Journal of Development Effectiveness
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rjde20

FairTrade's theory of change: an evaluation based on the cooperative life cycle framework and mixed methods

Gian Nicola Francesconi[a] & Ruerd Ruben[bc]

[a] International Centre for Tropical Agriculture (CIAT-CGIAR) and International Food Policy Research Institute (IFPRI-CGIAR), Titre 3396, Lot n.2, Almadies, Dakar BP 24063, Senegal

[b] The Independent Policy and Operations Evaluation (IOB), The Netherlands Ministry of Foreign Affairs, the Hague, Netherlands

[c] Centre of Development Issues (CIDIN), Radboud University, Nijmegen, Netherlands

Published online: 14 May 2014.

To cite this article: Gian Nicola Francesconi & Ruerd Ruben (2014) FairTrade's theory of change: an evaluation based on the cooperative life cycle framework and mixed methods, Journal of Development Effectiveness, 6:3, 268-283, DOI: 10.1080/19439342.2014.918164

To link to this article: http://dx.doi.org/10.1080/19439342.2014.918164

Journal of Development Effectiveness, 2014
Vol. 6, No. 3, 268–283, http://dx.doi.org/10.1080/19439342.2014.918164

Routledge
Taylor & Francis Group

FairTrade's theory of change: an evaluation based on the cooperative life cycle framework and mixed methods

Gian Nicola Francesconi[a]* and Ruerd Ruben[b,c]

[a]*International Centre for Tropical Agriculture (CIAT-CGIAR) and International Food Policy Research Institute (IFPRI-CGIAR), Titre 3396, Lot n.2, Almadies, Dakar BP 24063, Senegal;* [b]*The Independent Policy and Operations Evaluation (IOB), The Netherlands Ministry of Foreign Affairs, the Hague, Netherlands;* [c]*Centre of Development Issues (CIDIN), Radboud University, Nijmegen, Netherlands*

This study presents a quasi-experimental analysis of the impact of FairTrade certification on the commercial performance of coffee farmers in Tanzania. In doing so the study emphasises the importance of a well-contextualised theory of change as a basis for evaluation design. It also stresses the value of qualitative methods to control for selection bias. Based on a longitudinal (pseudo-panel) dataset comprising both certified and conventional farmers, it shows that FairTrade certification introduced a disincentive to farmers' commercialisation. We explain this counterintuitive conclusion on the basis of the 'cooperative life cycle' theory developed by US agribusiness scholars.

Keywords: FairTrade; impact evaluation; coffee; Tanzania; smallholder agriculture

1. Introduction

The objective of this study is to improve general understanding about the impact of FairTrade (FT) certification on the income of coffee farmers. Impact evaluators from a range of perspectives all stress the importance for analyses of this kind to be embedded in a well-contextualised theory of change, which helps in understanding the causal pathways behind why an intervention works or does not work (White 2011, 2009; Rogers 2007; Carvalho and White 2004; Pawson and Tilley 1997). The theory of change claimed by FT coffee is that certification provides farmers with an economic safety net, in the form of a guaranteed floor price, and a social premium (or equity investment), which respectively reduce rural livelihoods' vulnerability and add value to agricultural production. In particular, certified smallholders are paid a floor price of 1.40 USD per pound for washed Arabica and 1.35 USD for unwashed Arabica, or the market price if the latter is higher than the FT floor price. In addition to this, FT coffee provides a social premium of USD 20 cents per pound (with USD 5 cents earmarked for productivity and quality improvements) to be invested for the development of rural communities. It is important to stress that Fairtrade coffee does not certify individuals, but organisations owned and controlled by small-scale farmers, such as cooperatives and similar associative and community-based institutions. Certified cooperatives are expected to enforce environmental and quality standards, restricting the use of agrochemicals and encouraging the adoption of efficient production practices.[1]

*Corresponding author. Email: g.n.francesconi@cgiar.org

The commercialisation of FT coffee has grown significantly over the past 20 years, contributing to sensitisation of urban consumers to the economic difficulties faced by rural smallholders (Potts, van der Meer, and Daitchman 2010). However, the effect of FT certification on rural livelihoods appears to be increasingly contested. Although early evaluations indicated that FT certification produced the intended impacts, recent progress in evaluation methodology has started to shed some light on unintended outcomes as well. The first evaluations of FT initiatives were based on anecdotal evidence from Costa Rica (Ronchi 2002), Nicaragua (Bacon 2005; Bacon et al. 2008) and Mexico (Jaffee 2007; Calo and Wise 2005; Milford 2004). These studies highlighted that the floor price provided by FT did help farmers better cope with the coffee crisis of the early 1990s. Additional anecdotal evidence, produced at the beginning of the new millennium and compiled by the European Fair Trade Association (EFTA), further concluded that the premium provided by FT promoted accumulation of equity capital by producers' organisations.

However, consequent evaluations based on quantitative analyses contributed to emphasis an important trade-off associated with FT coffee certifications. Using longitudinal data from a certified coffee cooperative in Central America, De Janvry, McIntosh, and Sadoulet (2010) demonstrated that the positive impact of the FT floor price on farmers' income tends to be offset by the 'crowding-in' of certified organisations. According to these authors, when market price is below the FT floor price, farmers face a strong incentive to join certified organisations; but as membership increases, per capita price incentives tend to decrease. The authors explain that this crowding-in, or over-certification, problem arises because the demand for FT coffee is fixed at any given point in time. The underlying rationale of FT certification thus appears to be misleading since it aims to control price without controlling supply volumes. It follows that the introduction of a floor price can improve farmers' revenues only if certified cooperatives enforce adequate membership rules for coordinating members' entry and exit.

The theory of change elaborated by De Janvry, McIntosh, and Sadoulet (2010) is supported by the mixed and inconclusive results reported by recent evaluations of FT schemes based on cross-section data and quasi-experimental design (see Blackman and Rivera 2010). The crowding-in theory is further supported by the work of Balineau (2013), who showed that conventional farmers operating in the proximity of certified cooperatives can easily adapt their production practices in order to meet the sustainability and quality standards imposed by FT. These authors confirm that the barriers faced by farmers to enter FT schemes are low, favouring over-certification. Additional support for the theory of change advanced by De Janvry, McIntosh, and Sadoulet (2010) is found in agribusiness theory. In particular, the 'cooperative life cycle' theory, developed by Cook and Chambers (2007), suggests that growth in membership tends to favour the rise of free-rider and agency-cost problems. The free-rider problem occurs because the incorporation of new members dilutes returns for existing members, who then face an incentive to sell their coffee outside the cooperative channel (that is, to 'side-sell'). The agency-cost problem arises because membership expansion adds complexity to internal monitoring and decision-making efforts, creating an additional disincentive for members to engage in collective commercialisation. Compared with the theory by De Janvry, McIntosh, and Sadoulet (2010), Cook and Chambers (2007) further stress that growth-related problems are more likely to occur in cooperatives characterised by membership with heterogeneous socio-economic preferences.

Nevertheless, the theory of change advanced by De Janvry, McIntosh, and Sadoulet (2010) does not take into account potential impacts generated by the FT premium, which

is arguably the most important output of the certification process. While the floor price mechanism generates economic incentives only when the market price is particularly low, certified cooperatives are always expected to benefit from the premium mechanism. In particular, the FT premium is expected to provide farmers with an incentive to make equity (or collective) investments in their cooperative organisation in order to support community-driven development. Available evaluations based on cross-section data and quasi-experimental design indeed suggest that FT certification promotes smallholders' access to credit and equity investments, which translate into improved social welfare, but not necessarily into better incomes for member-farmers (Ruben and Fort 2012; Méndez et al. 2010; Arnould, Plastina, and Ball 2009; Ruben, Fort, and Zúñiga-Arias 2009; Ruben 2008). Méndez et al. (2010) argue that income effects fail to materialise because farmers' sales through FT market channels tend to be too limited to generate significant economic returns to collective investments, while Becchetti and Constantino (2006) suggest that households' revenues are constrained by the lack of involvement of farmers in the investment decisions taken by certified cooperatives. This body of empirical evidence thus suggests that even the premium mechanism of FT may not always work as intended. However, the different explanations raised to date have still to be added together into a consistent theory of change.

The general hypothesis postulated for this study, in order to explain potential side-effects of the FT premium, draws once again on the 'cooperative life cycle' theory defined by agribusiness scholars (Cook and Chambers 2007). This theory suggests that incentives to make collective investments, like those provided by the FT premium, promote the rise of influence costs and horizon and portfolio problems in agricultural cooperatives. External incentives to invest encourage cooperatives to engage in more diversified and longer-term investment strategies, which exceed the needs of member-farmers in terms of investment portfolio and horizon. As a result, members face an incentive to side-sell, in order to reduce their contribution to collective investments. The severity of these horizon and portfolio problems intensifies when external incentives induce cooperatives to make investments that entail intangible economic returns, as may be the case for certain 'social' investments promoted by FT. In addition to this, investments in a cooperative add organisational complexity and tend to increase the need for managerial and technical expertise, which is usually found outside a cooperative. However, the involvement of professional outsiders in a cooperative contributes to reducing members' influence in decision-making processes. Due to increasing influence costs, a cooperative evolves from being a direct extension of farmers into a separate firm favouring elite capture – that is, the administrative elite comprising leaders, managers, technicians, and so forth captures most returns to investments – and encourages members' side-selling. Finally, Cook and Chambers (2007) stress that investment-induced problems, like growth-induced ones, are expected to be prevalent in cooperatives characterised by internal heterogeneity in terms of members' socio-economic preferences.

In conclusion, the general theory of change considered for this study suggests that the impact of FT on farmers' income is temporary, that it tends to vanish over time. This hypothesis was validated in contexts where FT certifications generated a price incentive, but it remains to be tested in situations where FT certifications provided only a social premium. To do so, the remainder of the study analyses the prevalence of horizon-portfolio and influence-cost problems, as well as side-selling behaviour in a Tanzanian association comprising both FT certified and conventional coffee cooperatives, over a period characterised by relatively high market prices (above the FT floor price).[2] In particular, the remainder of this paper provides: i) a qualitative, factual analysis of the

certification process and overall performance that underlined the development of the association under study; ii) a mixed, qualitative and quantitative analysis of potential sources of selection bias in the certification process; iii) a quantitative, counterfactual analysis of the impact generated by the FT premium on farmers' side-selling, based on a quasi-experimental, difference-in-difference evaluation design; and iv) conclusions that aims to add new insights and further streamline FT's theory of change, as well as implications to maximise benefits for coffee farmers over time.

2. Setting the stage

Impact evaluation requires a combination of factual and counterfactual analysis drawing upon a range of data sources and methodological approaches (White 2011). Impact evaluators from a range of perspectives all stress the importance of proper understanding of the specific intervention under study and its unique context, as a basis to perform a good evaluation (De Hoop 2011; White 2011; Karlan 2009; Ravallion 2009; Rogers 2009). To do so, an impact evaluation should not be solely concerned with producing a counterfactual measure of impact; it should also include an in-depth factual analysis drawing upon qualitative information retrieved through extensive interactions between researchers and the beneficiaries of the intervention. As such, this section presents a factual analysis of the FT certification process undergone by the Tanzanian association under study. This analysis is based on qualitative information retrieved by the first author during two field visits (two weeks in total) to one of the districts where the association operates.[3] Within this district the first author carried out focus group discussions involving a total of 10 cooperatives and 20 member-farmers, as well as several key informant interviews with the local officers of the overarching association.

During most of the previous century, the coffee association considered in this study was embedded in a parastatal structure supported by colonial and post-independence governments in order to centralise control over coffee production and commercialisation. However, over the last two decades the state has gradually withdrawn its participation from the national coffee sector. As a result, the association is nowadays operating as an independent marketing agency, purchasing and selling the coffee produced by its member-cooperatives. Yet, market liberalisation reforms have also opened the national coffee market to individual investors, working as coffee traders, who are nowadays competing with the association in procuring and marketing the coffee produced by cooperative farmers. The competitive pressure of these middlemen has been growing significantly over the past decade, to the extent that side-selling by member-cooperatives has grown to become the major problem faced by the association. Over time, member-cooperatives have been increasingly reluctant to sell their coffee to the association, favouring the expansion of trading relationships with middlemen. To retain its comparative advantage, the association has been developing services that provide all member-cooperatives with advice on technical and organisational issues, as well as agricultural inputs on credit. In addition to these services, in 2007 the association decided to enter an agreement with FT and started a campaign aiming to gradually certify all its cooperatives through a phasing-in strategy.

By the time the first author had completed the fieldwork (end of 2012), FT had granted certifications to the majority of the cooperatives affiliated to the association. Only a few member-cooperatives were still waiting to be incorporated in the certification scheme. It is important to stress that by the end of 2012 certified cooperatives had only received yearly premiums from FT buyers. The latter had not provided any price incentive

because market prices had been consistently above the FT floor price since the beginning of the certification process. FT premiums had been mostly used to: i) build the district office of the association, equip it with chairs and desks, computers, Internet and vehicles, and staff it with managers and technicians; and ii) develop multiple decentralised infrastructures, such as community-based warehouses, washing stations and water tanks, managed by the association's cooperatives. Despite these investments, by the end of 2012 the commercial relationship between the association and its cooperatives appeared to be tense. In particular, side-selling by member-cooperatives was so large that the association could not put together the minimum amount of coffee required by FT buyers. Due to increasing side-selling, the association could no longer repay its creditors and began to downsize its personnel. On the other hand, some member-cooperatives decided to drop out of the association and establish a new association.

Focus group discussions with the representatives of member-cooperatives revealed that by the end of 2012 the association was widely perceived as a separate firm, mainly concerned with its own well-being rather than that of farmers. The farmers that participated in these discussions justified their side-selling by the need to contrast embezzlement at the level of the association. By contrast, the association's officers complained that farmers were too opportunistic: 'they [the farmers] benefited a lot from the services provided by the association, but they do not want to repay the association for these services, that is why they sell their coffee to traders'. The specific hypothesis postulated for this study is that the FT premium contributed to aggravation of this tension, especially between the association and its certified cooperatives, by generating an incentive to pursue long-term (the association's office) and highly diversified (community-level infrastructures) investments, resulting in the rise of horizon and portfolio problems. We also argue that the FT premium contributed to increase the decision power of the association vis-à-vis member-farmers and thus influence costs, given that certified cooperatives could access FT buyers and receive premiums only through the association. In turn, horizon-portfolio and influence-cost problems created an incentive for certified cooperatives to side-sell.

3. Materials and methods

Because certification was not randomly allocated within the district under study, a naïve estimation of the impact of the FT premium could be significantly affected by selection bias. Different approaches can be adopted to identify potential sources of selection bias, depending on the available sample and data. In this case, the district officers of the association provided us with institutional records referring to 2010 and 2012 in regard to the farmers of 10 member-cooperatives. Out of these 10 cooperatives, nine had been already certified.[4] Such an unbalanced sample, including only one conventional cooperative, induced us to deploy a reversed matching strategy in order to minimise potential selection bias. In other words, we screened certified (treatment) cooperatives and identified those most similar to the conventional (control) cooperative, on the basis of the criteria used to allocate certification. Given that the records provided by the association comprised only three indicators referring to the productive assets and commercial performance of cooperatives' members, the matching exercise was complemented by qualitative information gathered through focus group discussions and key informant interviews. In particular, our matching approach considered those criteria applied in order to select cooperatives in the certification scheme, which also had the potential to influence farmers' performance, defined as the amount of coffee sold through the overarching association.

Two main selection criteria were identified at the level of cooperative organisations: their location and the amount of time spent as members of the association. The certification process aimed to cover all the cooperatives located within the district under study, through a phasing-in strategy that focused on one district area at a time. Differences in location could thus explain why some cooperatives were certified and others were not, as well as differences in the amount of coffee sold by farmers through the association. In fact, we might expect cooperatives located in particularly remote and less accessible areas to be less likely to side-sell their coffee to local traders. However, differences in location are expected to be negligible in our sample as the dataset obtained from the association included only cooperatives located along the main road of one specific district area. Such a sample, comprising neighbouring cooperatives, is thus expected to minimise selection bias due to location, but it could entail significant spill-over effects as farmers could have moved from conventional to certified cooperatives (or vice versa). However, spill-over effects can be assumed to be negligible since membership in cooperatives appeared to be defined on the basis of community or village linkages and boundaries. In other words, every cooperative in the district was clearly identified with a rural community or village, suggesting that cooperatives were not likely to include members from outside the community-village of origin.

According to the association's officers, the FT selection process also gave priority to those cooperatives that had been part of the overarching association for longer. The amount of coffee sold through the association is also expected to vary depending on the exposition of a cooperative to the commercial and production services provided by the association. Since the conventional cooperative joined the association in 2009 and there is only one certified cooperative in our sample that joined in that year, the latter was identified as the treatment group for the remainder of this analysis. Given that this treatment group obtained its FT certification in 2011 and that available data refer to 2010 and 2012, the remainder of the analysis adopts a difference-in-difference evaluation design.

The qualitative analysis above describes the rationale deployed to match the only conventional cooperative with the nearest matching certified cooperative in our sample. Residual differences between these two groups are expected to be related to the characteristics and preferences of member-farmers. These differences can be partly identified on the basis of available data. In particular, the records retrieved from the association provided two explanatory variables referring to farm assets, and in particular: i) the number of acres of land used for coffee production per farm; and ii) the number of coffee trees per acre, per farm. Table 1 shows that the number of acres used by the average farm for coffee production was significantly lower for the treatment group in 2010 (that is, before certification). However, the average density of coffee trees per farm was higher within the treatment group before certification. Table 1 also shows that these differences between treatment and control groups appear to be reversed in 2012, because the average farm registered with the conventional (or control) cooperative appears to have held significantly less land is characterised by higher tree density.

First, the statistics described above indicate the presence of an inverse relationship between the amount of land held for coffee production and the density of coffee trees. Second, it suggests that the certification was granted to the group comprising farms with less land, but higher density of coffee trees. Third, it shows that unlike the treatment cooperative, the control group witnessed a significant change in farming assets over time. This change implies that farmers holding a larger acreage of coffee land were induced to quit the control group after 2010. One reason for better-off member-farmers to quit a

Table 1. Farmer-level data provided by the association.

Variables	Certified coop. Mean (Std. Dev.)	Conventional coop. Mean (Std. Dev.)	t-value
2010 (before certification): Kg of coffee cherries sold to the association	2163 (1785)	1583 (1696)	1.86*
No. of coffee trees owned per acre	595 (155)	505 (143)	3.34**
Acres of coffee land owned	1.77 (0.61)	2.31 (1.05)	−3.30**
Heterogeneity in acres of land owned[a]	0.27 (0.21)	0.38 (0.24)	−2.62**
Degrees of freedom (no. of farmers)	*50*	*80*	*130*
2012 (after certification): Kg of coffee cherries sold to the association	448 (540)	1654 (872)	−9.14**
No. of coffee trees owned per acre	605 (92)	665 (242)	−1.87*
Acres of coffee land owned	1.76 (0.83)	1.44 (0.75)	2.11**
Heterogeneity in acres of land owned[a]	0.30 (0.37)	0.46 (0.24)	−2.70**
Degrees of freedom (no. of farmers)	*67*	*48*	*115*

Notes: *Denotes statistical significance at the 10% level.
**Denotes statistical significance at the 5% level.

[a]These indicators were computed as: $\frac{\sqrt{(x_i - X_i)^2}}{X_i}$, where x is the value reported for a farm i and X is the mean value for the cooperative i.

cooperative is internal disagreements and conflicts, inducing those who are more independent to quit the organisation. Internal disagreements and conflicts are expected to arise when a group is heterogeneous in terms of members' productive capacity. Since the level of internal cohesion could have been another important reason for not selecting the control group in the certification scheme, we captured it on the basis of a variable measuring the deviation between the area of coffee land held by a farmer and the mean value reported for his/her cooperative. Table 1 shows that in 2010 the control group was significantly more heterogeneous than the certified group in terms of landholding, and this difference had become even more significant by 2012.

All these observed differences between treatment and control farmers are also expected to influence the amount of coffee sold through the association by cooperative farmers. The total quantity of coffee commercialised is expected to increase with the amount produced by a farmer. Since the association provided no data on total coffee production at the farm level, we argue that the latter indicator can be proxied on the basis of farmers' productive assets (that is, acreage of coffee-land and density of coffee trees). In fact, all cooperatives, including our control group, had to comply with the production practices specified by FT in order to join the association in the first place. Furthermore, the association provided all member cooperatives and farmers with the same services (inputs on credit and technical advice). It thus seems reasonable to expect limited variability in production practices between treatment and control farmers. As such, variability in production volumes can be mostly attributed to differences in farm assets, given that coffee production is expected to increase with the amount of farm-land and the density of coffee trees. It is however important to note that when the density of coffee trees becomes too high productivity tends to decrease, implying that this relationship cannot be expected to be linear, but rather exponential and concave. Finally, heterogeneity in farmers' productive assets also has the potential to encourage their side-selling, as this is expected to favour the rise of disagreements and conflicts within a cooperative.

Differences in observed farmers' characteristics need to be controlled for in order to identify the impact attributable to certification and to the related premium. Before doing so, however, it is important to describe the differences estimated in regard to the amount of coffee sold through the association on the basis of naïve comparisons of treatment and control groups, for 2010 and 2012. In particular, Table 1 shows that in 2010 (before certification) the farmers registered with the treatment cooperative sold slightly more coffee through the association when compared with those in the control group. However, in 2012 the amount of coffee sold by certified farmers appeared to be significantly lower than that from conventional farmers. It is also important to note that the amount of coffee sold through the association by the average farmer from the control group stayed more or less the same over time, while that sold by the average farmer from the treatment group reported a fivefold decrease between 2010 and 2012. Although the FT impact cannot be properly identified on the basis of these naïve comparisons, the latter provides a general sense of the direction taken by similar cooperatives in the presence and absence of certification, which is consistent with our previous theoretical and factual analyses and therefore is expected also to be consistent with the results of the counterfactual analysis presented in the next section. The theoretical and factual analyses presented in the previous sections also justify our choice of indicator, as they suggest that farmers' increasing side-selling can be reasonably and, to a significant extent, attributed to the certification process and more specifically to the FT premium. Although such an indicator provides intermediate measurements of impact, alternative indicators measuring the ultimate impact of certification on farmers' income, welfare and so forth are more likely to be subject to measurement errors and confounding effects. Unlike for these noisy indicators, the variations observed in our indicator of choice can be more directly and confidently attributed to certification.

4. Results

The counterfactual analysis presented in this section draws on the work of Abate, Francesconi, and Getnet (2014), Francesconi and Heerink (2011), Bernard, Taffesse, and Gabre-Madhin (2008) and Godtland et al. (2004), in order to address the specific question raised in this study: would farmers be side-selling less in the absence of a FT certification scheme? To address this question we estimated the difference in the amount of coffee sold through the association, before (2010) and after certification (2012), by the treatment and control groups identified in the previous section, controlling for differences in observed characteristics at the farmer level. We did so using two estimation techniques: propensity score matching (PSM) and regression analysis (RA). Estimations based on PSM are particularly sensitive to datasets characterised by a small number of observations, especially within the control group. Since our dataset is rather small and the control group includes fewer observations than the treatment group for 2012, estimations based on regression analysis are thus expected to be more reliable than those obtained through PSM. However, PSM allows one to define areas of common support (for both 2010 and 2012), which are the areas where treatment and control farmers overlap in terms of their propensity to participate in the certification scheme, given their observed characteristics. By restricting RA to the sub-sample including only those observations falling within the common support area, we expect to further increase the efficiency of impact estimates. In addition to this, the comparison of impact estimates produced using PSM and RA is expected to stress the consistency of our findings.

Table 2. Probit regressions explaining farmer participation in the treatment cooperative.

Membership in treatment coop. = 1 Membership in control coop. = 0	2010 (before certification)	2012 (after certification)
Acres of coffee land owned	−1.16 (0.35)**	1.27 (0.26)**
No. of coffee trees owned per acre	−0.00 (0.00)	0.01 (0.01)
(No. of coffee trees owned per acre)^2	0.00 (0.00)	−0.00 (0.00)
Heterogeneity in acres of land owned	−4.31 (1.06)**	−2.94 (0.75)**
No. of farmers	*130*	*115*
Pseudo-R^2	*0.24*	*0.30*

Notes: Robust standard error in parentheses.
**Denotes statistical significance at the 5% level.

The PSM analysis involved two steps. First, we estimated the propensity of farmers to be in the treatment group on the basis of observed farmers' characteristics, using a Probit estimator (Table 2). It is important to note that the relationships estimated in Table 2 are consistent with the statistics estimated in Table 1 (on the basis of t-tests) and described in the previous section. Second, we omitted a few observations that fell outside the common support area and matched the remainder across treatment and control groups on the basis of their propensity scores, using both kernel and nearest-neighbour techniques. Figure 1 indicates that the area of common support for 2010 corresponds to propensity scores higher than 0.07 and lower than 0.75, while Figure 2 shows that the area of common support for 2012 corresponds to propensity scores lower than 0.68. The robustness of the results is supported by a satisfactory balance, at the 10 per cent level, in the number of observations for treatment and control groups for each block of propensity scores (Table 3). Table 4 shows the results obtained through PSM analysis, which stress that

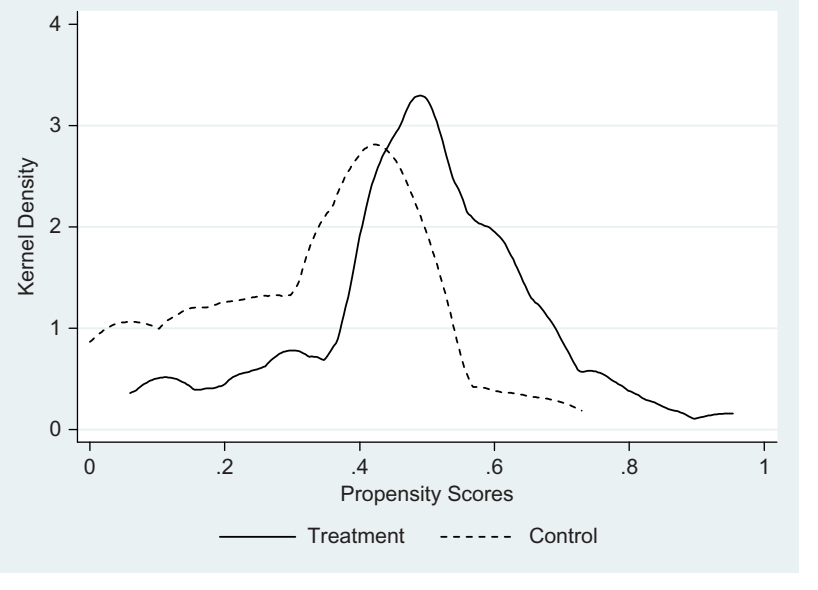

Figure 1. Common support area, 2010.

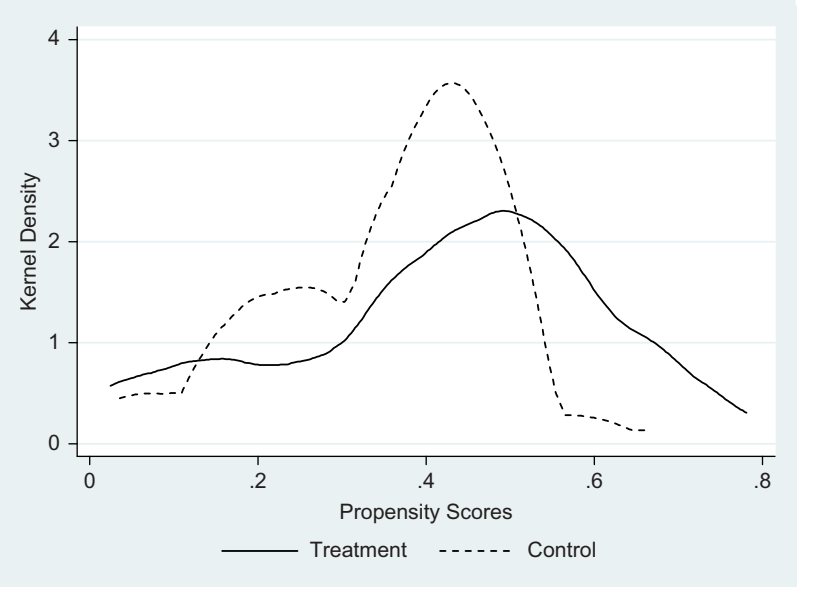

Figure 2. Common support area, 2012.

Table 3. Balancing properties after propensity scoring within common support area.

Year	Blocks of p-score	Control coop.	Treatment coop.	Total
	0.03	14	2	16
	0.2	25	4	29
2010	0.4	23	20	43
	0.6	2	17	19
	0.8	0	3	3
	Total	64	46	110
	0.06	1	2	3
	0.1	9	0	9
2012	0.2	4	7	11
	0.4	1	8	9
	0.45	13	1	14
	0.5	4	0	4
	0.6	11	0	11
	0.7	0	7	7
	0.8	0	42	42
	Total	43	67	110

certification may have had a negative impact on the amount of coffee sold by farmers through the association. According to Table 4, members of the treatment group were on average side-selling significantly less than the control group in 2010 (at the 5% level based on kernel matching technique and at the 10% level based on the nearest-neighbour technique). However, this scenario changed radically after certification, given that the treatment group appeared to be side-selling significantly more (at the 5% level or less, for both matching techniques) by 2012. These results lend general support to the naïve impact

Table 4. PSM results (kg sold by average treatment farmer less kg sold by average control farmer).

Production year	PSM Kernel	PSM Nearest neighbour
2010	878 (246)**	841 (485)*
(before certification)	*Farmers in treatment coop.: 46*	*Farmers in treatment coop.: 46*
	Farmers in control coop.: 64	*Farmers in control coop.: 34*
2012	−816 (147)**	−895 (234)**
(after certification)	*Farmers in treatment coop.: 67*	*Farmers in treatment coop.: 67*
	Farmers in control coop.: 43	*Farmers in control coop.: 18*

Notes: Standard error in parentheses.
*Denotes statistical significance at the 10% level.
**Denotes statistical significance at the 5% level.

estimates presented in Table 1, even if the latter reports a smaller difference in the performance of treatment and control groups in 2010, and a larger difference for 2012. Despite these inconsistencies, the magnitude of the overall impact inferred from Tables 1 and 4 appears to be very similar.

The results obtained through RA are presented in Tables 5 and 6. These OLS regressions explain the amount of coffee sold through the association by a farmer on the basis of participation in the treatment cooperative and farm household characteristics observed. It is important to note that all estimations reported in Tables 5 and 6 were obtained using robust standard errors adjusted for two clusters: treatment and control groups. However, the estimates reported in Table 5 are for the full sample, while those presented in Table 6 are restricted to the common support areas identified above (Figures 1 and 2). Both Tables 5 and 6 indicate that the sign of the relationships between farmers' observed characteristics and the amount sold through the association is in line with our expectations (see previous section). In fact, the amount of coffee sold increases with the amount of land available to a farmer for coffee production and with the amount of coffee trees per acre, but decreases when the density of coffee trees becomes too high (affecting productivity) and when members' landholding is particularly heterogeneous within a cooperative. However, these characteristics observed appear to be generally insignificant in explaining the amount of coffee sold through the association, given

Table 5. OLS regressions explaining the amount of coffee sold by a farmer through the association (no common support).

Kg of coffee cherries sold through the association	2010 (before certification)	2012 (after certification)
Membership in treatment cooperative	778 (118)*	−1327 (6.95)**
Acres of coffee land owned	509 (184)	431 (190)
No. of coffee trees owned per acre	2.79 (4.02)*	3.66 (0.99)
(No. of coffee trees owned per acre)^2	−0.00 (0.00)	−0.00 (0.00)
Heterogeneity in acres of land owned	−801 (12.8)**	−207 (215)
No. of farmers	*130*	*115*
R^2	*0.10*	*0.64*

Notes: Figures in parenthesis () denote robust standard errors adjusted for two clusters: treatment and control groups.
*Denotes statistical significance at the 10% level.
**Denote statistical significance at the 5% level.

Table 6. OLS regressions explaining the amount of coffee sold by a farmer through the association (with common support).

Kg of coffee cherries sold through the association	2010 (before certification)	2012 (after certification)
Membership in treatment cooperative	787 (163)	−1332 (17.7)**
Acres of coffee land owned	225 (322)	416 (202)
No. of coffee trees owned per acre	1.35 (1.35)	3.83 (0.89)
(No. of coffee trees owned per acre)^2	−0.00 (0.00)	−0.00 (0.00)
Heterogeneity in acres of land owned	−176 (188)	−160 (177)
N. of farmers	*110*	*110*
R^2	*0.21*	*0.66*

Notes: Figures in parentheses denote robust standard errors adjusted for two clusters: treatment and control groups.
**Denotes statistical significance at the 5% level.

participation in the treatment group. Only in Table 5 (RA on the full sample) do tree density and heterogeneity in landholding appear to have a significant effect on the amount sold through the association in 2010.

The most important difference between the results presented in Tables 5 and 6 is, however, the effect estimated for participation in the treatment cooperative. In particular, Table 5 indicates that member-farmers of the treatment cooperative sold a significantly larger (at the 10% level) amount of coffee through the association by 2010. By contrast, Table 6 shows that once we restricted the analysis to the area of common support, the effect of participation in the treatment cooperative remains positive but is no longer significant. Since the estimates in Table 6 are expected to be the most efficient and consistent, and considering also that Tables 1 and 4 indicate that differences in side-selling between the two groups were less markedly significant in 2010 (except for the kernel matching technique in Table 4), we can conclude that the amount of coffee sold through the association by treatment and control groups was similar before certification. This is an important result, because it allows us to argue, within reasonable doubt, that the two cooperatives were similar before certification – not only in terms of location and affiliation with the overarching association (see previous section), but also in terms commercial behaviour, ceteris paribus – and therefore it makes sense to compare them.

Finally, both Tables 5 and 6 indicate that after certification (in 2012) the members of the treatment group sold a significantly lower (at the 5% level or less) amount of coffee through the association compared with the members of the control group. Side-selling by the average farmer in the certified group is estimated at 1327–1332 kg of coffee cherries by 2012 (that is, after certification). This finding thus suggests that after certification, 66 per cent of the coffee commercialised by the members of the certified cooperative was sold outside the association, to local traders. Given that side-selling was estimated to be approximately the same between treatment and control farmers in 2010, the estimates reported in Tables 5 and 6 for 2012 suggest that certification may have had a large and negative effect on the amount of coffee sold through the association by farmers. RA thus emphasises that the naïve comparisons in Table 1 and the PSM estimates in Table 4 may have generally underestimated the negative impact generated by the FT premium.

It is, however, important to note that even the impact estimates reported by our most reliable estimator (RA in Table 6) could still be somewhat biased. Residual bias in the

estimation can be due to unobserved differences between treatment and control groups, especially with regard to farmers' motivations. For instance, we cannot exclude that the association disclosed the intention to certify our treatment group (and not to certify our control group) well in advance. This could have contributed to motivate treatment farmers, explaining the larger amount of coffee sold by them through the association in 2010. Such a difference in the motivation of the two groups could have thus introduced an anticipation bias, resulting in lower side-selling within the treatment group before certification. Unfortunately, we cannot possibly control for such an anticipation bias, since we do not have data from before 2010. However, we can argue that anticipation bias could have contributed to an overestimation of side-selling differences between the two groups in 2010. As such, this argument supports the results reported in Table 6 in suggesting that no significant difference existed between the two groups before certification.

Furthermore, the impact estimates presented in Table 6 could be affected by self-selection bias, given that membership in the treatment group increased over time (from 50 members in 2010 to 67 in 2012), whereas the size of the control group diminished (from 80 to 48 members). According to the cooperative life cycle theory discussed in the introduction, membership growth can promote the rise of the free-rider problem and agency costs. As such, farmers' self-selection into the treatment group, and self-exclusion from the control group, could have induced an overestimation of the impact measured in 2012. It follows that the overall impact of FT could have been less negative than reported in Table 6, or even positive. Yet, self-selection bias would have to be significantly large to change the sign of the estimated impact. Although the large R^2 in Table 6 for 2012 suggests that our estimation may indeed be affected by significant missing variable bias, which could be attributed to changes and differences in membership between the two groups, previous theoretical and factual analyses suggest that it is unreasonable to expect a positive impact for FT certification. However, it is not unreasonable to conclude that the impact of the FT premium could be less negative than reported in Table 6, or even insignificant.

5. Conclusions and implications

This study attempts to streamline FairTrade's intended and unintended impact pathways in order to contribute to the definition of a general theory of change for its increasingly widespread certification scheme. In particular, FT certification is defined as an intervention involving two incentive mechanisms that aim to improve the income and welfare of coffee farmers organised in cooperatives: a guaranteed floor price and a yearly premium. Previous research demonstrated that when the market price for coffee is particularly low, the introduction of a floor price is intended to generate a price incentive for certified cooperatives to keep procuring and selling coffee, minimising income losses among member-farmers. However, individual price incentives tends to diminish over time, since certified cooperatives tend to witness rapid growth in membership as a result of high external pressure for inclusion. Rapid growth in membership promotes a 'crowding-in' effect in certified cooperatives, which dilutes the initial incentive and increases organisational inefficiency, inducing member farmers to sell their coffee outside the FT marketing channel into spot markets. This theory of change suggests that FT price incentives are expected to be beneficial only when directed at cooperatives that restrict farmers' entry or facilitate members' exit. This conclusion implies that, before introducing price incentives, FT ought to ensure that target cooperatives are able to enforce adequate membership rules. One could argue that this recommendation would facilitate FT

becoming an instrument to promote social exclusion and the consolidation of rural elites. However, certified cooperatives that do not confront the price incentives generated by FT, by minimising membership growth, may be more likely to degenerate and collapse than conventional cooperatives. It follows that in the name of social inclusion, FT certification may end up depriving farmers of a key market outlet and undermine their capacity to cooperate and embark on collective marketing, which is exactly the opposite of what FT is meant to do.

This study adds to these concerns and recommendations by looking into the other, and arguably most important, incentive mechanism associated with FT certification: a yearly premium to be invested in collective assets (warehouses, washing stations, offices and so forth) for the development of cooperative organisations and rural communities. Our empirical analysis suggests that the prospect of receiving a premium may induce farmers to maximise the amount of coffee commercialised through the FT marketing channel (given by the cooperative association and FT buyers), but this positive effect tends to decrease rapidly after certification and can eventually become negative over time. This empirical finding is supported by theoretical and contextual analyses, explaining why the introduction of a FT premium did not produce the intended impact on the Tanzanian cooperative association under study. On the one hand, the FT premium may have contributed to shifting of decision-making power from member-farmers to the technicians and managers hired by the association, given that only the latter had direct access to FT buyers and thus to the premium. On the other hand, the FT premium may have induced the association to engage in an excessively long-term and diversified investment strategy, adding costs but generating limited benefits for member-farmers' over the short term. This would explain why certified farmers' started to raise embezzlement allegations against their administrators and to sell increasing amounts of coffee through alternative channels, inducing the association to declare bankruptcy.

This theory of change suggests that the FT premium can be expected to add value to smallholders' coffee production only when certified organisations (in this case the co-operative association) recognise and enforce members' decisions and claim rights over equity investments. For example, this can be done by allowing member-farmers to buy and trade investment shares in such a way as to maximise the appreciability, transparency and accountability of investment decisions. Therefore, we conclude that both price and investment incentives associated with FT certification are expected to benefit farmers only if selectively targeted to well-designed cooperative organisations, characterised by clearly defined and enforceable rules for coordinating members' entry and exit and allocating their decision-making and claim rights. In the absence of properly designed organisations, FT incentives may lead to unintended impacts over time, which can undermine rural collective action and the development of value chains.

Acknowledgement

This study benefited from the logistical support and in-depth local knowledge of Eliudi Doto and Kaminjoge Mhamerd.

Notes

1. The information provided in this first paragraph was retrieved from FairTrade's official website (www.fairtrade.net).
2. The name of the association involved in this study is omitted in respect of confidentiality agreements.

3. The selected district is a major production site for Arabica coffee in Tanzania (no Robusta coffee is grown in this district). The topography of the district includes mountains, valleys and plateaus, ranging from 800 to 1900 m above sea level. The micro-climate of the district is described as temperate to cool tropical climate with a unimodal rainfall pattern (December–April). The average annual rainfall is 1224 mm, ranging from 1000 to 1600 mm. Coffee harvesting is done during the dry season, from August to November. The average minimum and maximum temperatures for the district are 19 and 23°C, respectively.
4. It is important to stress that the authors of this study did not have any influence on the sampling process. The 10 cooperatives considered in this study were selected by the association's officers, who stated they had no clearance from the central administration to disclose additional information and data regarding the other member-cooperatives operating in the district. As such, the representativeness of the sample used in this study at the district and association levels could not be established.

References

Abate, G. T., G. N. Francesconi, and K. Getnet. 2014. "Impact of Agricultural Cooperatives on Smallholders' Technical Efficiency: Empirical Evidence from Ethiopia." *Annals of Public and Cooperative Economics* 85: 2.

Arnould, E. J., A. Plastina, and D. Ball. 2009. "Does Fair Trade Deliver on its Core Value Proposition? Effects on Income, Educational Attainment, and Health in Three Countries." *Journal of Public Policy and Marketing* 28 (2): 186–201. doi:10.1509/jppm.28.2.186.

Bacon, C. 2005. "Confronting the Coffee Crisis: Can Fair Trade, Organic, and Specialty Coffees Reduce Small-Scale Farmer Vulnerability in Northern Nicaragua?" *World Development* 33 (3): 497–511. doi:10.1016/j.worlddev.2004.10.002.

Bacon, C. M., V. Méndez, M. E. Flores Gómez, D. Stuart, and S. R. Díaz Flores. 2008. "Are Sustainable Coffee Certifications Enough to Secure Farmer Livelihoods? The Millenium Development Goals and Nicaragua's Fair Trade Cooperatives." *Globalizations* 5 (2): 259–274. doi:10.1080/14747730802057688.

Balineau, G. 2013. "Disentangling the Effects of Fair Trade on the Quality of Malian Cotton." *World Development* 44: 241–255. doi:10.1016/j.worlddev.2012.12.005.

Becchetti, L., and M. Constantino. 2006. *Fair Trade on Marginalized Producers: An Impact Analysis on Kenyan Farmers*. Rome: University of Rome Working Paper CEIS 220/ ECINEQ2006.

Bernard, T., A. S. Taffesse, and E. Z. Gabre-Madhin. 2008. "Impact of Cooperatives on Smallholders' Commercialization Behavior: Evidence from Ethiopia." *Agricultural Economics* 39: 147–161. doi:10.1111/j.1574-0862.2008.00324.x.

Blackman, A., and J. Rivera, 2010. "The Evidence Base for Environmental and Socioeconomic Impacts of Sustainable Certification." Resources for the Future Discussion Paper. Washington, DC: Resources for the Future. http://rff.org/rff/documents/rff-dp-10-17.pdf

Calo, M., and T. A. Wise. 2005. *Revaluing Peasant Coffee Production: Organic and Fair Trade Markets in Mexico*. Medford, MA: Mimeo, Global Development and Environment Institute, Tufts University.

Carvalho, S., and H. White. 2004. "Theory-Based Evaluation: The Case of Social Funds." *American Journal of Evaluation* 25 (2): 141–160. doi:10.1177/109821400402500202.

Cook, M. L., and M. Chambers. 2007. "The Role of Agricultural Cooperatives in Global Netchains." Working Paper Presented at the Workshop on Small Farmer Organization and Markets in Developing Countries Organized by INRA-MOISA and Wageningen University, Montpellier, April 26–27.

De Hoop, T. J. 2011. *How Context Matters for Development Effectiveness: A Study into Social Norms and Heterogeneous Impacts*. Arnhem: Radboud University Nijmegen. ISBN 978-94-6191-148-3.

De Janvry, A., C. McIntosh, and E. Sadoulet. 2010. "Fair Trade and Free Entry: Generating Benefits in a Disequilibrium Market". USAID-BASIS Working Paper, Washington, DC.

Francesconi, G. N., and N. Heerink. 2011. "Ethiopian Agricultural Cooperatives in an Era of Global Commodity Exchange: Does Organizational Form Matter?" *Journal of African Economies* 20 (1): 153–177. doi:10.1093/jae/ejq036.

Godtland, E. M., E. Sadoulet, A. de Janvry, R. Murgai, and O. Ortiz. 2004. "The Impact of Farmer-Field-Schools on Knowledge and Productivity: A Study of Potato Farmers in the Peruvian Andes." *Economic Development and Cultural Change* 53: 63–92. doi:10.1086/423253.

Jaffee, D. 2007. *Brewing Justice: Fair Trade Coffee, Sustainability and Survival*. Berkeley: University of California Press.

Karlan, D. 2009. "Thoughts on Randomized Trials for Evaluation of Development: Presentation to the Cairo Evaluation Clinic." *Journal of Development Effectiveness* 1 (3): 237–242. doi:10.1080/19439340903134519.

Méndez, V. E., C. M. Bacon, M. Olson, S. Petchers, D. Herrador, C. Carranza, L. Trujillo, C. Guadarrama-Zugasti, A. Cordón, and A. Mendoza. 2010. "Effects of Fair Trade and Organic Certifications on Small-Scale Coffee Farmer Households in Central America and Mexico." *Renewable Agriculture and Food Systems* 25 (3): 236–251. doi:10.1017/S1742170510000268.

Milford, A. 2004. *Coffee, Co-Operatives and Competition: The Impact of Fair Trade*. Bergen: Mimeo, Chr. Michelsen Institute.

Pawson, R., and N. Tilley. 1997. *Realistic Evaluation*. London: Sage Publications.

Potts, J., J. van der Meer, and J. Daitchman. 2010. *The State of Sustainability Initiatives Review 2010: Sustainability and Transparency.* Mimeo, International Institute for Sustainable Development, Winnipeg, Manitoba, and, International Institute for Environment and Development, London.

Ravallion, M. 2009. "Evaluating Three Stylised Interventions." *Journal of Development Effectiveness* 1 (3): 227–236. doi:10.1080/19439340903114644.

Rogers, P. J. 2007. "Theory-Based Evaluation: Reflection Ten Years On." *Enduring Issues in Evaluation: The 20th Anniversary of the Collaboration between NDE and AEA. New Directions for Evaluation* 114: 63–67. doi:10.1002/ev.225.

Rogers, P. J. 2009. "Impact Evaluation Design: Matching Impact Evaluation Design to the Nature of the Intervention and the Purpose of the Evaluation." *Journal of Development Effectiveness* 1 (3): 217–226. doi:10.1080/19439340903114636.

Ronchi, L. 2002. *The Impact of Fair Trade on Producers and their Organisations: A Case Study with Coocafe in Costa Rica*. Sussex: Policy Research Unit. University of Sussex.

Ruben, R. 2008. *The Impact of Fair Trade*. Wageningen: Wageningen Academic Publishers.

Ruben, R., and R. Fort. 2012. "The Impact of Fair Trade Certification for Coffee Farmers in Peru." *World Development* 40 (3): 570–582. doi:10.1016/j.worlddev.2011.07.030.

Ruben, R., R. Fort, and G. Zúñiga-Arias. 2009. "Measuring the Impact of Fair Trade on Development." *Development in Practice* 19 (6): 777–788. doi:10.1080/09614520903027049.

White, H. 2009. "Theory-Based Impact Evaluation: Principles and Practice." *Journal of Development Effectiveness* 1 (3): 271–284. doi:10.1080/19439340903114628.

White, H. 2011. "Achieving High-Quality Impact Evaluation Design through Mixed Methods: The Case of Infrastructure." *Journal of Development Effectiveness* 3 (1): 131–144. doi:10.1080/19439342.2010.547588.