74	159.61	0.07	86.22	23
Exclude outlier	1.50	1.32	1.70	158.35	0.06	84.84	25
Research design							
IV	1.51	1.31	1.74	40.38	0.04	70.28	13
PSM	1.47	1.16	1.86	48.93	0.08	87.74	7
Switching regressions	2.02	0.84	4.87	56.47	0.98	91.15	6
Outcome measurement							
Expenditure	1.15	0.92	1.42	9.40	0.02	89.36	2
Income	1.69	1.45	1.96	108.28	0.09	78.76	24
Definition of income							
Household	1.32	1.13	1.54	40.27	0.03	82.62	8
Farming	1.65	1.17	2.33	23.36	0.13	78.60	6
Contracted crop	1.92	1.47	2.50	67.11	0.15	83.61	12
Revenue or net income							
Revenue	1.68	1.33	2.11	6.25	0.03	51.97	4
Net income	1.70	1.42	2.04	100.71	0.11	81.13	20

4.4.3 Moderator analysis

A bivariate moderator analysis examined if the estimated pooled effects differed with respect to (1) type of production; (2) continent, and (3) contract characteristics (Table 10). Contract farming of animal products was more remunerative (RR=2.69, CI=1.55-4.65) than annual (RR=1.50, CI=1.30-1.73) or perennial products (RR=1.49, CI=1.13-1.96). The forest plot (figure 13) shows that this higher effectiveness of contract farming in animal products was driven by the example of broiler farming. We argued in 4.3.1. that a counterfactual analysis through a comparison with alternative investment opportunities in these semi-urban areas would most likely have resulted in smaller net-effect estimates.

We are aware of the wide heterogeneity included under these labels, and, therefore, present in each forest plot the type of production and location of each empirical instance, to facilitate interpretation of the results.

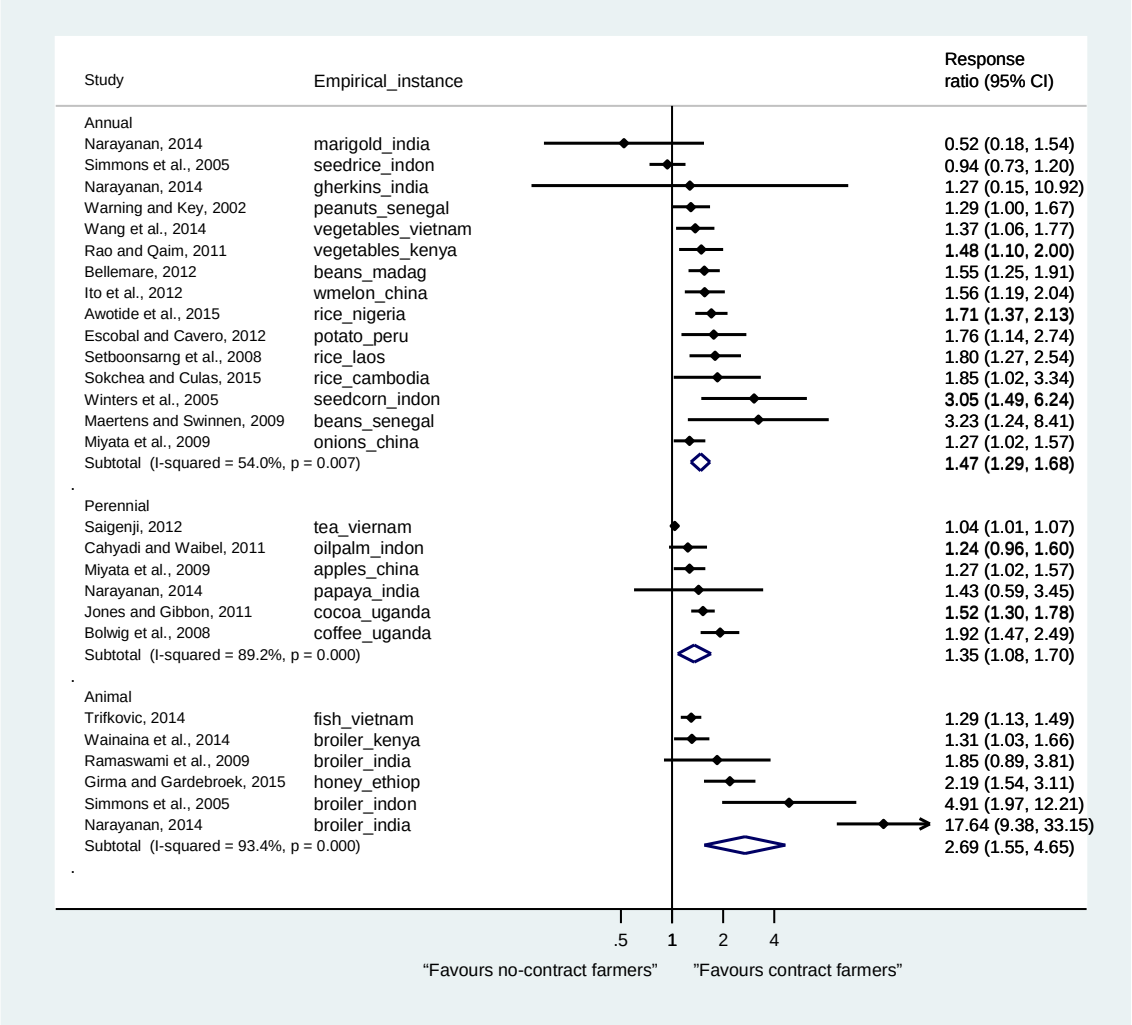
Table 10: Moderator analysis

	RR	95% confidence intervals		Q	tau-sq	I-sq	Sample size
Crop characteristics							
Animal husbandry	2.69**	1.55	4.65	75.42	0.40	93.37	6
Annual crop	1.47	1.29	1.68	30.46	0.03	54.04	15
Perennial crop	1.35	1.07	1.70	46.50	0.06	89.25	6
Perishable product	1.61	1.31	1.98	129.61	0.10	89.20	16
Non-perishable product	1.63	1.37	1.94	29.92	0.05	69.92	10
Region							
Africa	1.59	1.42	1.77	13.22	0.01	39.47	9
Asia	1.64	1.34	2.00	142.27	0.11	88.75	17
South-America	1.76	1.14	2.74	0.00	0.00	100.00	1
Contract characteristics							
Existing cooperative							
Yes	1.41	1.17	1.70	60.53	0.06	85.13	10
No	1.80**	1.47	2.20	89.90	0.11	82.20	17
Price premium							
Yes	1.65**	1.51	1.81	12.64	0.00	5.07	13
No	1.50	1.23	1.84	133.80	0.10	91.03	13
Transport							
Yes	2.49**	1.69	3.66	65.46	0.24	89.31	8
No	1.41	1.24	1.61	99.26	0.05	81.87	19
Credit							
Yes	1.65**	1.51	1.81	12.64	0.00	5.07	13
No	1.43	1.28	1.61	27.25	0.02	52.30	14
Seeds							
Yes	1.64	1.36	1.98	172.23	0.11	89.55	19
No	1.58	1.38	1.81	11.86	0.01	40.97	8
Key inputs							
Yes	1.72*	1.43	2.06	187.20	0.11	89.85	20
No	1.47	1.22	1.77	18.43	0.04	67.44	7
Extension services							
Yes	1.70**	1.43	2.02	220.18	0.12	90.01	23
No	1.36	1.22	1.52	2.14	0.00	0.00	4
Level of services provided							
High	1.78**	1.44	2.19	181.22	0.15	90.07	19
Low	1.43	1.23	1.66	20.08	0.03	65.14	8

[a]. Apples and onions in Miyata et al (2009) are analysed as two separate instances of contract farming.

*. ** Significantly higher than contract farmers without this characteristic with $p<0.10$, $p<0.05$.

Figure 13: Forest plot of the effect of contract farming on income per crop type.



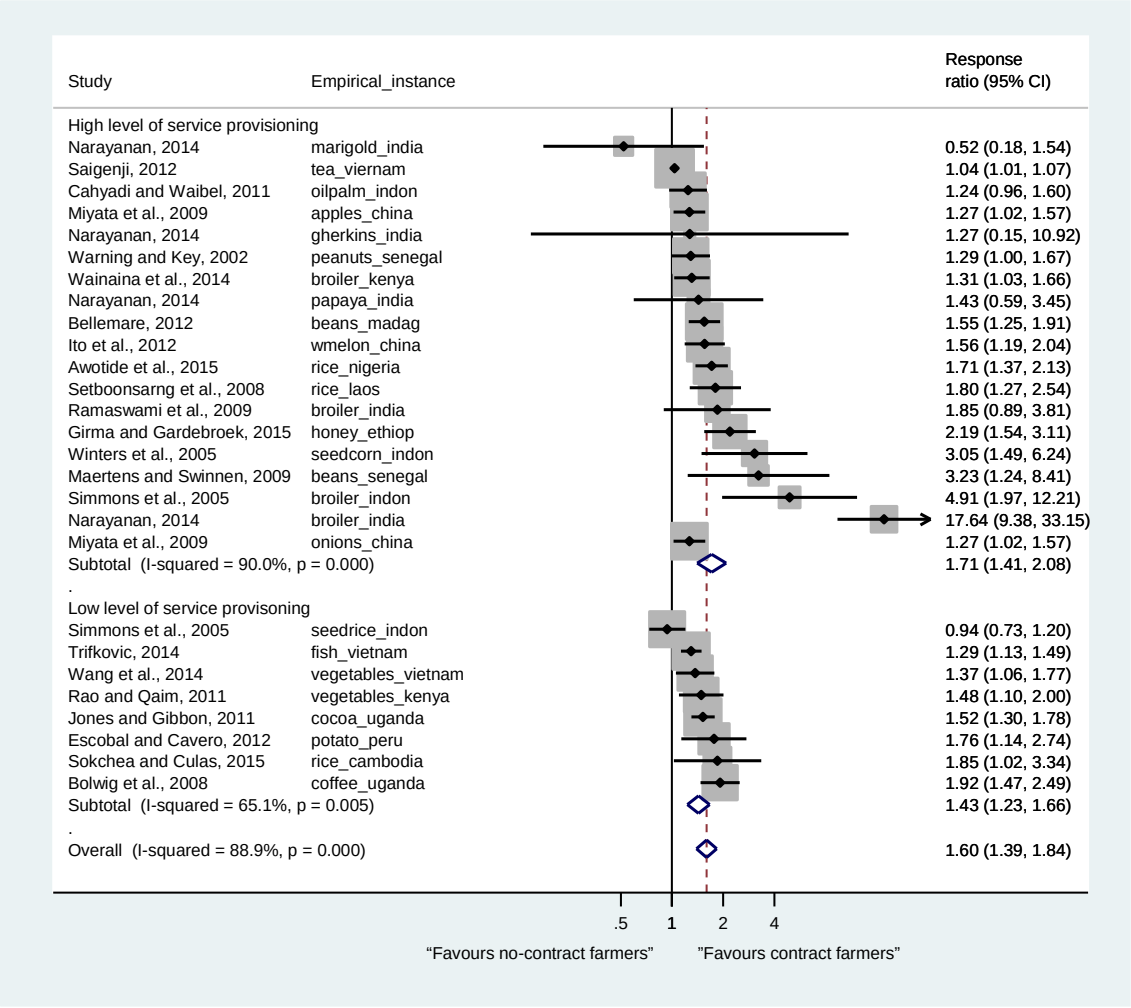
We also analysed the moderating effect of geographical location, using the continent as a proxy indicator. The positive effect of contract farming on income did not differ between Africa and Asia. No comparison was done with the single study from South America. Figure 14 shows the forest plot according to the level of service provisioning. All contractual arrangements studied included services provided by the firm to the farmer. Five services were frequently encountered as part of the contractual arrangement: credit, (improved) seeds, agrochemicals, extension and transport services (see Table 11). Unfortunately, not all studies detailed the inclusion of the services in the contract.

Table 11: Overview of the heterogeneity of service provisioning in the empirical instances covered by the meta-analysis

	Study	Empirical instance	transport	credit	seeds	inputs	extension
1	Awotide et al., 2015	rice_nigeria	1	0	1	1	1
2	Bellemare, 2012	beans_madag	1	1	1	1	1
3	Bolwig et al., 2009	coffee_uganda	1	0	0	0.2	1
4	Cahyadi and Waibel, 2011	oilpalm_indon	0	0.8	1	1	1
5	Escobal and Cavero, 2012	potato_peru	0	0	1	0	0
6	Girma and Gardebroek, 2015	honey_ethiop	0	1	0	1	1
7	Ito et al.,2012	wmelon_china	0	0	1	1	1
8	Jones and Gibbon, 2011	cocoa_uganda	0	0	0	0	1
9	Maertens and Swinnen, 2009	beans_senegal	1	1	1	1	1
10	Miyata et al.,2009	apples_china	0	0.8	0.2	1	1
11	Miyata et al.,2009	onions_china	0	0.8	1	1	1
12	Narayanan, 2014	marigold_india	0	0	1	0.8	1
13	Narayanan, 2014	papaya_india	0	0.2	1	0.8	1
14	Narayanan, 2014	broiler_india	1	1	1	1	1
15	Narayanan, 2014	gherkins_india	0	0.2	1	1	1
16	Ramaswami, 2005	broiler_india	1	1	1	1	1
17	Rao and Qaim, 2011	vegetables_kenya	0	0	0	0	0
18	Saigenji, 2012	tea_vietnam	0	0	1	1	1
19	Saigenji, 2012	tea_viernam	0	1	1	1	1
20	Setboonsarng et al.,2008	rice_laos	0	0.8	1	1	1
21	Simmons et al., 2005	seedrice_indon	0	0	1	0	1
22	Simmons et al., 2005	broiler_indon	1	1	1	1	1
23	Sokchea and Culas, 2015	rice_cambodia	0	0	0	0	1
24	Trifković, 2014, 2016	fish_vietnam	0	0	1	1	0
25	Wainaina et al., 2014	broiler_kenya	0.8	0.8	0.8	0.8	0.8
26	Wang et al., 2014	vegetables_vietnam	0	0	0	0	0
27	Warning and Key, 2002	peanuts_senegal	0	0	1	1	1
28	Winters et al., 2005	seedcorn_indon	0	1	1	1	1

Note: 0 means absence, 1 means presence: when information was not provided we inferred: 0.2 means likely absence, 0.8 means likely present.

Figure 14: Forest plot by level of service provisioning



Contractual arrangements were classified as having a ‘high level of service provisioning’ if they provided at least three of the five services, and as having a ‘low level of service provisioning’ if they offered fewer than three services (Table 10). It is noteworthy that contracts that offered transport to farmers had a substantially larger impact on income (RR=2.49, CI=1.69-3.66) than contracts that did not include transport (RR=1.41, CI=1.24-1.61). This highlights that access to distant markets through a firm might be one of the main benefits of contract farming for farmers. The service index confirms that contractual arrangements offering at least three services were associated with significantly larger effects on income (RR=1.78, CI=1.44-2.19) than those offering fewer services (RR=1.43, CI=1.23-1.66). These results appear not to be driven by the single empirical instance of broiler farming in India. As these moderator variables may suffer from collinearity, we explore these differences in more detail in the next section.

4.4.4 Meta-regressions

Meta-regressions were conducted in order to control for several sensitivity and moderator variables simultaneously (Table 12). We only included variables that were shown to be statistically significant in the bivariate analyses, in order to preserve degrees of freedom. Nevertheless, due to the limited sample size, the power of the analyses was low and the results should be interpreted with care.

Table 12: Meta-regression on moderator and sensitivity variables for income effects of contract farming

	(1)	(2)	(3)	(4)	(5)
Animal husbandry	0.518*		0.168	0.205	0.159
	(2.02)		(1.04)	(1.40)	(1.12)
Measurement of household income (expenditure=0; other=1)		0.206	0.284	0.349**	0.231
		(1.26)	(1.43)	(2.18)	(1.38)
Income level (baseline: household)					
Farming income		0.0214	0.0312		
		(0.14)	(0.19)		
Income from contracted crop		0.156	0.103		
		(1.18)	(0.68)		
Low level of service provisioning	-0.132		-0.0193	-0.0147	-0.0020
	(-0.59)		(00.16)	(-0.13)	(-0.02)
Dummy for outlier of broiler farming in India (Narayan)		2.371***	2.244***	2.265***	2.175***
		(5.59)	(5.07)	(3.31)	(5.09)
Standard error (log)					0.849
					(1.44)
Constant	0.444***	0.136	0.070	0.051	0.032
	(3.00)	(1.05)	(0.40)	(0.31)	(0.21)
Observations	26	26	26	26	26
R-squared adjusted	9.1%	90.2%	85.7%	86.3%	89.0%
F-statistic	2.464	10.31	6.782	10.79	9.35
tau-sq	0.202	0.0222	0.0318	0.0305	0.0245
I-sq	87.5%	59.4%	57.3%	57.1%	55.3%

***,**,* coefficient statistically significant at the 1%, 5% and 10% levels, respectively; values are unstandardized meta-regression coefficients, with their corresponding standard errors in parenthesis.

The first and second regressions only included the moderator and sensitivity variables, respectively. The third regression included all moderator and sensitivity variables simultaneously, whereas the fourth regression included all variables that were significant in at least one of the previous regressions. In the fifth regression, the log of the standard errors was included to reduce publication bias. The analyses confirmed the sensitivity of the results to the single empirical instance of broiler farming in India (Narayanan 2014). The meta-regressions suggested that there was no statistically significant effect of any of the moderating factors. Measuring income by expenditure was associated with lower RRs compared with other methods. The direction of the coefficients is indicative only, and some moderators might become significant with higher sample size. The adjusted R-squared and tau-squared show that several models explained most of the heterogeneity between studies. Note, however, that the outlier of broiler farming contributed most to explaining the variance.

4.4.5 Factors influencing participation

All included studies had a design and reporting quality that made it possible to extract the indicators used to assess differences between participating and non-participating farmers. Table 12 summarises the indicators used to assess whether the contractual arrangements were accepted by larger or smaller farmers, and by asset-rich or asset-poor farmers. These indicators were extracted from the descriptive statistics on

characteristics of the treatment and comparison group (t- or Z-tests) or, when not reported, from the results of the econometric analysis to determine the factors that influence participation in the contract farming arrangement, usually a Probit analysis. Generally, the results of the t-tests on the descriptives, and the results of the first stage regressions (usually Probit models) coincided. However, when the results did differ, we used the more straightforward t-test on the treatment and comparison groups, because in the Probit analysis, differences between participants and non-participants are not always reflected as significant coefficients in the regression due to collinearity with other variables in the equation.

Table 13 shows that in 52% of the studies, the contracted farmers had significantly more land (12 out of 23). Only in 17% of the cases (4 out of 23) did participation in the contract tend to involve the relatively smaller farms. Except broiler farming, the instances of contract farming that involved relatively smaller farms showed substantially lower income effects (response ratios) than the average, which suggests that for the larger farmers these relatively lower benefits might not outweigh the costs of participating in the contract. When differences in asset endowments between contracting and non-contracted farmers were taken into account - for those studies containing such information – the tendency to contract relatively better-endowed farmers was even more accentuated. Only in one of the 15 studies that provided data on assets, did contracted farmers have significantly fewer assets than the comparison group; in 66% of the cases, contracted farmers were significantly wealthier in terms of assets (10 out of 15). When we combine both analyses, we see that in 61% of the empirical instances covered in the studies (14 out of 23) the contracted farmers were better off than the non-contracted farmers in the sample.

The above analysis suggests that contract farming tends to attract relatively larger and wealthier farmers. However, scale and asset ownership are relative. It might well be that even the relatively larger and wealthier strata are still small scale in absolute terms (small scale was indeed a selection criterion in the search procedure). To verify this, data on average landholdings was extracted from the various studies. We selected this variable for two reasons. First, it was reported in all studies, except for the study on honey collectors of Girma and Gardebroek (2015) , and the study of Ito et al. (2012), who registered the cultivated area of the contracted crop instead of total land area owned. Second, no assumptions are required to compare landholdings between regions and over time. The boxplot (Figure 14) shows that, with the exception of the study of Warning and Key (Warning and Key 2002) in Senegal, landholdings were below 6 ha for all studies. Most studies covered smallholder farmers, with landholdings in the range of 0.5 to 3 ha.

Figure 15: Average landholding in selected studies.

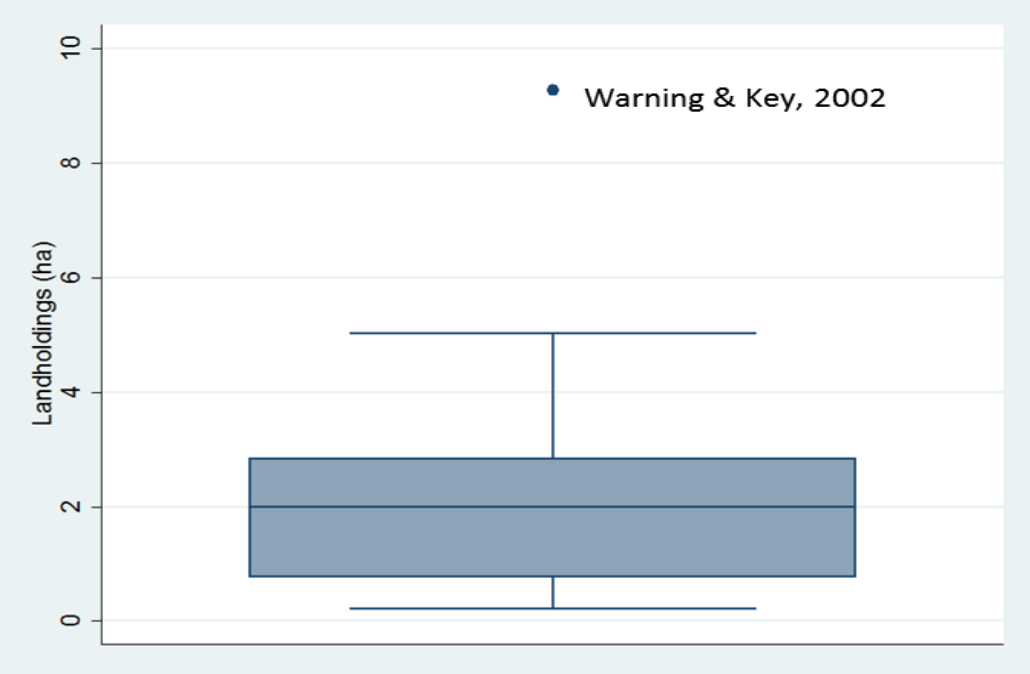


Table 13: Assessment of differences in scale and wealth between participants and non-participants

	Author	Product	Proxy-indicators for scale	++	+-	--	Proxy-indicators for wealth	++	+-	--	Income effect (RR)
1	Awotide et al. (2015)	rice	none				none				1.71
2	Bellemare (2012);	various crops	landholding	1			equipment and inputs	1			1.55
3	Bolwig et al. (2009)	coffee	number of trees, farm size	1			(house) walls	1			1.92
4	Cahyadi and Waibel (2011)	oil palm	land size	1			total assets	1			1.24
5	Escobal and Cavero (2012)	potato	land size	1			productive assets	1			1.76
6	Girma and Gardebroek (2015)	honey	traditional hives	1			none				2.19
7	Ito et al. (2012)	watermelon	none (area watermelon)				total assets	1			1.56
8	Jones and Gibbon (2011)	cocoa	farm size, cocoa trees	1			none				1.52
9	Maertens and Swinnen (2009)	green beans	landholding in 1995	1			non-land assets		1		3.23
10*	Miyata et al. (2009)	apples	land cultivated, irrigated			1	agricultural assets	1			1.27
11*	Miyata et al. (2009)	green onions	land cultivated, irrigated	1			agricultural assets		1		
12	Narayanan (2014)	gherkins	land owned			1	none				1.27
13	Narayanan (2014)	marigold	land owned		1		none				0.52
14	Narayanan (2014)	papaya	land owned		1		none				1.43
15	Narayanan (2014)	broiler	land owned		1		none				17.64
16	Ramaswami (2009)	broiler	(un)irrigated lands			1	none				1.85
17	Rao and Qaim (2011)	various crops	land area	1			none				1.48
18**	Saigenji (2012)	state - tea	none				none				no data
19	Saigenji (2012)	private - tea	none				none				1.04
20	Setboonsarng et al. (2008)	rice	none				none				1.8
21	Simmons et al. (2005)	broiler	dry and irrigated land	1			(non)agric. assets,	1			4.91
22	Simmons et al. (2005)	seed rice	dry and irrigated land	1			(non)agric. assets		1		0.94
23	Sokchea and Culas (2015)	rice	cultivated land		1		agric. equipment		1		1.85
24	Trifković (2014, 2016)	catfish	aquaculture area		1		asset index			1	1.29
25	Wainaina et al. (2014)	broiler	land size		1		total assets	1			1.31
26	Wang et al. (2014a)	various crops	total land area			1	none				1.37
27	Warning and Key (2002)(2002)	peanuts	land cultivated		1		agric. equipment	1			1.29
28	Winters et al. (2005)	seed corn	(irrigated) area operated	1			assets incl. land	1			3.05
TOTAL				12	7	4		10	4	1	

*. Analysed as one instance in the meta-analysis; **. Not part of meta-analysis

Table 14: Dataset of cases and conditions for the qualitative comparative analysis

		Conditions					Response ratio	Outcome	
Studies	QCA name	existing farmer org	price premium	credit in cash	seeds/ breeds	key inputs	income effect size	high effective	low effective
Awotide et al., 2015	rice_nigeria	0.8	0	0	1	1	1.71	1	0
Bellemare, 2012	beans_madag	0.8	1	1	1	1	1.55	0.97	0.03
Bolwig et al., 2009	coffee_uganda	0	1	0	0	0.2	1.92	1	0
Cahyadi and Waibel, 2011	oilpalm_indon	0	0	0.8	1	1	1.24	0.47	0.53
Escobal and Cavero, 2012	potato_peru	0	1	0	1	0	1.76	1	0
Girma and Gardebroek, 2015	honey_ethiop	0	1	1	0	1	2.19	1	0
Ito et al.,2012	wmelon_china	0.2	1	0	1	1	1.56	0.98	0.02
Jones and Gibbon, 2011	cocoa_uganda	0	1	0	0	0	1.52	0.96	0.04
Maertens and Swinnen, 2009	beans_senegal	0.2	1	1	1	1	3.23	1	0
Miyata et al.,2009	apples_china	0.8	0	0.8	0.2	1	1.27	0.56	0.44
Miyata et al.,2009	onions_china	0	0	0.8	1	1	1.27	0.56	0.44
Narayanan, 2014	broiler_n_india	0	0	1	1	1	17.64	1	0
Narayanan, 2014	gherkins_india	0	1	0.2	1	1	1.27	0.56	0.44
Narayanan, 2014	marigold_india	0	0	0	1	0.8	0.52	0	1
Narayanan, 2014	papaya_india	0	1	0.2	1	0.8	1.43	0.9	0.1
Rao and Qaim, 2011	vegetables_kenya	0.8	0	0	0	0	1.48	0.94	0.06
Saigenji, 2012	broiler_w_india	0	0	1	1	1	1.85	1	0
Saigenji, 2012	tea_viernam	1	0	1	1	1	1.04	0.07	0.93
Setboonsarng et al., 2008	rice_laos	0.2	1	0.8	1	1	1.8	1	0
Simmons et al., 2005	broiler_indon	0	0	1	1	1	4.91	1	0
Simmons et al., 2005	seedcorn_indon	1	1	1	1	1	3.05	1	0
Sokchea and Culas, 2015	rice_cambodia	1	1	0	0	0	1.85	1	0
Trifković, 2014, 2016	fish_vietnam	0	0	0	1	1	1.29	0.62	0.38
Wainaina et al., 2014	broiler_kenya	0	0	0.8	0.8	0.8	1.31	0.67	0.33
Wang et al., 2014	vegetables_vietnam	0.8	1	0	0	0	1.37	0.81	0.19
Warning and Key, 2002	peanuts_senegal	0.2	0	0	1	1	1.29	0.62	0.38
Winters et al,. 2005	seedrice_indon	1	1	0	1	1	0.94	0.02	0.98

Source: Authors

4.4.6 Factors associated with effectiveness

For half of the 28 empirical instances of contract farming, we could link the study selected for meta-analysis with other (peer-reviewed) literature that offered additional information on contextual conditions or content of the service package delivered under contract to the participating farmers (Table 14). Using qualitative comparative analysis, we created an overview of all possible combinations of conditions (configurations) and listed the cases that shared the same combination. This matrix is called a 'truth-table'. We analysed the truth-table for combinations of conditions that were consistently related with being highly effective or being a less effective. To make the analysis less sensitive for the threshold chosen to differentiate between highly and less effective instances of contract farming, we used fuzzy-set scores. We used the pooled average effect size of 38 percent as the cross-over point, and a response rate of 1.50 as the threshold value to define the relatively high and 1.25 to define relatively the low effective contractual arrangements. Intermediate values (1.25<RR<1.50) were converted into fuzzy-set scores using the logistic function provided by the software application fsQCA™ (Ragin and Davey 2009). The truth-table shows which cases (empirical instances of contract farming) and associated combinations of conditions were consistent with a relatively high or relatively low effectiveness on income. The set-consistency score is the lowest value of the fuzzy-score of the case, be it in the set of conditions or the in the outcome.

With the 27 empirical instances for which we had sufficient information, we could only use a limited number of conditions for the QCA. We focused on the five conditions that appeared most important as predictors of effectiveness, based on the moderator analysis and our reading of the literature. For some studies, we could not classify the empirical instance as 'present' or 'absent' without caveats, due to missing information, and we used fuzzy-set scores to account for this uncertainty. A fuzzy-set score of 0.2 means that the condition is most likely absent, whereas 0.8 indicates that it is most likely present. The scoring was based on contextual information. This fuzzy-set ambiguity was taken into account when the set consistency score was calculated for each truth-table row, or for each combination of conditions in the QCA-solutions.

Table 15 provides an overview of all empirical instances covered in the studies according to the logical combination of incentives and services provided in the contracts. The study of Trifković (2014) on catfish in Vietnam combined different types of marketing contracts. Using additional data provided by the author (Trifković 2016), we were able to disaggregate her analysis and use the information on farmers that had input-provisioning contracts in the qualitative comparative analysis. Sixteen of these combinations (the rows in Table 13) were unique and covered by only one empirical instance. This was a reflection of the *diversity in extremis* that characterised the contractual arrangements included in the meta-analysis. Nevertheless, some conclusions may be drawn based on this table. The broiler cases demonstrate some uniformity, with four cases sharing a similar combination. The contracts for broiler production provided full-service packages, and the destination of the chicken was a market that did not provide a higher price than the local market. The two studies that analysed vegetables as a generic crop type (Rao and Qaim 2011, Wang et al. 2014a), documented no service provisioning and can be considered borderline cases of contract farming. The contract included only support for compliance with quality requirements in addition to the forward-sales agreement. These studies would have been excluded had we applied a narrower interpretation of our

definition of contract farming. Several other studies on the impact of supermarkets were excluded because these marketing contracts did not include technical support in production as a service provided by the firm to the farmers, but only provided marketing contracts.

Table 15: Combination of services and incentives embedded in the contracts

	Existing cooperative involved	Price-premium above local market	Transport provided	Credit in cash provided	Seeds provided	Key agro-inputs provided	On-farm extension provided	Empirical instances covered
1	Present	Present	Present	Present	Present	Present	Present	beans_madagascar
2	Present	Present	Not	Present	Present	Present	Present	seedcorn_indonesia
3	Present	Present	Not	Not	Present	Present	Present	seedrice_indonesia
4	Present	Present	Not	Not	Not	Not	Present	rice_cambodia
5	Present	Present	Not	Not	Not	Not	Not	vegetables_vietnam
6	Present	Not	Present	Not	Present	Present	Present	rice_nigeria
7	Present	Not	Not	Present	Present	Present	Present	tea_vietnam
8	Present	Not	Not	Present	Not	Present	Present	apples_china
9	Present	Not	Not	Not	Not	Not	Not	vegetables_kenya
10	Not	Not	Present	Present	Present	Present	Present	broiler_n_india broiler_w_india broiler_indonesia broiler_kenya
11	Not	Present	Not	Not	Present	Present	Present	wmelon_china gherkins_india papaya_india
12	Not	Not	Not	Present	Present	Present	Present	oilpalm_indonesia onions_china
13	Not	Not	Not	Not	Present	Present	Present	marigold_india peanuts_senegal
14	Not	Present	Present	Present	Present	Present	Present	beans_senegal
15	Not	Present	Present	Not	Not	Not	Present	coffee_uganda
16	Not	Present	Not	Present	Present	Present	Present	rice_laos
17	Not	Present	Not	Present	Not	Present	Present	honey_ethiopia
18	Not	Present	Not	Not	Present	Not	Not	potato_peru
19	Not	Present	Not	Not	Not	Not	Present	cocoa_uganda
20	Not	Not	Not	Not	Present	Present	Not	fish_vietnam

The conditions are considered present or absent and do not take into consideration the 'intensity' of the service provisioning, e.g. the quality of the services or quantity of the resources implied. Furthermore, to limit the large number of combinations (2^7= 128) that are logically possible when using the seven conditions presented in Table 14, we reduced the number of conditions to the four that appeared to make most difference between the cases. These were: existing cooperative involved (present/not present), price premium above local market (true/false), provisioning of credit in cash (present/not present), and provisioning of key inputs (present/not present). These four conditions result in a truth-table with 16 rows (2^4) containing all possible combinations, with 11 of these rows covered by one or more empirical case (Table 16).

Table 16: Truth-table with all contract farming arrangements in the sample

Conditions used as possible predictors				Sets of cases		Outcome pattern		
Existing farmer organisation	Price premium	Credit	Key inputs	N	Set consis-tency	Outcome	High-effectiveness	Low-effectiveness
Present	Present	Present	Present	2	1	True	beans_madagascar; seedcorn_indonesia	-
Present	Present	Present	Not	0	n/a	-	-	-
Present	Present	Not	Present	1	0.3	False	-	seedrice_indonesia
Present	Present	Not	Not	2	1	True	rice_cambodia; vegetables_vietnam	-
Present	Not	Present	Present	2	0.35	False	-	apples_china; tea_vietnam
Present	Not	Present	Not	0	n/a	-	-	-
Present	Not	Not	Present	1	1	True	rice_nigeria	-
Present	Not	Not	Not	1	1	True	vegetables_kenya	-
Not	Present	Present	Present	3	1	True	honey_ethiopia; beans_senegal; rice_laos	-
Not	Present	Present	Not	0	n/a	-	-	-
Not	Present	Not	Present	3	0.91	Unclear	wmelon_china; papaya_india	gherkins_india
Not	Present	Not	Not	3	0.99	True	coffee_uganda; potato_peru; cocoa_uganda	-
Not	Not	Present	Present	6	0.88	Unclear	broiler_n_india; broiler_w_india; broiler_indonesia	oilpalm_indonesia onions_china; broiler_kenya
Not	Not	Present	Not	0	n/a	-	-	-
Not	Not	Not	Present	3	0.62	False	-	marigold_india; fish_vietnam; peanuts_senegal
Not	Not	Not	Not	0	n/a	-	-	-

We expected, based on the moderator analysis, that some combinations of conditions could be enablers or barriers of effectiveness in a certain type of production and not in other. Therefore, to limit the diversity in contractual arrangements due to crop characteristics, we present the truth-tables separately for perennial crops, annual crops and animal husbandry. We are aware of the wide heterogeneity included under these labels, and, therefore, added the country and crop name to facilitate interpretation of the results. These reduced truth-tables (Tables 17, 18 and 19) excluded the logical combinations of conditions that are not covered by the cases in the sample (so-called ‘logical remainders’).

For six contractual arrangements related to perennial crops (Table 17), we found suggestive evidence that a price premium may be essential in the incentive structure for high effectiveness. It was present in the two high-effective cases and absent in the two less-effective cases. Interestingly, the absence of credit seemed consistently related to higher effectiveness. This suggests that in perennial crops, a ‘lock-in’ situation due to credit obligations may explain why farmers did not opt-out of the contract even if it was not very effective. In the studies on oil palm (Cahyadi and Waibel 2011) and tea (Saigenji 2012), the authors explicitly mentioned the possibility of farmers becoming locked into less effective contracts, though they did not find evidence that this indeed took place.

For annual crops (Table 18), the price premium in combination with credit and inputs was also an important ingredient of the recipe for high effectiveness. It was absent in only two of the highly effective cases. In both cases, however, we see that existing farmer organisations were involved in brokering and governing the contractual arrangements. This suggests that a price premium is required for a contractual arrangement to become highly effective on income, especially when cooperatives cannot serve as an intermediary between the firm and the farmer.

For the animal husbandry cases (Table 19), the role of an existing farmer organisation as broker or intermediary in a contractual arrangement between a firm and farmers was not reported. The provision of both inputs and credit did not result in high effectiveness in the Kenyan broiler case. However, this combination appears to be a likely enabler of effectiveness, as evidenced by its presence in all other highly effective contracts arrangements related to animal husbandry.

Table 17: Truth-table of contract farming related to perennial crops

Existing farmer organization	Price premium	Credit	Key inputs	N	Set consis-tency	Outcome	High-effective	Low-effective
Not	Present	Not	Present	1	1	True	papaya_india	-
Not	Present	Not	Not	2	0.98	True	coffee_uganda; cocoa_uganda	-
Present	Not	Present	Present	2	0.35	False	-	apples_china; tea_vietnam
Not	Not	Present	Present	1	0.67	False	-	oilpalm_indonesia

Table 18: Truth-table of contract farming related to annual crops

Existing farmer organization	Price premium	Credit	Key inputs	N	Set consis-tency	Outcome	High-effective	Low-effective
Present	Present	Present	Present	2	1	True	beans_madagascar; seedcorn_indonesia	-
Present	Present	Not	Not	3	0.65	Unclear	rice_cambodia; vegetables_vietnam	seedrice_indonesia
Present	Not	Not	Present	1	1	True	rice_nigeria	-
Present	Not	Not	Not	1	1	True	vegetables_kenya	-
Not	Present	Present	Present	2	1	True	beans_senegal; rice_laos	-
Not	Present	Not	Present	2	0.87	Unclear	watermelon_china	gherkins_india
Not	Present	Not	Not	1	1	True	potatoes_peru	-
Not	Not	Present	Present	1	0.7	False	-	onions_china
Not	Not	Not	Present	2	0.51	False	-	marigold_india; peanuts_senegal

Table 19: Truth-table with contract farming arrangements related to animal husbandry

Existing farmer organization	Price premium	Credit	Key inputs	N	Set consis-tency	Outcome	High-effective	Low-effective
Not	Present	Present	Present	1	1	True	honey_ethiopia	-
Not	Not	Present	Present	4	0.97	Likely	broiler_indonesia; broiler_n_india; broiler_w_india	broiler_kenya
Not	Not	Not	Present	1	0.68	False	-	fish_vietnam

5. Discussion

5.1 Summary of main results

5.1.1 What is known about the effect size of contract farming?

The search and screening process identified 23 studies. Only one study reported on food security (Bellemare and Novak 2016). They measure the 'length of the hungry season', and report the reduction in time (9% effect, RR=0.92) and compute 'the likelihood that a household's hungry season will end at any given time' (18% effect, RR=1.18). This study relied on the same methodology and data as another study by the same author, which was included in the meta-analysis of income effects (Bellemare, 2012). However, this indicates that there is a lack of serious research on the effects of contract farming on food security, whereas this issue is prominent in the critique of certain international NGOs (for example, ActionAid 2015, Holt-Giménez, Williams and Hachmyer 2015).

For the meta-analysis of income effects, we could use 22 studies. The meta-analysis suggests that on average, the income effect of participation in a contractual arrangement, measured with the respective proxy-indicator of income in each study, is 62% (RR=1.62, CI=1.40-1.88). The empirical instance of broiler production in India (Narayanan 2014) was a clear outlier and had a substantial effect on the pooled effect estimate. Replacing the four empirical instances reported in Narayanan's study by their synthetic effect reduced the overall response ratio from 1.62 to 1.53 (CI=1.35-1.74). For the eight studies that used household income as the proxy-indicator, the pooled effect size was 1.32 (CI=1.13-1.54), for the six studies that used farm income it was 1.65 (CI=1.17-2.33) and for the 12 studies that measured effects in the income derived from the crop or animal husbandry, it was 1.92 (CI=1.47-2.50).

To assess publication bias, we regressed the effect sizes on the standard errors and showed that there was a large asymmetry in the funnel plot, which is considered a strong indicator of small study effects and therefore possible publication bias. We showed that all included studies reported positive effects for at least one of the empirical instances covered in their study. A trim-and-fill exercise to partially reduce this publication bias resulted in a pooled average effect-size of 1.38 (CI=1.23-1.55). We generated the hypothesis that these relatively large positive effects on income might be an important incentive for farmers to give up their autonomy in markets, production, and quality handling. Contract farming arrangements need to offer clear benefits to farmers in order to survive over time, especially in a context of free entry and exit of farmers (Narayanan 2013, Wendimu, Henningsen and Gibbon 2016).

We showed that even this partially corrected pooled average effect size of 38% is upwardly flawed due to survivor bias: most studies took place in empirical instances of contract farming that had survived their first difficult years. This implies that the meta-analysis did not account for the effects of the empirical instances that did cease to exist and that it will inevitably overestimate the contract farming average treatment effect. It is unlikely that new contract farming arrangements will on average result in these large positive effects. Modest expectations and careful planning is needed for contract farming to be effective and sustainable. All empirical instances covered by the studies related to contract schemes that had already been operational for several years at the time of data collection. We found information on only one empirical instance that was reported as being collapsed after the study period, the Laos organic rice case, reported in Campbell et al. (2012). The professional literature suggest that failed attempts to establish contract farming arrangements with smallholder farmers are quite common, and stress the need for adaptive management and dispute settlement mechanisms to prevent failure (Da Silva and Rankin 2013, Da Silva and Shepherd 2013, Eaton and Shepherd 2001, FAO 2008, Prowse 2012, Ton and Mheen-Sluijer 2009, UNIDROIT-FAO 2016, Will 2013).

The information retrieved from the effectiveness studies did not permit strong inferences on spill-over effects. Other literature strongly indicated the positive effects of contract farming pilots for sector-wide technological innovation in the area of incidence (Euler, Schwarze, Siregar et al. 2016, Otsuka, Nakano and Takahashi 2016, Schipmann and Qaim 2010), especially in the adoption of new crops, cultivation practices, post-harvest handling and access to new markets. We could not assess potential negative effects of, for example, indebtedness, land ‘grabbing’, gender discrimination and crop intensification, whereas these are key concerns for organisations that are critical of contract farming, as for example ActionAid (2015). However, the negative livelihood effects are unlikely to endure if farmers are free to step in or out of the contract. Unintended environmental effects of the contract farming arrangements have not been reported in any of the studies. Interestingly, many of the studies selected in this review cover production contracts under improved environmental management (e.g. organic, certified) that are better than conventional production.

5.1.2 Under which enabling or limiting conditions are contract farming arrangements more effective?

The analyses of possible enablers and barriers show that in most of the studies, the contractual arrangements that were offered by the firms tended to be agreed with farmers that were relatively asset-richer. In 61 percent of the studies, the contracted farms were significantly larger or farmers were richer than non-participating farmers. Only in 11 percent of the cases (3 out of 26), did participation in the contract tend to involve relatively less-endowed farmers, and, two of these instances showed relatively low income effects compared to the average in the sample. This might suggest that the better-off farmers, having multiple market alternatives and higher risk-bearing capacities, opt out of contracts, because the relatively low income effects in these contractual arrangements may not compensate for the costs implied. For example, better-off farmers may have more autonomy obtaining (unsecured) credit from a bank or using personal savings than accessing the credit provided under the contractual arrangement.

We found suggestive evidence on combinations of conditions that may predict relatively high income effects for farmers from the contractual arrangement that is offered to them. For perennial crops, the presence of a price premium was consistently present in all empirical instances with a relatively high effectiveness (RR>1.50). In the high effectiveness cases, the firms did not offer credit as part of the service package, but only provided a sales point where farmers could buy key inputs with cash. For annual crops, a price premium also seems to be a necessary part of the service package for the arrangement to become highly effective in raising income, especially when no cooperative is involved. Where no price premium exists, and a competitive price is paid on local markets, the intermediary role of farmer organisations may become more important for enabling higher income effects of the contract farming arrangement. For animal husbandry, there is suggestive evidence that the package of 'inputs plus credit' is an enabling factor for higher effectiveness. All highly effective cases offered such a package.

5.2 Overall completeness and applicability of evidence

The contractual arrangements covered by this meta-analysis varied very much in terms of crops grown, embedded services and the role of intermediaries, such as cooperatives, state agencies, and NGOs. Therefore, pooled meta-analyses of these studies resulted in inferences with important threats to construct validity (Shadish, Cook and Campbell 2002). Even though we used a quite restrictive definition of contract farming, the diversity was still huge, not only because the type of production under contract differed, but also due to differences in the socio-economic and agro-ecological context of each country and region in which the study was done. We demonstrated that there were important differences in the service package and incentive structure offered to farmers. It often proved difficult to extract sufficient detailed information about these incentives and services from the effectiveness studies alone. Studies would benefit from a better specification and description of what the contract farming arrangements comprise in practice.

We are aware that the inferences derived from this review have a constrained generalisation domain. The mandatory screening of studies on the risk of bias to their net-effect estimates, resulted in a limited number of studies, which do not cover the whole range of sectors in which contract farming is an important modality of procurement (e.g. sugar, dairy, barley, banana, asparagus, fresh fruits). Inevitably, meta-analyses on a small but widely heterogeneous sample of empirical instances of contract farming will have important construct validity threats. Therefore, in the forest plots and tables, we also provide information about the type of production and location of each empirical instance, in order to facilitate interpretation and reduce the chances of misinterpretation of our findings.

5.3 Quality of the evidence

We noticed that the academic community is highly interested in studying the effectiveness of contract farming. We found 22 studies - containing 26 empirical instances - with a strong counterfactual design. All studies selected for the meta-analysis

were observational, cross-sectional studies that used advanced econometric methods to resolve he challenges related to selection and self-selection of farmers into contracts.

All experimental and quasi-experimental study designs are vulnerable to the fact that they assess net-effects in specific (months of the) year(s). Most panel studies (RCT, Diff-in-diff) use two measurements to derive inference of impact. The years used for baseline or endline may, however, refer to ‘unusual’ situations (e.g. weather, markets). These sources of variability are especially relevant when the contract farming arrangement covers news crops or livestock that are different from the traditional (counterfactual) activities. If conditions differ between years and affect these crops differently (e.g. terms of trade, drought resistance, etc.) a comparison between baseline and endline, and contracted farmer versus non-contracted farmer, will result in net-effect estimates with limited predictive value.

Because the content of the ‘treatment’, targeted geographical area, and the incentives and services embedded in the contractual arrangement may vary and evolve over time, there are operational challenges to clean baseline surveys. Cross-sectional study designs with econometric controls for selection bias are arguably more robust to these changes and may result in better estimates of net-effects. Notwithstanding this operational robustness, one-off cross-sectional surveys (with no baseline) only take a snapshot and therefore suffer from survivor bias. The repetition of the survey in the same empirical instance in a longitudinal panel design with multiple follow-up measurements, as advised by McKenzie (2012), would increase the robustness of the inferences, and help to shed light on the step-in and drop-out dynamics of contract farming.

We found no randomised control trials or longitudinal studies with a quasi-experimental design. The absence of RCT impact studies is not surprising, as the nature of contract farming implies deliberate self-selection by farmers and deliberate targeting by firms. Theoretically, an RCT with an encouragement design – randomly providing an additional stimulus to convince farmers to participate in the contract, and assessing the difference between those that do and those that do not respond to this stimulus - might work to assess the net-effects of treatments characterised by self-selection. In such a design, it might be feasible to experiment with different (intensity of) service packages, or contractual modalities (Saenger, Torero and Qaim 2013).

5.4 Limitations and potential biases in the review process

We are aware that the screening process to select the studies for meta-analysis resulted in a sample of empirical instances of contract farming that is not necessarily representative of the whole population of empirical instances. The systematic review yielded only one study of the effects of contract farming on food security with an econometric design to reduce selection bias. However, this precludes more generalised inferences, beyond the context of Madagascar. For income effects, we cover more empirical instances but also in this collection we observe limitations in coverage, especially of traditional crops where contract farming arrangements are common, such as sugar and cotton production.

For the analysis of enablers of effectiveness, we could have used the wider body of literature, identified during the search. Time and budget constraints, however, limited the scope of this review to only those empirical instances of contract farming covered by the effectiveness studies with a sufficiently strong econometric design to be selected for meta-analysis. However, there is more scope to learn from the wider body of literature identified in the search process. Ideally, we would have applied the analysis of enablers and barriers to a larger set of empirical instances. Therefore, we propose a follow-up explorative review of the qualitative and quantitative literature identified after title-abstract screening, in order to generate hypotheses on enablers and barriers of effectiveness of contract farming, differentiating the farmers according to types of crops and service packages which we identified. This may also allow us to find more qualitative evidence on sector innovation, drop-out dynamics and spillover effects, for which we found little evidence in the studies selected for meta-analysis but that may have been recorded by studies that we excluded.

The process of deriving the pooled average results of the meta-analysis followed the protocol (Ton et al. 2015). All studies were reviewed by at least two independent reviewers, except for the screening of relevance based on title and abstract. The reason for this single-screening was the heavy time investment needed to scan literature retrieved from electronic libraries with the broad search terms that were necessary to capture the fuzzy concept ‘contract farming’. However, as the criteria for exclusion used in title-abstract screening proved quite straightforward, we do not expect this decision to have had an influence on the inferences from the meta-analysis. Normative bias is more likely in the qualitative comparative analysis, which explored for plausible enablers and barriers of effectiveness. We focused on the incentives and services embedded in the contracts. As with realist synthesis, other researchers are likely to put other emphases, which could have resulted in other inferences (Wong, Greenhalgh, Westhorp et al. 2013).

Finally, the search in October 2015 might not have captured all studies that were published in the months immediately preceding October 2015 due to the time lag between the date of publication as a working paper or academic article and the date that the reference is included in the electronic libraries.

5.5 Agreements and disagreements with other studies or reviews

The meta-analysis covered studies discussed in earlier literature reviews (Minot and Ronchi 2015, Otsuka et al. 2016, Oya 2012, Prowse 2012, Wang et al. 2014b). Three of these reviews had similar objectives to our review, namely to derive inferences about effects on income, selecting only econometric studies (Minot and Ronchi 2015, Otsuka et al. 2016, Wang et al. 2014b). Two other reviews were less restrictive and reviewed a broader body of evidence (Oya 2012, Prowse 2012).

Our conclusions on the relatively high average effect-size corroborate those of Minot and Ronchi (2015) who mention an overall increase in income “between 25 and 75 percent”. Our findings also coincide on the importance for firms to offer an attractive price premium to compensate for the transaction costs involved in contract farming (Minot and Ronchi 2015:2).We disagree with Minot and Ronchi (2015:4) when they state that “most studies detect no significant difference in farm size between contracted farmers and other farms

in a given region, a finding that points to a role for contract farming in inclusive growth and poverty." Instead, we found strong evidence of bias towards the participation of relatively larger farms or richer farmers in areas where contract farming takes place. We agree that contract farming has a role in obtaining a more inclusive growth because it creates possibilities for upward social mobility of some smallholders and may be catalytic for technological upgrading and innovation in agriculture. It is, however, not the panacea for the poorest portion of smallholder farmers.

In our conclusion that participation in contract farming tends to be biased towards the better-off farmers, we coincide with Wang et al. (2014b), Prowse (2012:85), and Otsuka, Nakano and Takahashi (2016). These reviews found that although most contract farming schemes target and include small farms, most included farms are larger.

We coincide with most authors that the transaction costs embedded in contract farming need to be outweighed by the benefits, both for firms and farmer. Wang et al. (2014b:1260) stress the transaction costs savings for the processor, while we stress the benefits of transaction costs for farmers involved in contracting. We suggest that there is a need for a relatively high level of effectiveness of the contractual arrangement to compensate for the loss of autonomy in production and marketing.

Our review covers primarily empirical instances where the farmer had the option to step in and out of the contract (Narayanan 2013), whereas other scholars emphasise the empirical instances in which contracts are 'imposed' on farmers without these options. For example, Wendimu et al. (2016:84) argue that "... it is not surprising that most existing studies find a positive effect of (private-led voluntary) contract farming on the participating households' income. In contrast, it is questionable whether out-grower schemes with compulsory participation also provide benefits for farmers because the farmers are forced to participate regardless of whether the participation is beneficial for them or not." We agree with this observation.

We consider cooperatives to be important mediators of effectiveness, especially in sectors where there is no clear price premium compared to the local market. Our findings agree with those of Prowse (2012), Wang et al. (2014b), Minot and Ronchi (2015), and Otsuka et al. (2016). Oya (2012) does not mention the role of previous collective marketing experience, though he does emphasise the importance of collective action for the negotiation capacities of farmers when contracting with the firm.

6. Authors' conclusions

6.1 Implications for practice and policy

The publication and survivor bias detected in the studies prevents us from drawing strong conclusions on the effectiveness of contract farming arrangements. The effect sizes reported in published academic literature is prone to several biases, hence they are likely to paint too favourable a picture of contract farming as a way to increase smallholder incomes. The pooled average effects that result from the meta-analysis refers to empirical instances of contract farming that have existed for several years (which points to survivor bias), and where the income effects for smallholder farmers are large enough to show up even in relatively small samples, (which points to publication bias).

All studies covered in the meta-analysis related to empirical instances of contract farming that had managed to survive the initial years. The professional literature indicates that there are many factors (e.g. lack of trust between firm and farmers, fragility of market access for the firm, low knowledge and skills on the new crop/livestock on the part of farmers) that result in a high likelihood of failure and subsequent abandonment of contract farming as a modality for a firm to source products from smallholders.

Only one study assessed the effects on food security (Bellemare and Novak 2016). They measure the ‘length of the hungry season’, and report the reduction in duration of 8 percent, and compute ‘the likelihood that a household's hungry season will end at any given time’ at 18 percent larger for contract farmers than for non-contract farmers. The lack of studies shows that more research is needed on food security effects.

The average increase in income for contract farmers is 62 percent due to the contractual arrangement, over the income of non-contract farmers. However, when we used a method which aims to control for publication bias, we estimate income among contract farmers is only 38 percent bigger than income for non-contract farmers. This is still a large effect of contract farming and suggests that the income effects of participating in the contractual arrangement need to be relatively high in order to be attractive and to compensate for the transaction costs and loss of autonomy and to prevent farmers from stepping out.

The services offered by a firm in exchange for the pre-planting sales agreement imply opportunities and risks for smallholders. Contract farming is especially attractive for those farmers who can bear more risk and investments. The qualitative synthesis in this review suggests that contract farming is biased towards the farmers that are wealthier in terms of land or other assets. Only 3 of the 22 empirical instances showed that contracted farms were statistically significantly smaller or farmers poorer than the comparison group, whereas in 16 out of 22 cases the contract farms were significantly larger or farmers wealthier. Plausible explanations for this phenomenon -frequently mentioned in the studies - are the lower transaction costs for the firm when procuring the needed volume from larger farmers, and the capacity of the farmer to bear the associated production or post-harvest quality risks.

We found some suggestive evidence on conditions that could be used by practitioners and policy makers to assist with enabling effectiveness. These differ per crop type. In perennial crops, a price premium was consistently present in all empirical instances with a relatively high effectiveness. It seemed to be preferable for firms not to offer credit as part of their services to farmers, but have farmers pay cash for the inputs. In the highly effective cases covered by the studies, the firm served as an access point for these key inputs (compost, fertiliser, pesticides), providing timely access and proper quality, but did not provide credit. In annual crops, a price premium seemed to be a necessary component of the service package in order to result in high income effects for farmers, especially in situations where there was no cooperative involved as an intermediary between the firm and the farmers. If the local market pays competitive prices, the intermediary role of farmer organisations at the moment of starting the firm-farm relationship may work as an enabler of effectiveness. In cases in which these organisations were not involved, contractual arrangements that provided a higher price than available on the spot market were consistently more effective. In animal husbandry,

there was suggestive evidence that providing the package of ‘inputs plus credit’ is an enabling factor for higher effectiveness, especially in intensive animal husbandry, such as broiler production.

We showed that all studies covered farmers in and around the geographical locations where the empirical instance of contract farming was operating. These geographical locations are not necessarily representative for the country as a whole. The contractual arrangements were most likely set up in regions which benefitted from external factors such as relatively good access to infrastructure or nearby processing facilities. Hence, one cannot conclude that contract farming would have similar effects in every region of these countries. National surveys or agricultural censuses should include questions on the service packages available to farmers as well as the source of funding for this service delivery, in order to get a better idea of the importance of contract farming in agriculture.

6.2 Implications for research

Contract farming is a container concept that covers a wide range of contractual arrangements, which makes it difficult to draw overly general conclusions. This does not necessarily change if more studies become available. Nevertheless, new research to assess the income effects in specific instances of contract farming may be relevant, especially when it assesses the effects of various well-specified service packages. Our findings lead to some recommendations for future research on contract farming.

First, the reporting about the intervention can be improved. Because many contract farming arrangements imply the adoption of new inputs, new crops, and new ways of horizontal coordination (farmer groups), the effect of the ‘contract farming’ could likewise be framed as the result of ‘new inputs’, ‘access to credit’ or ‘collective marketing’, without changing the empirical analysis, but making the fact that a contract is involved an irrelevant detail. A better specification of the services provided by the firm to the farmer, and (quasi) experimental research with varying services packages in each empirical instance, would help to disentangle the effects and help to better identify the drivers and mechanisms of effectiveness in each contractual arrangement. In most studies selected in this review, the emphasis in the texts was on the econometric methods used to derive income effect estimates, whereas the description of the services, incentives and contractual clauses of these arrangements lacked sufficient detail, and there was no data on the drop-out or step-in dynamics of farmers. Although econometric rigour is valuable, we noticed that the literature often focused more on the econometric methods used than the description of the services provided under the contracts.

Second, in order to obtain better insight into the enablers and barriers of effectiveness and drop-out dynamics, and to compensate for the apparent publication and survivor bias in the existing knowledge base on contract farming, new research should also document the less-successful instances of contract farming, and report inconclusive results (insignificant effects). Research should start earlier and take particular care to cover the performance and dynamics of contract farming in the first years. When a baseline study is not possible, rigour will increase using repeated measurements or longitudinal monitoring. There is ample room for making research more comparative, by applying similar analyses on multiple instances of contract farming. Together, this would

increase the relevance of meta-analysis, which is now highly influenced by the prevalent publication and survivor bias. There are important methodological advances in how net-effect estimates are derived from (cross-sectional) survey data. These include new survey questions to get appropriate instrumental variables to control for participation bias in net-effect estimates, as well as switching regressions that model farm outcomes according to crop-specific production functions instead of estimating these with an overall production function. New studies need to build on these methodological advances. Nevertheless, we show that most studies still face a risk of bias due to imperfect group identification and weak instruments, especially because they relied on only one measurement in time.

Third, in addition to income effects, other outcomes of contract farming are important to assess. Apart from food security effects, the role of contract farming in rural development, such as (sector-wide) innovation, and livelihood resilience, will need more research. Ideally, new studies should rely on common indicators and questions that improve the comparability of findings between studies. This should be possible to attain, as most studies in this review used fairly similar outcome areas and farm characteristics. However, because authors chose slightly different methods and slightly different definitions of variables and proxy-indicators, results were not always easy to compare. National surveys or agricultural censuses should include questions on the service packages available to farmers as well as the source of funding for this service delivery, to get a better idea of the importance of contract farming in agriculture.

Appendix

Characteristics of included studies

Awotide et al. 2015

Methods	PSM
Sample	341
Context of intervention	Nigeria; rice; contracted by private firms.
Interventions	Seeds, agrochemicals, extension. No unique firm or contractual arrangement.
Outcomes	2013: Crop income, yield, poverty
Bellemare 2012	
Methods	IV
Sample	1178
Context of intervention	Madagascar; several crops (green beans, leek, snow peas, barley, rice); contracted by private firms.
Interventions	Seeds (for some), agrochemicals, extension No unique firm or contractual arrangement. Most of the sample, however, concerns one firm (Lecofruit) that has operated in Madagascar since the early 1990s
Outcomes	2008 - Household income
Bellemare and Novak 2016	
Methods	IV
Sample	1178
Context of intervention	Madagascar; several crops (green beans, leek, snow peas, barley, rice); contracted by private firms.
Interventions	Seeds (for some), agrochemicals, extension No unique firm or contractual arrangement. Most of the sample, however, concerns one firm (Lecofruit) that has operated in Madagascar since the early 1990s
Outcomes	2008 – Duration (months) of the hungry season(s).
Bolwig et al. 2009	
Methods	Heckman
Sample	160
Context of intervention	Uganda; coffee; contracted by private firm
Interventions	Seeds (on a limited scale), agrochemicals (on a limited scale),

	extension Single firm (Kawacom) started certification in 2000
Outcomes	2005 - Crop income, yield, practice adoption
Cahyadi and Waibel 2011	
Methods	Heckman
Sample	245
Context of intervention	Indonesia; oil palm; contracted by private firm
Interventions	Seeds, extension, other services unclear One oil palm company with 15,441 has. The contracts were offered in two periods, 1989-94 and 1995-2000.
Outcomes	2010 - Household income
Escobal and Cavero 2012	
Methods	Switching Regression
Sample	360
Context of intervention	Peru; potatoes; contracted by private firm in cooperation with NGO
Interventions	Extension, other services unclear Dominated by one firm, facilitated by the GO FOVIDA from 2000 onwards.
Outcomes	2002/2003 - Crop income
Girma and Gardebroek 2015	
Methods	IV
Sample	195
Context of intervention	Ethiopia; honey; contracted by private firm
Interventions	Credit, extension, inspections, honey containers, and other necessary inputs. Agreements between producers and processors in the region started in 2005. The firm (Beza Mar Agro) started contracting in 2007.
Outcomes	2010/2011 - Crop income, price margin, labor productivity
Ito et al. 2012	
Methods	PSM
Sample	318
Context of intervention	China; watermelon; contracted by farmers' cooperative

Interventions	Seeds, agrochemicals, extension, subsidized land improvement, wholesale market. An executive of a cooperative established the firm in 2000 with 170 farmers. By 2009, 2,300 households grew watermelon.
Outcomes	2009 - Farming income
Jones and Gibbon 2011	
Methods	IV
Sample	222
Context of intervention	Uganda; cocoa; contracted by private firm. Important changes between 2005 and 2009 in the contractual arrangements.
Interventions	Seeds and agrochemicals (only for some and briefly at the beginning of the program), extension, initiation of savings society Set up with SIDA support in 2001/2002. I
Outcomes	2005 & 2009 - Crop income, yield, price
Maertens and Swinnen 2009	
Methods	PSM
Sample	217
Context of intervention	Senegal; French beans; contracted by private firms.
Interventions	Credit (for some), seeds, agrochemicals, extension, and land preparation, coordination and financing of planting and harvesting (on demand) Survey in the main horticultural zone that asked for contract with an agro-exporting company. This practice is common with FFV exporters organized in ONAPES since 1999.
Outcomes	2005 - Household income
Miyata et al. 2009	
Methods	Heckman
Sample	162
Context of intervention	China; apples and green onions; contracted by private firm; for super market channels vs. local markets
Interventions	Green onions: Seeds, agrochemicals (for most), extension, spraying services (sometimes) Apples: Agrochemicals (for most), extension, spraying services (sometimes) Two companies for each crop are investigated, with no detail on the time they were already active in the area.
Outcomes	2005 - Household income

Narayanan 2014	
Methods	Switching regression
Sample	262-289
Context of intervention	India; marigold, papaya, broiler and gherkins
Interventions	Marigold: Seeds, extension, other services unclear. Single firm. Papaya: Seeds, extension, other services unclear, organization and training of hired labor for latex extraction. Single firm. Broiler: Breeds, agrochemicals, extension, other services unclear. Multiple firms, one focus firm. Gherkins: Seeds, extension, other services unclear. Multiple firms, one focus firm. Companies are not explicitly mentioned, nor the time they were already active in the area.
Outcomes	2009/2010 - Crop income
Ramaswami 2009	
Methods	IV
Sample	50
Context of intervention	India; broiler
Interventions	Production management'' contract in which the integrator supplies inputs and extension, advances credit (in kind), provides price insurance and monitors grower effort through frequent inspection. Twenty growers were associated with Venkateshwara Hatcheries and remainder with two other firms. No details on the time that the companies already engaged in contract farming.
Outcomes	2002/2003 - Crop income
Rao and Qaim 2011	
Methods	Switching regression
Sample	402
Context of intervention	Kenya; vegetables.
Interventions	Services provided are unclear. Multiple firms (supermarkets, intermediaries) involved.
Outcomes	2008 - Household income
Saigenji 2012	
Methods	PSM
Sample	88, 90
Context of	Vietnam; tea; contracted by private firms and farmers'

intervention	cooperatives or state-owned enterprises
Interventions	Private contracts: Credit, seeds, agrochemicals, extension, input application for production is implemented by the company. Several firms offer contract, most established only a decade ago. SEO: Agrochemicals, extension, other services unclear. Established in the 1950s.
Outcomes	2007 - Household expenditure, technical efficiency
Setboonsang et al. 2008	
Methods	PSM
Sample	585
Context of intervention	Laos; rice;.
Interventions	Seeds, agrochemicals, extension Contracted by private firm through a farmers' cooperative Lao Arrowny Corporation contracts since 2002.
Outcomes	2004 - Crop income, yield , price
Simmons et al. 2005	
Methods	Heckman
Sample	300 (seed rice), 200 (broiler), 300 (seed corn – covered by Winters et al. [2005])
Context of intervention	Indonesia; seed corn, seed rice; contracted by private form through farmer groups; broiler, contracted by private firm; also contracts with state-owned businesses
Interventions	Seed corn: Credit (for land preparation), seeds, agrochemicals, extension, risk insurance (all production is accepted regardless of quality). Pioneer started to offer contracts in 1986. Seed rice: Seeds, extension, other services unclear. State-company started contracting in 1988. Broiler: Credit, breeds, agrochemicals, extension, veterinary services. Nusantara Unggasjava Mataram operates since 1998.
Outcomes	2002 - Farming income, family, non-family, female and off-farm labor use
Sokchea and Culas 2015	
Methods	Heckman
Sample	75
Context of intervention	Cambodia; rice;
Interventions	The farmers' cooperative RSSADC started in 2003 supported by GIZ.

	Credit (through organization of collective savings), extension (market information), investing members receive annual dividends and share in earnings of credit program, transportation cost, packaging
Outcomes	2010 - Farming income
Trifković 2014	
Methods	Heckman
Sample	191
Context of intervention	Vietnam; catfish; contracted by private firm
Interventions	Outgrower (17%): fingerlings, fry, feed, medicines. No details on the background of the firm. Marketing contract (83%): No services provided.
Outcomes	2010 - Household expenditure
Wainaina et al. 2014	
Methods	PSM
Sample	180
Context of intervention	Kenya; poultry.
Interventions	Services unclear. Multiple firms, not detail on time these already were active.
Outcomes	2010/2011 - Crop income
Wang et al. 2014	
Methods	Heckman
Sample	107
Context of intervention	Vietnam; 'Safe' vegetables; private firms through farmer cooperatives
Interventions	No services provided Multiple firms involved. The 'safe vegetables' program started in 1995. As of 2008, firms needed to organize VietGAP inspection/traceability.
Outcomes	2007/2008 - Farming income
Warning and Key 2002	
Methods	Heckman
Sample	26
Context of intervention	Senegal; peanuts; contracted by investor-owned firm (after initiation by the government, which still holds shares in the company)

Interventions	Seeds, agrochemicals, extension, collection NOVASEN started operations in 1990.
Outcomes	1992/1994 - Farming income
Winters et al. 2005	
Methods	Heckman
Sample	300
Context of intervention	Indonesia; seed corn; contracted by private firm and/or farmers' cooperative
Interventions	Credit (for land preparation), seeds, agrochemicals, extension Pioneer started to source hybrid seeds in 1986.
Outcomes	2002 - Farming income, input use, family and hired labor use

Characteristics of excluded studies

Anim, Raphala et al. (2008)

Reason for exclusion	Inadequate design T-tests only, no control for self-selection
Bamiro, Momoh et al. (2009)	
Reason for exclusion	No control for selection bias; not clear it is contract farming; mentions non-integrated vs. partially integrated vs. fully integrated
Barrett, Bachke et al. (2010)	
Reason for exclusion	It is about contract farming effectiveness, in the broad sense of the concept. It summarizes several studies, and provides insufficient data and information on methodologies
Barrett, Bachke et al. (2012)	
Reason for exclusion	Summary of five other studies
Begum and Alam (2005)	
Reason for exclusion	Inadequate counterfactual design, descriptions only
Begum, Osanami et al. (2005)	
Reason for exclusion	Inadequate counterfactual design, simple comparison of net return and other indicators, no self-selection control
Begum (2005)	
Reason for exclusion	Effectiveness study of contract farming; inadequate counterfactual design, simple comparison of mean profitability, no control for self-selection
Begum (2008)	
Reason for exclusion	Similar paper as 13566135. No counterfactual design, simple comparisons only, no control for self-selection

Begum, Alam et al. (2012)

Reason for exclusion	Inadequate counterfactual design, efficiency frontier analysis without correction for self-selection bias

Berdegué, Reardon et al. (2007)

Reason for exclusion	No counterfactual design, summary of other studies, simple comparisons, no control for self-selection

Birthal, Joshi et al. (2005)

Reason for exclusion	Quantitative effectiveness study without adequate counterfactual design; focus on participation 'choice'

Birthal et al. (2009)

Reason for exclusion	Not possible to calculate effect size

Boulay, Tacconi et al. (2013)

Reason for exclusion	No control for self-selection bias

Briones (2014)

Reason for exclusion	Not possible to calculate effect size reported in study

Cai and Han (2011)

Reason for exclusion	Regression analysis on many variables, including marketing and production contracts, ceteris paribus effects on use of organic fertilizer and off-farm income

Cai et al. (2008)

Reason for exclusion	Effect sizes could not be calculated due to missing information

Chengappa, Nagaraj et al. (2012)

Reason for exclusion	Inadequate counterfactual methodology, simple comparisons, no control for self-selection

Boulay (2013)

Reason for exclusion	Comparison of contract with non-contract growers, but no control for self-selection

Costales, Delgado et al. (2007)

Reason for exclusion	No counterfactual design; comparative study of different marketing options

Da Silva and Rankin (2013)

Reason for exclusion	Excluded from core set because no control groups (only detailed discussion of several contract schemes)

Dedehouanou, Swinnen et al. (2013)

Reason for exclusion	Only study with happiness as dependent variable; no reference group for meta-analysis

Freguin-Gresh et al. (2012)

Reason for exclusion	Very weak instruments; average farm size of 50 hectares

Gibbon and Bolwig (2007)	
Reason for exclusion	Means and correlation analysis
Herck, Noev et al. (Herck, Noev and Swinnen 2012)	
Reason for exclusion	No control for self-selection bias on contracts. Only study with growth rates as a dependent variables (and discussion where this is smallholder farming)
Indarsih (2012)	
Reason for exclusion	Inadequate counterfactual method; very small sample, no control for self-selection, simple comparison
Jabbar and Akter (2006)	
Reason for exclusion	Not really a quantitative effectiveness study, although contracts are mentioned. No control for self-selection
Kennedy and Oniang'o (1990)	
Reason for exclusion	Inadequate counterfactual design, mainly qualitative, simple comparisons, no control for self-selection
Khamphone and Sato (2011)	
Reason for exclusion	Inadequate counterfactual control, no control group, only comparisons over time
Kongchheng (2010)	
Reason for exclusion	Effectiveness study of contract farming; inadequate counterfactual method, only regression of only contracted farmers and a Probit model
Kumar (Kumar 2006)	
Reason for exclusion	Inadequate methods, only OLS, no control for self-selection bias
Lee, Ghazoul et al. (2014)	
Reason for exclusion	No control for selection bias
Mabila (Mabila 2006)	
Reason for exclusion	Comparative data envelopment analysis, no counterfactual design
Maertens, Dries et al. (2007)	
Reason for exclusion	Effectiveness study of contract farming, but stays descriptive
McCulloch and Ota (2002)	
Reason for exclusion	No control for self-selection
Michelson (Michelson 2013)	
Reason for exclusion	Only study with assets as dependent variable; no reference group for meta-analysis
Michelson, Reardon et al. (2012)	
Reason for exclusion	Comparison, mainly on prices, between traditional and supermarket channels. No effectiveness studies controlling

	for self-selection
Munongo (2012)	
Reason for exclusion	Insufficient information provided: unreliable?
Nagaraj, Chandrakanth Mysore et al. (2008)	
Reason for exclusion	No control for selection bias
Patrick (2004)	
Reason for exclusion	No adequate control for self-selection
Rana, Neeraj et al. (2014)	
Reason for exclusion	Inadequate counterfactual method, contract farming not the focus of the study, few simple comparisons without correction of self-selection bias
Roy and Thorat (2008)	
Reason for exclusion	Not possible to calculate effect size
Saenz and Ruben (2004)	
Reason for exclusion	No counterfactual design, not primarily impact study but focused on participation choice
Sáenz-Segura D'Haese et al. (2009)	
Reason for exclusion	No counterfactual design, comparison of different marketing options without adequate control for self-selection
Sänger (Sänger 2012)	
Reason for exclusion	Same as publications with Torero; no control group of independent farmers
Schipmann and Qaim (2011)	
Reason for exclusion	No counterfactual design (choice experiment on contracts)
Setboonsarng, Leung et al. (2006)	
Reason for exclusion	No control for self-selection
Setboonsarng, Leung et al. (2006)	
Reason for exclusion	Authors mention control for unobservables, but no first-stage results shown. Only results of simulation models are reported
Sharma (2008)	
Reason for exclusion	No counterfactual design (Heckman model, but without IV)
Tadesse and Guttormsen (2009)	
Reason for exclusion	Only study with marketed surplus as dependent variable; no reference group for meta-analysis
Tatlidil and Akturk (2004)	
Reason for exclusion	No control for selection bias: comparative analysis (as per the title)
Tongchure and Hoang (2013)	
Reason for exclusion	Means comparison; regression on participation choice;

	verbal contracts
Tripathi, Singh et al. (2005)	
Reason for exclusion	No control for selection bias; only linear regression

Search strategy

Elaborated with the kind assistance of John Eyers (3ie)

Scopus search – Searched 30th September 2015

((((TITLE-ABS-KEY ((afghanistan OR albania OR algeria OR angola OR argentina OR armenia OR armenian OR aruba OR azerbaijan OR bangladesh OR benin OR byelarus OR byelorussian OR belarus OR belorussian OR belorussia OR belize OR bhutan OR bolivia OR bosnia OR herzegovina OR hercegovina OR botswana OR brasil OR brazil OR bulgaria OR "Burkina Faso" OR "Burkina Fasso" OR "Upper Volta" OR burundi OR urundi OR cambodia OR "Khmer Republic" OR kampuchea OR cameroon OR cameroons OR cameron OR camerons OR "Cape Verde" OR "Central African Republic" OR chad OR china OR colombia OR comoros OR "Comoro Islands" OR comores OR mayotte OR congo OR zaire OR "Costa Rica*" OR "Cote d'Ivoire" OR "Ivory Coast" OR cuba OR djibouti OR "French Somaliland" OR dominica OR "Dominican Republic" OR "East Timor" OR "East Timur" OR "Timor Leste" OR ecuador OR egypt OR "United Arab Republic" OR "El Salvador" OR eritrea OR ethiopia OR fiji OR gabon OR "Gabonese Republic" OR gambia OR gaza OR "Georgia Republic" OR "Georgian Republic" OR ghana OR grenada OR guatemala OR guinea OR guiana OR guyana OR haiti OR hungary OR honduras OR india OR maldives OR indonesia OR iran OR iraq OR jamaica OR jordan OR kazakhstan OR kazakh OR kenya OR kiribati OR korea OR kosovo OR kyrgyzstan OR kirghizia OR "Kyrgyz Republic" OR kirghiz OR kirgizstan OR "Lao PDR" OR laos OR lebanon OR lesotho OR basutoland OR liberia OR libya OR macedonia OR madagascar OR "Malagasy Republic" OR malaysia OR malaya OR malay OR sabah OR sarawak OR malawi OR mali OR "Marshall Islands" OR mauritania OR mauritius OR "Agalega Islands" OR mexico OR micronesia OR "Middle East" OR moldova OR moldovia OR moldovian OR mongolia OR montenegro OR morocco OR ifni OR mozambique OR myanmar OR myanma OR burma OR namibia OR nepal OR "Netherlands Antilles" OR "New Caledonia" OR nicaragua OR niger OR nigeria OR pakistan OR palau OR palestine OR panama OR paraguay OR peru OR philippines OR philipines OR phillipines OR phillippines OR "Puerto Ric*" OR romania OR rumania OR roumania OR rwanda OR ruanda OR "Saint Lucia" OR "St Lucia" OR "Saint Vincent" OR "St Vincent" OR grenadines OR samoa OR "Samoan Islands" OR "Navigator Island" OR "Navigator Islands" OR "Sao Tome" OR senegal OR serbia OR montenegro OR seychelles OR "Sierra Leone" OR "Sri Lanka" OR "Solomon Islands" OR somalia OR "South Africa" OR sudan OR suriname OR surinam OR swaziland OR syria OR tajikistan OR tadzhikistan OR tadjikistan OR tadzhik OR tanzania OR thailand OR togo OR togolese republic OR tonga OR tunisia OR turkey OR turkmenistan OR turkmen OR uganda OR ukraine OR uzbekistan OR uzbek OR vanuatu OR "New Hebrides" OR venezuela OR vietnam OR "Viet Nam" OR "West Bank" OR yemen OR yugoslavia OR zambia OR zimbabwe))) OR (TITLE-ABS-KEY ("Developing Countries" OR africa OR asia OR caribbean OR "West Indies" OR "South America" OR

"Latin America" OR "Central America" OR ((developing OR "less* developed" OR "under developed" OR underdeveloped OR "middle income" OR "low* income" OR underserved OR "under served" OR deprived OR poor*) W/1 (countr* OR nation* OR population* OR world)))) OR (TITLE-ABS-KEY (((developing OR "less* developed" OR "under developed" OR underdeveloped OR "middle income" OR "low* income") W/1 (economy OR economies)) OR (low* W/1 (gdp OR gnp OR "gross domestic" OR "gross national")) OR (low W/3 middle W/3 countr*))) OR (TITLE-ABS-KEY (((lmic OR lmics OR "third world" OR "lami countr*")) OR "transitional countr*"))) AND (TITLE-ABS-KEY ((contract* W/2 farm*))))) OR ((((TITLE-ABS-KEY ((afghanistan OR albania OR algeria OR angola OR argentina OR armenia OR armenian OR aruba OR azerbaijan OR bangladesh OR benin OR byelarus OR byelorussian OR belarus OR belorussian OR belorussia OR belize OR bhutan OR bolivia OR bosnia OR herzegovina OR hercegovina OR botswana OR brasil OR brazil OR bulgaria OR "Burkina Faso" OR "Burkina Fasso" OR "Upper Volta" OR burundi OR urundi OR cambodia OR "Khmer Republic" OR kampuchea OR cameroon OR cameroons OR cameron OR camerons OR "Cape Verde" OR "Central African Republic" OR chad OR china OR colombia OR comoros OR "Comoro Islands" OR comores OR mayotte OR congo OR zaire OR "Costa Rica*" OR "Cote d'Ivoire" OR "Ivory Coast" OR cuba OR djibouti OR "French Somaliland" OR dominica OR "Dominican Republic" OR "East Timor" OR "East Timur" OR "Timor Leste" OR ecuador OR egypt OR "United Arab Republic" OR "El Salvador" OR eritrea OR ethiopia OR fiji OR gabon OR "Gabonese Republic" OR gambia OR gaza OR "Georgia Republic" OR "Georgian Republic" OR ghana OR grenada OR guatemala OR guinea OR guiana OR guyana OR haiti OR hungary OR honduras OR india OR maldives OR indonesia OR iran OR iraq OR jamaica OR jordan OR kazakhstan OR kazakh OR kenya OR kiribati OR korea OR kosovo OR kyrgyzstan OR kirghizia OR "Kyrgyz Republic" OR kirghiz OR kirgizstan OR "Lao PDR" OR laos OR lebanon OR lesotho OR basutoland OR liberia OR libya OR macedonia OR madagascar OR "Malagasy Republic" OR malaysia OR malaya OR malay OR sabah OR sarawak OR malawi OR mali OR "Marshall Islands" OR mauritania OR mauritius OR "Agalega Islands" OR mexico OR micronesia OR "Middle East" OR moldova OR moldovia OR moldovian OR mongolia OR montenegro OR morocco OR ifni OR mozambique OR myanmar OR myanma OR burma OR namibia OR nepal OR "Netherlands Antilles" OR "New Caledonia" OR nicaragua OR niger OR nigeria OR pakistan OR palau OR palestine OR panama OR paraguay OR peru OR philippines OR philipines OR phillipines OR phillippines OR "Puerto Ric*" OR romania OR rumania OR roumania OR rwanda OR ruanda OR "Saint Lucia" OR "St Lucia" OR "Saint Vincent" OR "St Vincent" OR grenadines OR samoa OR "Samoan Islands" OR "Navigator Island" OR "Navigator Islands" OR "Sao Tome" OR senegal OR serbia OR montenegro OR seychelles OR "Sierra Leone" OR "Sri Lanka" OR "Solomon Islands" OR somalia OR "South Africa" OR sudan OR suriname OR surinam OR swaziland OR syria OR tajikistan OR tadzhikistan OR tadjikistan OR tadzhik OR tanzania OR thailand OR togo OR togolese republic OR tonga OR tunisia OR turkey OR turkmenistan OR turkmen OR uganda OR ukraine OR uzbekistan OR uzbek OR vanuatu OR "New Hebrides" OR venezuela OR vietnam OR "Viet Nam" OR "West Bank" OR yemen OR yugoslavia OR zambia OR zimbabwe))) OR (TITLE-ABS-KEY ("Developing Countries" OR africa OR asia OR caribbean OR "West Indies" OR "South America" OR "Latin America" OR "Central America" OR ((developing OR "less* developed" OR "under developed" OR underdeveloped OR "middle income" OR "low* income" OR underserved OR "under served" OR deprived OR

poor*) W/1 (countr* OR nation* OR population* OR world)))) OR (TITLE-ABS-KEY (((developing OR "less* developed" OR "under developed" OR underdeveloped OR "middle income" OR "low* income") W/1 (economy OR economies)) OR (low* W/1 (gdp OR gnp OR "gross domestic" OR "gross national")) OR (low W/3 middle W/3 countr*))) OR (TITLE-ABS-KEY (((lmic OR lmics OR "third world" OR "lami countr*")) OR "transitional countr*"))) AND ((TITLE ("Food security" OR poverty OR "household* income*") OR ABS ("Food security" OR poverty OR "household* income*") OR TITLE ((increas* OR improv* OR lower* OR decreas* OR diminish* OR reduc* OR loss OR declin* OR slump OR dwindl* OR curtail* OR restrict* OR shrink* OR fall) W/3 (income* OR revenue* OR yield* OR productivity)) OR ABS ((increas* OR improv* OR lower* OR decreas* OR diminish* OR reduc* OR loss OR declin* OR slump OR dwindl* OR curtail* OR restrict* OR shrink* OR fall) W/3 (income* OR revenue* OR yield* OR productivity)) OR TITLE ("market power" OR net-return* OR "net return*" OR outcome* OR effect* OR impact) OR ABS ("market power" OR net-return* OR "net return*" OR outcome* OR effect* OR impact)) AND (TITLE (contract* OR "nucleus estate*" OR cooperative* OR "producer* association*") OR ABS (contract* OR "nucleus estate*" OR cooperative* OR "producer* association*") OR TITLE (embedded W/3 service*) OR ABS (embedded W/3 service*) OR TITLE ((pre-harvest) W/2 (agreement* OR sales)) OR ABS ((pre-harvest) W/2 (agreement* OR sales)) OR TITLE ("value chain*" OR farm-firm* OR outgrow*) OR ABS ("value chain*" OR farm-firm* OR outgrow*) OR TITLE ((vertical) W/3 (integration OR coordination OR linkage*)) OR ABS ((vertical) W/3 (integration OR coordination OR linkage*))) AND ((TITLE (farm* OR smallhold* OR "small hold*" OR small-hold*) OR ABS (farm* OR smallhold* OR "small hold*" OR small-hold*) OR TITLE ((small-scale OR "small scale") W/3 (producer*)) OR ABS ((small-scale OR "small scale") W/3 (producer*))) OR (TITLE (agricultur* OR outgrower* OR "small farmer*" OR "small grower*") OR ABS (agricultur* OR outgrower* OR "small farmer*" OR "small grower*")) OR (TITLE ((vegetable* OR fruit OR livestock OR dairy OR milk OR beef OR poultry OR pig* OR flower* OR cereal OR tea OR soybean* OR rice OR coffee OR potato* OR sugarcane OR mushroom* OR maize OR millet OR pepper* OR crop OR crops) W/3 (produc* OR grow*)) OR ABS ((vegetable* OR fruit OR livestock OR dairy OR milk OR beef OR poultry OR pig* OR flower* OR cereal OR tea OR soybean* OR rice OR coffee OR potato* OR sugarcane OR mushroom* OR maize OR millet OR pepper* OR crop OR crops) W/3 (produc* OR grow*))) OR (TITLE (floriculture) OR ABS (floriculture)))))) **- 973 hits**

CAB Abstracts search – Searched 29th September 2015

- Database: CAB Abstracts

19 ("Food security" or poverty or "household* income*" or ((lower* or decreas* or diminish* or reduc* or loss or declin* or slump or dwindl* or curtail* or restrict* or shrink* or fall) adj3 (income* or revenue* or yield* or productivity)) or "market power" or net-return* or "net return*" or outcome* or effect* or impact).ti,ab,sh. (2095228)

21 (Contract* or "nucleus estate*" or cooperative* or "producer* association*" or (embedded adj3 service*) or (pre-harvest adj2 (agreement* or sales)) or "value chain*" or farm-firm* or outgrow* or (vertical adj3 (integration or coordination or linkage*))).ti,ab,sh. (52881)

23 (Farm* or smallhold* or "small hold*" or small-hold* or ((small-scale or "small scale") adj3 producer*) or agricultur* or outgrower* or "small farmer*" or "small grower*" or ((vegetable* or fruit or livestock or dairy or milk or beef or poultry or pig* or flower* or cereal or tea or soybean* or rice or coffee or potato* or sugarcane or mushroom* or maize or millet or pepper* or crop or crops) adj3 (produc* or grow*)) or floriculture).ti,ab,sh. (943886)

24 (Afghanistan or Angola or Albania or "American Samoa" or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Belarus or Belize or Benin or Bolivia or Bosnia or Herzegovina or Botswana or Brazil or Bulgaria or Burkina Faso or Burkina Fasso or Burundi or Urundi or Cambodia or Cameroon or Cameroons or Cameron or Camerons or Central African Republic or Chad or China or Colombia or Comoros or Comoro Islands or Comores or Congo or Costa Rica or Cuba or Zaire or Cote d'Ivoire or Ivory Coast or Djibouti or Dominica* or East Timor or East Timur or Timor Leste or Ecuador or Egypt or United Arab Republic or El Salvador or Eritrea or Ethiopia or Fiji or Gabon or Gambia or Gaza or Georgia Republic or Georgian Republic or Ghana or Grenada or Guatemala or Guinea or Guiana or Guyana or Haiti or Honduras or Hungary or India or Indonesia or Iran or Iraq or Kazakhstan or Kenya or Kiribati or Korea or Kosovo or Kyrgyzstan or Kirghizia or Kyrgyz Republic or Kirghiz or Kirgizstan or Lao PDR or Laos or Lebanon or Lesotho or Liberia or Libya or Macedonia or Madagascar or Malagasy Republic or Malawi or Malaysia or Maldives or Marshall Islands or Mali or Mauritania or Mauritius or Agalega Islands or Mexico or Micronesia or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Morocco or Ifni or Mozambique or Myanmar or Myanma or Burma or Namibia or Nepal or Nicaragua or Niger or Nigeria or Pakistan or Palau or Palestine or Panama or Paraguay or Peru or Philippines or Philipines or Phillipines or Phillippines or Romania or Rwanda or Ruanda or Samoa or Samoan Islands or Sao Tome or Senegal or Serbia or Seychelles or Sierra Leone or Sri Lanka or Solomon Islands or Somalia or South Africa or St Lucia or St Vincent or Grenadines or Sudan or Suriname or Swaziland or Syria or Tajikistan or Tadzhikistan or Tadjikistan or Tadzhik or Tanzania or Thailand or Tonga or Togo or Togolese Republic or Tunisia or Turkey or Turkmenistan or Tuvalu or Uganda or Ukraine or Uzbekistan or Uzbek or Vanuatu or Venezuela or New Hebrides or Vietnam or Viet Nam or West Bank or Yemen or Zambia or Zimbabwe).hw,ti,ab,cp. (2004287)

25 ((developing or less* developed or under developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor*) adj (countr* or nation? or population? or world)).ti,ab. (44493)
26 ((developing or less* developed or under developed or underdeveloped or middle income or low* income) adj (economy or economies)).ti,ab. (712)
27 (low* adj (gdp or gnp or gross domestic or gross national)).ti,ab. (43)
28 (low adj3 middle adj3 countr*).ti,ab. (1881)
29 (lmic or limits or third world or lami countr*).ti,ab. (2264)
30 transitional countr*.ti,ab. (78)
31 exp developing countries/ (1347769)
32 or/24-31 (2068603)
33 23 and 32 (9796)
34 19 and 23 and 32 (3496)
****

References

References to Included Studies

Awotide, B. A., A. Fashogbon and T. T. Awoyemi. 2015. "Impact of Agro-Industrial Development Strategies on Smallholder Rice Farmers' Productivity, Income and Poverty: The Case of Contract Farming in Nigeria."

Bellemare, Marc F. 2012. "As You Sow, So Shall You Reap: The Welfare Impacts of Contract Farming." World Development 40(7):1418-34. doi: http://dx.doi.org/10.1016/j.worlddev.2011.12.008.

Bolwig, S., P. Gibbon and S. Jones. 2009. "The Economics of Smallholder Organic Contract Farming in Tropical Africa." World Development 37:1094-104.

Cahyadi Eko, Ruddy and Hermann Waibel. 2011. "Do Smallholders Gain from Contract with an Oil Palm Company? Lessons Learned from Jambi, Indonesia." Conference on International Research on Food Security, Natural Resource Management and Rural Development.

Escobal, J. A. and D. Cavero. 2012. "Transaction Costs, Institutional Arrangements and Inequality Outcomes: Potato Marketing by Small Producers in Rural Peru." World Development 40:329-41.

Girma, J. and C. Gardebroek. 2015. "The Impact of Contracts on Organic Honey Producers' Incomes in Southwestern Ethiopia." Forest Policy and Economics 50:259-68.

Ito, J., Z. Bao and Q. Su. 2012. "Distributional Effects of Agricultural Cooperatives in China: Exclusion of Smallholders and Potential Gains on Participation." Food Policy 37(6):700-09.

Jones, S. and P. Gibbon. 2011. "Developing Agricultural Markets in Sub-Saharan Africa: Organic Cocoa in Rural Uganda." Journal of Development Studies 47(10):1595-618.

Maertens, Miet and Johan FM Swinnen. 2009. "Trade, Standards, and Poverty: Evidence from Senegal." World Development 37(1):161-78.

Miyata, S., N. Minot and D. Hu. 2009. "Impact of Contract Farming on Income: Linking Small Farmers, Packers, and Supermarkets in China." World Development 37(11):1781-90.

Narayanan, Sudha. 2014. "Profits from Participation in High Value Agriculture: Evidence of Heterogeneous Benefits in Contract Farming Schemes in Southern India." Food Policy 44:142-57.

Ramaswami, Bharat, Pratap Singh Birthal and PK Joshi. 2009. "Grower Heterogeneity and the Gains from Contract Farming: The Case of Indian Poultry." Indian Growth and Development Review 2(1):56-74.

Rao, Elizaphan JO and Matin Qaim. 2011. "Supermarkets, Farm Household Income, and Poverty: Insights from Kenya." World Development 39(5):784-96.

Saigenji, Yoshiko. 2012. "Contract Farming and Its Impact on Production Efficiency and Rural Household Income in the Vietnamese Tea Sector." Vol. Faculty of Agricultural Sciences, Institute of Agricultural Economics and Social Sciences in the Tropics and Subtropics. University of Hohenheim, Germany.

Sethboonsarng, Sununtar, PingSun Leung and Adam Stefan. 2008. "Rice Contract Farming in Lao Pdr: Moving from Subsistence to Commercial Agriculture." Vol. 90. ADB Institute Discussion Papers.

Simmons, P., P. Winters and I. Patrick. 2005. "An Analysis of Contract Farming in East Java, Bali, and Lombok, Indonesia." Agricultural Economics 33(3 supplement):513-25.

Sokchea, A. and R. J. Culas. 2015. "Impact of Contract Farming with Farmer Organizations on Farmers' Income: A Case Study of Reasmey Stung Sen Agricultural Development Cooperative in Cambodia." Australasian Agribusiness Review 23.

Trifković, N. 2014. "Governance Strategies and Welfare Effects: Vertical Integration and Contracts in the Catfish Sector in Vietnam." Journal of Development Studies 50(7):949-61.

Wainaina, Priscilla W., Julius J. Okello and Jonathan M. Nzuma. 2014. "Blessing or Evil? Contract Farming, Smallholder Poultry Production and Household Welfare in Kenya." Quarterly Journal of International Agriculture 53:319-40.

Wang, Holly, P. Moustier and N. T. T. Loc. 2014. "Economic Impact of Direct Marketing and Contracts: The Case of Safe Vegetable Chains in Northern Vietnam." Food Policy 47:13-23.

Warning, M. and N. Key. 2002. "The Social Performance and Distributional Consequences of Contract Farming: An Equilibrium Analysis of the Arachide De Bouche Program in Senegal." World Development 30(2):255-163.

Winters, P., P. Simmons and I. Patrick. 2005. "Evaluation of a Hybrid Seed Contract between Smallholders and a Multinational Company in East Java, Indonesia." Journal of Development Studies 41:62-89.

References to Excluded Studies

Anim, F. D. K., Raphala, S. B., & Mandleni, B. (2008). Contract farming: a case study of small scale sunflower farmers in Bojanala district of North West Province. Unpublished research report. Pretoria: University of South Africa.

Bamiro, O. M., Momoh, S., & Phillip, D. O. A. (2009). Vertical integration and profitability in poultry industry in Ogun and Oyo States, Nigeria. J Hum Ecol, 27, 149-154.

Barrett, C. B., Bachke, M. E., Bellemare, M. F., Michelson, H. C., Narayanan, S., & Walker, T. F. (2010). Smallholder participation in agricultural value chains: Comparative evidence from three continent. Available at SSRN 1733942.

Barrett, C. B., Bachke, M. E., Bellemare, M. F., Michelson, H. C., Narayanan, S., & Walker, T. F. (2012). Smallholder participation in contract farming: Comparative evidence from five countries. World Development, 40(4), 715-730.

Begum, Alam, & Buysse. (2012). Contract farmer and poultry farm efficiency in Bangladesh: a data envelopment analysis. Applied economics (Online), 44, 3737-3747.

Begum, I. A. (2005). Vertically integrated contract and independent poultry farming system in Bangladesh: a profitability analysis. Livestock Research for Rural Development, 17.

Begum, I. A. (2008). Prospects and potentialities of vertically integrated contract farming in Bangladesh poultry sector development. Hokkaido University, Japan.

Begum, I. A., & Alam, M. J. (2005). Contract Farming and Small Farmers: A case study of the Bangladeshi poultry sector Journal of Bangladesh Studies, 7.

Begum, I. A., Osanami, F., & Kondo, T. (2005). Performance of vertically integrated contract poultry farming and independent poultry farms in Bangladesh: a comparative study. Review of Agricultural Economics, 61, 101–117.

Bellemare, M. F., & Novak, L. (2015). Contract Farming and Food Security.

Berdegué, J., Reardon, T., Balsevich, F., Flores, L., Hernández, R., & Swinnen, J. F. M. (2007). Supermarkets and Small Horticultural Product Farmers in Central America Global Supply Chains, Standards and the Poor: CAB International.

Birthal, P. S., Jha, A. K., Tiongco, M. M., & Narrod, C. (2009). Farm-level impacts of vertical coordination of the food supply chain: Evidence from contract farming of milk in India. Indian Journal of Agricultural Economics, 64, 481-496.

Birthal, P. S., Joshi, P. K., & Gulati, A. (2005). Vertical coordination in high-value food commodities: Implications for smallholders: International Food Policy Research Institute (IFPRI).

Boulay, A. (2013). Contract farming in Thailand: provision of genetically improved eucalypt plants and extension support. Rome: Food and Agriculture Organization of the United Nations (FAO).

Boulay, A., Tacconi, L., & Kanowski, P. (2013). Financial Performance of Contract Tree Farming for Smallholders: The Case of Contract Eucalypt Tree Farming in Thailand. Small-scale Forestry, 12(2), 165-180.

Briones, R., M. (2014). Small farmers in high-value chains: Binding or relaxing constraints to inclusive growth? Policy Notes: Philippine Institute for Development Studies.

Cai, J., Ung, L., Setboonsarng, S., & Leung, P. (2008). Rice contract farming in Cambodia: Empowering farmers to move beyond the contract toward independence.

Cai, R., & Han, H. (2011). Contract production and farmer's organic fertilizer use behavior - an empirical analysis based on the survey data of apple growers in Shandong Province. Scientia Agricultura Sinica, 44, 1277-1282.

Chengappa, P. G., Nagaraj, N., Mangala, K. P., & Vijayalakshmi, D. (2012). Improving the supply chain management for vegetables through contract farming and organized

food retail chains: Indian experience. Proceedings SEAVEG.

Costales, A., Delgado, C., Catelo, M. A., Lapar, M. L., Tiongco, M., Ehui, S., & Bautista, A. Z. (2007). Scale and access issues affecting smallholder hog producers in an expanding peri-urban market: Southern Luzon, Philippines. Research Report of the International Food Policy Research Institute(151), 1-72.

Dedehouanou, S. F. A., Swinnen, J., & Maertens, M. (2013). Does contracting make farmers happy? Evidence from Senegal. Review of Income and Wealth, 59, S138-S160.

Freguin-Gresh, S., D'Haese, M., & Anseeuw, W. (2012). Demythifying contract farming: Evidence from rural South Africa. Agrekon, 51(3), 24-51.

Gibbon, P., & Bolwig, S. (2007). The economics of certified organic farming in tropical Africa: a preliminary assessment DIIS Working Paper - Danish Institute for International Studies; 2007. (2007/3):34 pp. 22 ref. Copenhagen: Danish Government Institute of Seed Pathology for Developing Countries.

Herck, K. v., Noev, N., & Swinnen, J. F. M. (2012). Institutions, exchange and firm growth: evidence from Bulgarian agriculture. European Review of Agricultural Economics, 39, 29-50.

Indarsih, B. (2012). Performance of integrated contract and independent broiler production in East Lombok, West Nusa Tenggara Journal of the Indonesian Tropical Animal Agriculture, 37(1), 50-58.

Jabbar, M. A., & Akter, S. (2006). Market and other factors affecting farm-specific production efficiency in poultry production in Vietnam. Journal of Food Products Marketing, 12, 99-113.

Kennedy, E., & Oniang'o, R. (1990). Health and nutrition effects of sugar cane production in south-western Kenya. Food and Nutrition Bulletin, 12, 261-267.

Khamphone, B., & Sato, N. (2011). Effectiveness of Rubber Plantation on Villagers' livelihood Improvement in the Northern Part of Laos. Journal of the Faculty of Agriculture, Kyushu University, 56(1), 185-191.

Kongchheng, P. (2010). Transboundary Production in Agriculture: A Case Study of Maize Contract Farming in Cambodia. Mekong Institute: Research Working Papers Series, 1.

Kumar, P. (2006). Contract farming through agribusiness firms and state corporation: a case study in Punjab. Economic and Political Weekly, 5367-5375.

Lee, J. S. H., Ghazoul, J., Obidzinski, K., & Koh, L. P. (2014). Oil palm smallholder yields and incomes constrained by harvesting practices and type of smallholder management in Indonesia. Agronomy for Sustainable Development, 34(2), 501-513.

Mabila, M. M. (2006). Comparative Analysis of Contract and Non-Contract Broiler Fanners, a Case study of the Manzini Region.

Maertens, M., Dries, L., Dedehouanou, F. A., & Swinnen, J. F. M. (2007). High-value

supply chains, food standards and rural households in Senegal (pp. 159-172 136 notes and ref). Wallingford: Global supply chains, standards and the poor CABI.

McCulloch, N., & Ota, M. (2002). Export Horticulture and Poverty in Kenya.

Michelson, H., Reardon, T., & Perez, F. (2012). Small farmers and big retail: trade-offs of supplying supermarkets in Nicaragua Staff Paper - Department of Agricultural, Food and Resource Economics, Michigan State University; 2010. (2010-02):31 pp. 21 ref. East Lansing: Michigan State University, Department of Agricultural Economics.

Michelson, H. C. (2013). Small farmers, NGOs, and a Walmart World: Welfare effects of supermarkets operating in Nicaragua. American Journal of Agricultural Economics, 95(3), 628-649.

Munongo, S. (2012). Welfare impact of private sector interventions on rural livelihoods: The case of Masvingo and Chiredzi smallholder farms. Russian Journal of Agricultural and Socio-Economic Sciences, 10(10), 3-9.

Nagaraj, N., Chandrakanth Mysore, G., Chengappa, P. G., Roopa, H. S., & Chandakavate Pramod, M. (2008). Contract farming and its implications for input-supply, linkages between markets and farmers in Karnataka. Agricultural Economics Research Review, 21, 307-316.

Patrick, I. (2004). Contract farming in Indonesia: smallholders and agribusiness working together ACIAR Technical Reports Series; 2004. (54):xvi + 72 pp. 29 ref. Canberra: Australian Centre for International Agricultural Research (ACIAR).

Rana, R. K., Neeraj, S., Arya, S., Kadian, M. S., Singh, B. P., & Pandey, S. K. (2014). Status of potato husbandry and farmer's socio-economic profile in moisture and heat prone Karnataka, India. Pakistan Journal of Agricultural Sciences, 51, 7-16.

Roy, D., & Thorat, A. (2008). Success in High Value Horticultural Export Markets for the Small Farmers: The Case of Mahagrapes in India. World Development, 36(10), 1874-1890. doi: 10.1016/j.worlddev.2007.09.009

Sáenz-Segura, F., D'Haese, M., & Speelman, S. (2009). The influence of contracts on smallholder pepper (Piper nigrum L.) producers in Costa Rica under different market conditions. Fruits, 64, 371-382.

Saenz, F., & Ruben, R. (2004). Export contracts for non-traditional products: chayote from Costa Rica. Journal on Chain and Network Science, 4, 139-150.

Sänger, C. (2012). On small farms and the design of contracts in agricultural markets – Experimental evidence from Vietnam. Göttingen.

Schipmann, C., & Qaim, M. (2011). Supply chain differentiation, contract agriculture, and farmers' marketing preferences: The case of sweet pepper in Thailand. Food Policy, 36(5), 666-676.

Setboonsarng, S., Leung, P., & Cai, J. (2006). Contract Farming and Poverty Reduction: The Case of Organic Rice Contract Farming in Thailand Poverty Strategies in Asia: A Growth Plus Approach. Weiss, John. Khan, Haider A., eds., Cheltenham, U.K. and

Northampton, Mass.: Elgar; Tokyo: Asian Development Bank Reserve. (pp. 266-299).

Sharma, V. P. (2008). India's agrarian crisis and corporate-led contract farming: Socio-economic implications for smallholder producers. International Food and Agribusiness Management Review, 11, 25-48.

Tadesse, G., & Guttormsen, A. (2009). Commercializing smallholder's through interlinked contracts: prospects and challenges in the rift valleys of Ethiopia.

Tatlidil, F. F., & Akturk, D. (2004). Comparative analysis of contract and non-contract farming model in tomato production. Journal of Agronomy, 3, 305-310.

Tongchure, S., & Hoang, N. (2013). Cassava Smallholders' Participation in Contract Farming in Nakhon Ratchasrima Province, Thailand. Journal of Social and Development Sciences, 4, 332-338.

Tripathi, R. S., Singh, R., & Singh, S. (2005). Contract Farming in Potato Production: An Alternative for Managing Risk and Uncertainty. Agricultural Economics Research Review, 18, 47-60.

Complementary Literature on empirical instances

Andersson, Camilla IM, Christine GK Chege, Elizaphan JO Rao and Matin Qaim. 2015. "Following up on Smallholder Farmers and Supermarkets in Kenya." American Journal of Agricultural Economics 97(4):1247-66.

Bellemare, Marc F. 2010. "Agricultural Extension and Imperfect Supervision in Contract Farming: Evidence from Madagascar." Agricultural Economics 41(6):507-17.

Bellemare, Marc F and Lindsey Novak. 2015. "Contract Farming and Food Security."

Bolwig, Simon. 2012. "Poverty and Gender Effects of Smallholder Organic Contract Farming in Uganda." IFPRI Uganda Strategy Support Program Working Paper 8.

Campbell, Roderick, Tristan Knowles and Amphaphone Sayasenh. 2012. "Business Models for Foreign Investment in Agriculture in Laos." The International Institute for Sustainable Development.

Guimas, Céline. 2012. "Effect of Organic Contract Farming on Labor Demand. A Study Case in the Western Uganda." Vol.

Johnson, Michael, Hiroyuki Takeshima and Kwabena Gyimah-Brempong. 2013. "Assessing the Potential and Policy Alternatives for Achieving Rice Competitiveness and Growth in Nigeria."

Lao P.D.R. Ministry of Agriculture and Forestry. 2006. "Diagnostic Study of the Agricultural and Agribusiness Sectors." Mimeo. Laos.

Ma, Wanglin and Awudu Abdulai. 2016. "Does Cooperative Membership Improve Household Welfare? Evidence from Apple Farmers in China." Food Policy 58:94-102.

Minten, Bart, Lalaina Randrianarison and Johan FM Swinnen. 2009. "Global Retail Chains and Poor Farmers: Evidence from Madagascar." World Development

37(11):1728-41.

Pérez Rozzi, Sergio. s/d. "Cadena De La Papa Andina.: Valle Chillón Mantaro. Perú." Mimeo.

Seville, Don, Abbi Buxton and Bill Vorley. 2011. "Under What Conditions Are Value Chains Effective Tools for Pro-Poor Development." Accessed 21st August], available from http://pubs. iied. org/16029IIED. html.

Trifković, N. 2016. "Comparison of Farm/Household Characteristics Using T-Tests - Table 5." Email to G. Ton. Copenhagen.

Vermeulen, Sonja and Nathalie Goad. 2006. Towards Better Practice in Smallholder Palm Oil Production: Iied.

Additional References

ActionAid. 2015. "Contract Farming and out-Grower Schemes Appropriate Development Models to Tackle Poverty and Hunger?" Vol. Johannesburg: ActionAid.

Barrett, C. B., M. E. Bachke, M. F. Bellemare, H. C. Michelson, S. Narayanan and T. F. Walker. 2010. "Smallholder Participation in Agricultural Value Chains: Comparative Evidence from Three Continent." Available at SSRN 1733942.

Borenstein, Michael, Larry V Hedges, Julian PT Higgins and Hannah R Rothstein. 2011. Introduction to Meta-Analysis: John Wiley & Sons.

Caliendo, Marco and Sabine Kopeinig. 2008. "Some Practical Guidance for the Implementation of Propensity Score Matching." Journal of economic surveys 22(1):31-72.

Campbell, Roderick, Tristan Knowles and Amphaphone Sayasenh. 2012. "Business Models for Foreign Investment in Agriculture in Laos." The International Institute for Sustainable Development.

Cartwright, Nancy. 2007. "Are Rcts the Gold Standard?". BioSocieties 2(1):11-20.

Chege, Christine GK, Camilla IM Andersson and Matin Qaim. 2015. "Impacts of Supermarkets on Farm Household Nutrition in Kenya." World Development 72:394-407.

Collier, David. 2014. "Comment: Qca Should Set Aside the Algorithms." Sociological Methodology:0081175014542568.

Da Silva, Carlos. A., and M. Rankin (2013). Contract farming for inclusive market access. Rome: FAO.

Da Silva, Carlos A and Andrew W Shepherd. 2013. "Editorial: The Multi-Faceted Dimensions of Contract Farming." Food Chain 3(3):127-30.

De Janvry, Alain, Andrew Dustan and Elisabeth Sadoulet. 2010. "Recent Advances in Impact Analysis Methods for Ex-Post Impact Assessments of Agricultural Technology: Options for the Cgiar." Unpublished working paper, University of California-Berkeley.

Deaton, Angus and Nancy Cartwright. 2016. "Understanding and Misunderstanding Randomized Controlled Trials." Vol.: National Bureau of Economic Research.

Duval, Sue and Richard Tweedie. 2000. "Trim and Fill: A Simple Funnel-Plot–Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis." biometrics 56(2):455-63.

Eaton, Charles and Andrew Shepherd. 2001. Contract Farming: Partnerships for Growth: Food & Agriculture Org. Agricultural Services Bulletin 145

Egger, Matthias, George Davey Smith, Martin Schneider and Christoph Minder. 1997. "Bias in Meta-Analysis Detected by a Simple, Graphical Test." BMJ 315(7109):629-34.

Euler, M., S. Schwarze, H. Siregar and M. Qaim. 2016. "Oil Palm Expansion among Smallholder Farmers in Sumatra, Indonesia." Journal of Agricultural Economics 67(3):658-76.

FAO. 2008, "Contract Farming Resource Centre", Rome: FAO. Retrieved 1 December, 2014.

Friese, Susanne. 2014. Qualitative Data Analysis with Atlas. Ti: Sage.

Glover, David and Ken Kusterer. 1990. Small Farmers, Big Business: Contract Farming and Rural Development: St. Martin's Press.

Harris, Ross, Mike Bradburn, Jon Deeks, Roger Harbord, Doug Altman, Thomas Steichen and Jonathan Sterne. 2010. "Metan: Stata Module for Fixed and Random Effects Meta-Analysis." Statistical Software Components.

Higgins, JPT and S Green. 2011. "Cochrane Handbook for Systematic Reviews of Interventions Version 5.1. 0 " Vol. www. cochrane-handbook. org. The Cochrane Collaboration.

Hombrados, J. and H Waddington. 2012. "Idcg Risk of Bias Tool." Vol. London.

Holt-Giménez, Eric, Justine Williams and Caitlyn Hachmyer. 2015. "The World Bank Group’s 2013-15 Agriculture for Action Plan: A Lesson in Privatization, Lack of Oversight." Vol. Oakland: Food First / Trans-National Institute.

Imbens, Guido M and Jeffrey M Wooldridge. 2008. "Recent Developments in the Econometrics of Program Evaluation." Vol.: National Bureau of Economic Research.

Ioannidis, John PA. 2005. "Why Most Published Research Findings Are False." PLoS Med 2(8):e124.

Ioannidis, John PA and Thomas A Trikalinos. 2007. "The Appropriateness of Asymmetry Tests for Publication Bias in Meta-Analyses: A Large Survey." Canadian Medical Association Journal 176(8):1091-96.

Jia, Xiangping and Jos Bijman. 2013. "Contract Farming: Synthetic Themes for Linking Farmers to Demanding Markets." Pp. 21-38 in Contract Farming for Inclusive Market Access, edited by C. Da Silva and M. Ranking. Rome: FAO.

Jones, Sam and Peter Gibbon. 2011. "Developing Agricultural Markets in Sub-Saharan Africa: Organic Cocoa in Rural Uganda." Journal of Development Studies 47(10):1595-618.

Keef, Stephen P and Leigh A Roberts. 2004. "The Meta-Analysis of Partial Effect Sizes." British Journal of Mathematical and Statistical Psychology 57(1):97-129.

Key, Nigel and David Runsten. 1999. "Contract Farming, Smallholders, and Rural Development in Latin America: The Organization of Agroprocessing Firms and the Scale of Outgrower Production." World Development 27(2):381-401.

Lindholm, Annika. 2014. "The (Re) Emergence of Contract Farming in Sub-Saharan Africa: Moving from Land Grab to Power Grab?".

Little, Peter D and Michael Watts. 1994. Living under Contract: Contract Farming and Agrarian Transformation in Sub-Saharan Africa: Univ of Wisconsin Press.

Masset, Edoardo, Lawrence Haddad, Alex Cornelius and Jairo Isaza-Castro. 2011. "A Systematic Review of Agricultural Interventions That Aim to Improve Nutritional Status of Children." London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

McKenzie, D. 2012. "Beyond Baseline and Follow-Up: The Case for More T in Experiments." Journal of Development Economics 99(2):210-21.

Minot, Nicholas. 1986. "Contract Farming and Its Effect on Small Farmers in Less Developed Countries." Vol.: Michigan State University, Department of Agricultural, Food, and Resource Economics.

Minot, Nicholas and Loraine Ronchi. 2015. "Contract Farming: Risks and Benefits of Partnership between Farmers and Firms."

Minten, Bart, Lalaina Randrianarison and Johan FM Swinnen. 2009. "Global Retail Chains and Poor Farmers: Evidence from Madagascar." World Development 37(11):1728-41.

Mitchell, Jonathan, Jodie Keane and Christopher Coles. 2009. Trading Up: How a Value Chain Approach Can Benefit the Rural Poor: Overseas development institute (ODI).

Narayanan, Sudha. 2013. "Smallholder Attrition in Contract Farming Schemes in India: Extent, Causes, and Concerns." Food Chain 3(3):155-70. doi: 10.3362/2046-1887.2013.015.

Otsuka, Keijiro, Yuko Nakano and Kazushi Takahashi. 2016. "Contract Farming in Developed and Developing Countries." Annual Review of Resource Economics 8:353-76.

Oya, Carlos. 2012. "Contract Farming in Sub-Saharan Africa: A Survey of Approaches, Debates and Issues." Journal of Agrarian Change 12(1):1-33.

Peters, Jaime L, Alex J Sutton, David R Jones, Keith R Abrams and Lesley Rushton. 2007. "Performance of the Trim and Fill Method in the Presence of Publication Bias and

between-Study Heterogeneity." Statistics in medicine 26(25):4544-62.

Prowse, Martin. 2012. "Contract Farming in Developing Countries - a Review." Vol. A Savoir. Antwerpen.

Ragin, Charles. 2008. Redesigning Social Inquiry: Fuzzy Sets and Beyond: University of Chicago Press.

Ragin, Charles and Sean Davey. 2009. "Fs/Qca, Version 2.5 (Computer Program)." Tucson: University of Arizona.

Ravallion, M. 2009. "Should the Randomistas Rule?". The Economists' Voice 6(2):6.

Reichert, Christopher and Claude Rubinson. 2014. "Kirq, Version 2.1. 12 [Computer Program]." Houston, TX: University of Houston-Downtown.

Rihoux, B and CC Ragin, eds. 2009. Configurational Comparative Methods: Qualitative Comparative Analysis (Qca) and Related Techniques, Vol. Applied Social Research Methods. Thousand Oaks and London: Sage.

Schipmann, Christin and Matin Qaim. 2010. "Spillovers from Modern Supply Chains to Traditional Markets: Product Innovation and Adoption by Smallholders." Agricultural Economics 41(3-4):361-71.

Setboonsarng, Sununtar and PingSun Leung. 2014. "Making Globalization Work Better for the Poor through Contract Farming."

Shadish, W.R., T. D. Cook and D. T. Campbell. 2002. Experimental and Quasi-Experimental Designs for Generalized Causal Inference: Houghton Mifflin Co. Boston, MA.

Sivramkrishna, Sashi and Amalendu Jyotishi. 2008. "Monopsonistic Exploitation in Contract Farming: Articulating a Strategy for Grower Cooperation." Journal of International Development 20(3):280-96.

Thomas, J., J. Brunton and S. Graziosi. 2010. "Eppi-Reviewer 4.0: Software for Research Synthesis." EPPI-Centre Software. London: Social Science Research Unit, Institute of Education.

Ton, G, and J Mheen-Sluijer. 2009. "Sense and Sensibilities in Contract Farming: A Checklist for Reflection on Critical Issues in Contract Farming Arrangements in Developing Countries." Wageningen: LEI Wageningen - ICCO - EVD.

Ton, G, W Vellema, M D'Haese, S Desiere, S Weituschat, A Brouwer and H van Hoeven. 2015. "3ie Systematic Review of the Effectiveness of Contract Farming for Income and Food Security of Smallholder Farmers in Low- and Middle Income Countries: Protocol Sr6.1088." Vol. Systematic Reviews. London and Delhi: International Initiative for Impact Evaluation (3ie).

Trifković, N. 2016. "Comparison of Farm/Household Characteristics Using T-Tests - Table 5." edited by G. Ton. Copenhagen.

UNIDROIT-FAO. 2016. "UNIDROIT/FAO Legal Guide on Contract Farming." Vol.: IFAD-FAO-UNIDROIT.

United Nations Population Division. 2014. World Urbanization Prospects: The 2014 Revision: United Nations Publications.

von Hagen, Oliver and G Alvarez. 2011. "The Impacts of Private Standards on Global Value Chains." Vol. Literature Review Series on the Impacts of Private Standards, Part I. Geneva: International Trade Centre.

Vorley, Bill and Felicity Proctor. 2008. "Inclusive Business in Agrifood Markets: Evidence and Action."

Waddington, Hugh, Birte Snilstveit, Jorge Garcia Hombrados, Martina Vojtkova, Jock Anderson and Howard White. 2014. "Farmer Field Schools for Improving Farming Practices and Farmer Outcomes in Low-and Middle-Income Countries: A Systematic Review." Campbell systematic reviews 10(6).

Wang, Holly, Yanbing Wang and Michael S. Delgado. 2014b. "The Transition to Modern Agriculture: Contract Farming in Developing Economies." American Journal of Agricultural Economics. doi: 10.1093/ajae/aau036.

Warning, M. and N. Key. 2002a. "The Social Performance and Distributional Consequences of Contract Farming: An Equilibrium Analysis of the Arachide De Bouche Program in Senegal." World Development 30(2):255-163.

Warning, Matthew and Nigel Key. 2002b. "The Social Performance and Distributional Consequences of Contract Farming: An Equilibrium Analysis of the Arachide De Bouche Program in Senegal." World Development 30(2):255-63.

Wendimu, Mengistu Assefa, Arne Henningsen and Peter Gibbon. 2016. "Sugarcane Outgrowers in Ethiopia:“Forced” to Remain Poor?". World Development 83:84-97.

Will, Margret. 2013. "Contract Farming Handbook: A Practical Guide for Linking Small-Scale Producers and Buyers through Business Model Innovation." Vol. Bonn and Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ).

Wong, Geoff, Trish Greenhalgh, Gill Westhorp, Jeanette Buckingham and Ray Pawson. 2013. "Rameses Publication Standards: Realist Syntheses." BMC medicine 11(1):21.

World Bank. 2007. World Development Report 2008: Agriculture for Development. Washington: World Bank.

Other publications in the 3ie Systematic Review Series

The following reviews are available at
http://www.3ieimpact.org/publications/systematic-review-publications/

Interventions to improve the labour market outcomes of youth: a systematic review of training, entrepreneurship promotion, employment services and subsidized employment interventions. 3ie Systematic Review 37. Kluve J, Puerto S, Robalino D, Romero JM, Rother F, Stöterau J, Weidenkaff F and Witte M, 2017.

Promoting handwashing and sanitation behaviour change in low- and middle-income countries: a mixed-method systematic review. 3ie Systematic Review 36. Buck, ED, Remoortel, HV, Hannes, K, Govender, T, Naidoo, S, Avau, B, Veegaete, AV, Musekiwa, A, Lutje, V, Cargo, M, Mosler, HJ, Vandekerckhove, P and Young T (2017)

Incorporating the life cycle approach into WASH policies and programmes: A systematic review. 3ie Systematic Review 35. Annamalai, TR, Narayanan, S, Devkar, G, Kumar, VS, Devaraj, R, Ayyangar, A and Mahalingam, A (2017)

Effects of certification schemes for agricultural production on socio-economic outcomes in low- and middle-income countries: a systematic review 34. Oya, C, Schaefer, F, Skalidou, D, McCosker, C and Langer, L (2017)

Short-term WASH interventions in emergency response: a systematic review. 3ie Systematic Review 33. Yates, T, Allen, J, Joseph, ML and Lantagne, D (2017)

Community monitoring interventions to curb corruption and increase access and quality of service delivery in low- and middle-income countries. 3ie Systematic Review 32. Molina E, Carella L, Pacheco A, Cruces, G and Gasparini, L (2016)

Effects and mechanisms of market-based reforms on access to electricity in developing countries: a systematic review. 3ie Systematic Review 31. Bensch, G, Sievert, M, Langbein, J, Kneppel, N (2016)

Youth gang violence and preventative measures in low- and middle-income countries: a systematic review (Part II), 3ie Systematic Review 30. Higginson, A, Benier, K, Shenderovich, Y, Bedford, L, Mazerolle, L, Murray, J (2016)

Youth gang membership and violence in low- and middle-income countries: a systematic review (Part I), 3ie Systematic Review 29. Higginson, A, Benier, K, Shenderovich, Y, Bedford, L, Mazerolle, L, Murray, J (2016)

Cash-based approaches in humanitarian emergencies: a systematic review, 3ie Systematic Review Report 28. Doocy, S and Tappis, H (2016)

Factors affecting uptake of voluntary and community-based health insurance schemes in low-and middle-income countries: a systematic review, 3ie Systematic Review 27. Panda, P, Dror, IH, Koehlmoos, TP, Hossain, SAS, John, D, Khan, JAM and Dror, DM (2016)

Parental, community and familial support interventions to improve children's literacy in developing countries: a systematic review, 3ie Systematic Review 26. Spier, E, Britto, P,

Pigott, T, Roehlkapartain, E, McCarthy, M, Kidron, Y, Song, M, Scales, P, Wagner, D, Lane, J and Glover, J (2016)

Business support for small and medium enterprises in low- and middle-income countries: a systematic review, 3ie Systematic Review 25. Piza, C, Cravo, T, Taylor, L, Gonzalez, L, Musse, I, Furtado, I, Sierra, AC and Abdelnour, S (2016)

Interventions for improving learning outcomes and access to education in low- and middle- income countries: a systematic review, 3ie Systematic Review 24. Snilstveit, B, Stevenson, J, Phillips, D, Vojtkova, M, Gallagher, E, Schmidt, T, Jobse, H, Geelen, M, Pastorello, M, and Eyers, J (2015)

Economic self-help group programmes for improving women's empowerment: a systematic review, 3ie Systematic Review 23. Brody, C, De Hoop, T, Vojtkova, M, Warnock, R, Dunbar, M, Murthy, P and Dworkin, SL (2016)

The identification and measurement of health-related spillovers in impact evaluations: a systematic review, *3ie Systematic Review 22*. Benjamin-Chung, J, Abedin, J, Berger, D, Clark, A, Falcao, L, Jimenez, V, Konagaya, E, Tran, D, Arnold, B, Hubbard, A, Luby, S, Miguel, E and Colford, J (2015)

The effects of school-based decision-making on educational outcomes in low- and middle-income countries: a systematic review, 3ie Systematic Review Report 21. Carr-Hill, R, Rolleston, C, Pherali, T and Schendel, R, with Peart, E, and Jones, E (2015)

Policing interventions for targeting interpersonal violence in developing countries: a systematic review, *3ie Systematic Review 20*. Higginson, A, Mazerolle, L, Sydes, M, Davis, J, and Mengersen, K (2015)

The effects of training, innovation and new technology on African smallholder farmers' wealth and food security: a systematic review, *3ie Systematic Review 19*. Stewart, R, Langer, L, Rebelo Da Silva N, Muchiri, E, Zaranyika, H, Erasmus, Y, Randall, N, Rafferty, S, Korth, M, Madinga, N and de Wet, T (2015)

Community based rehabilitation for people with disabilities in low- and middle-income countries: a systematic review, *3ie Systematic Review 18*. Iemmi, V, Gibson, L, Blanchet, K, Kumar, KS, Rath, S, Hartley, S, Murthy, GVS, Patel, V, Weber, J and Kuper H (2015)

Payment for environmental services for reducing deforestation and poverty in low- and middle-income countries: a systematic review, *3ie Systematic Review 17.* Samii, C, Lisiecki, M, Kulkarni, P, Paler, L and Chavis, L (2015)

Decentralised forest management for reducing deforestation and poverty in low- and middle- income countries: a systematic review, *3ie Systematic Review 16.* Samii, C, Lisiecki, M, Kulkarni, P, Paler, L and Chavis, L (2015)

Supplementary feeding for improving the health of disadvantaged infants and young children: a systematic and realist review, *3ie Systematic Review 15.* Kristjansson, E, Francis, D, Liberato, S, Greenhalgh, T, Welch, V, Jandu, MB, Batal, M, Rader, T, Noonan, E, Janzen, L, Shea, B, Wells, GA and Petticrew, M (2015)

The impact of land property rights interventions on investment and agricultural productivity in developing countries: a systematic review, 3ie Systematic Review Report 14. Lawry, S, Samii, C, Hall, R, Leopold, A, Hornby, D and Mtero, F, 2014.

Slum upgrading strategies and their effects on health and socio-economic outcomes: a systematic review, 3ie Systematic Review 13. Turley, R, Saith, R., Bhan, N, Rehfuess, E, and Carter, B (2013)

Services for street-connected children and young people in low- and middle-income countries: a thematic synthesis, 3ie Systematic Review 12. Coren, E, Hossain, R, Ramsbotham, K, Martin, AJ and Pardo, JP (2014)

Why targeting matters: examining the relationship between selection, participation and outcomes in farmer field school programmes, 3ie Systematic Review 11. Phillips, D, Waddington, H and White, H (2015)

The impact of export processing zones on employment, wages and labour conditions in developing countries, 3ie Systematic Review 10. Cirera, X and Lakshman, R (2014)

Interventions to reduce the prevalence of female genital mutilation/cutting in African countries, 3ie Systematic Review 9. Berg, RC and Denision, E (2013)

Behaviour change interventions to prevent HIV among women living in low and middle income countries, 3ie Systematic Review 8. McCoy, S, Kangwende, RA and Padian, NS (2009)

The impact of daycare programs on child health, nutrition and development in developing countries, 3ie Systematic Review 7. Leroy, JL, Gadsden, P and Guijarro, M (2011)

Willingness to pay for cleaner water in less developed countries: Systematic review of experimental evidence, 3ie Systematic Review 6. Null, C, Hombrados, JG, Kremer, M, Meeks, R, Miguel, E and Zwane, AP (2012)

Community-based intervention packages for reducing maternal morbidity and mortality and improving neonatal outcomes, 3ie Systematic Review 5. Lassi, ZS, Haider, BA and Langou, GD (2011)

The effects of microcredit on women's control over household spending: a systematic review, 3ie Systematic Review 4. Vaessen, J, Rivas, A, Duvendack, M, Jones, RP, Leeuw, F, van Gils, G, Lukach, R, Holvoet, N, Bastiaensen, J, Hombrados, JG and Waddington, H, (2013).

Interventions in developing nations for improving primary and secondary school enrolment of children: a systematic review, 3ie Systematic Review 3. Petrosino, A, Morgan, C, Fronius, T, Tanner-Smith, E, and Boruch, R, 2016.

Interventions to promote social cohesion in Sub-Saharan Africa, 3ie Systematic Review 2. King, E, Samii, C and Snilstveit, B (2010)

Water, sanitation and hygiene interventions to combat childhood diarrhoea in developing countries, 3ie Systematic Review 1. Waddington, H, Snilstveit, B, White, H and Fewtrell, L (2009)

International Initiative for Impact Evaluation
London International Development Centre
36 Gordon Square
London WC1H 0PD
United Kingdom

3ieuk@3ieimpact.org
Tel: +44 207 958 8351/8350

www.3ieimpact.org