Preliminary draft. Please do not quote or cite without authors' permission.

DOES ECO-CERTIFICATION HAVE ENVIRONMENTAL BENEFITS? ORGANIC COFFEE IN COLOMBIA

Marcela Ibañez
Courant Research Centre for Poverty, Equity and Growth
Wilhelm Weber Str. 2
D-37073 Göttingen, Germany.
+49 (0) 551-39-10677
mibanez@uni-goettingen.de

Allen Blackman
Resources for the Future
Washington, DC
blackman@rff.org
+001 (202) 328-5073
and
Environment for Development Center
for Central America (EfD-CA)

Draft: October 14, 2011

Acknowledgments: Funding for this research was provided by the Swedish Foundation for Strategic Environmental Research (MISTRA) through its Environment and Trade in a World of Interdepenence (ENTWINED) program, the Latin American Environmental Economics Program (LACEEP), the Adlerbert Research Foundation, and the Swedish Research Council Formas. We are grateful to Gerhard Riener and LACEEP participants for helpful comments, and to Tomas David Vargas for assistance with field research.

Preliminary draft. Please do not quote or cite without authors' permission.

DOES ECO-CERTIFICATION HAVE ENVIRONMENTAL BENEFITS? ORGANIC COFFEE IN COLOMBIA

Abstract

Eco-certification of coffee, timber, and other high-value agricultural commodities is increasingly widespread. In principle, it can improve commodity producers' environmental performance, even in countries where state regulation is weak. However, evidence needed to evaluate this hypothesis is virtually nonexistent. To help fill this gap, we use detailed farm-level data to analyze the environmental impacts of organic coffee certification in southeast Colombia. We use propensity score matching to control for self-selection bias. We find that organic certification improves coffee growers' environmental performance. It significantly reduces chemical input use and increases adoption of some environmentally friendly management practices.

Keywords: Certification, coffee, Colombia, propensity score matching.

JEL codes: Q13, Q20, 013, Q56

1. INTRODUCTION

Initiatives certifying that goods and services have been produced in an environmentally friendly manner are increasingly popular. In the past two decades alone, more than 300 have been launched in a wide range of countries and economic sectors (Ecolabel Index 2011). Eco-certification of agricultural commodities is particularly widespread. Today, 10% of the timber, 7% of the coffee, and 12% of the wild fish products traded in international markets are certified as having been sustainably produced by organizations such as the Forest Stewardship Council, Rainforest Alliance, and International Federation of Organic Agriculture Movements (Eilperin 2010).

According to proponents, commodity eco-certification schemes like these have the potential to improve producers' environmental performance (Giovannucci and Ponte 2005; Rice and Ward 1996). In theory, they can do this by enabling the consumer to differentiate among commodities based on their environmental attributes. This improved information facilitates price premiums for certified commodities, and these premiums, in turn, create financial incentives for producers to meet certification standards.

If that logic holds, certification may help address pressing environmental problems associated with agricultural commodities in developing countries. Growing and processing bananas, cocoa, coffee, timber, and other high-value agricultural products in poor countries often entails deforestation, soil erosion, and agrochemical pollution. These problems are difficult to tackle using conventional command-and-control regulation because producers are often small, numerous, and geographically dispersed while regulatory institutions are undermanned and underfunded (Wehrmeyer and Mulugetta 1999). Certification schemes have the potential to sidestep these constraints by creating a private sector system of economic incentives, monitoring, and enforcement.

Yet certification programs that aim to improve commodity producers' environmental performance also faces important challenges. They must use standards stringent enough and monitoring and enforcement strict enough to ensure that poorly performing producers are excluded. In addition, they must offer price premiums high enough to offset the costs of certification. Even if these two challenges are met, certification schemes still can be undermined by selection effects. Commodity producers already meeting certification standards have strong incentives to select into certification programs: they need not make additional investments in environmental management to pass muster and can obtain price premiums and other benefits. But certification programs that mainly attract such producers will have limited effects on producer behavior and few environmental benefits.

Although a growing academic literature examines commodity certification, we still know little about whether it actually affects producers' environmental performance. As discussed below, few studies evaluate the environmental impacts of certification, and many of those that do rely on problematic methods that bias their results. To identify certification impacts, an evaluation must construct a reasonable counterfactual outcome —that is, an estimate of what environmental outcomes for certified entities would have been had they not been certified. However, most evaluations use problematic counterfactual outcomes, either certified producers' precertification outcomes or uncertified producers' outcomes. In the first case, results are biased whenever outcomes change during the study period because of factors unrelated to certification (including changes in commodity prices, input prices, weather conditions, and technology, all of which are common). In the second case, results are biased whenever commodity producers already meeting certification standards select into certification.

A variety of *ex post* statistical methods are available to overcome these problems, including propensity score matching and instrumental variables (Ferraro 2009; Frondel and Schmidt 2005). A recent comprehensive review of the empirical studies of certification of agricultural commodities and tourism operations found only two that use such methods to identify environmental impacts (Blackman and Rivera In Press). Neither of these studies examined certification for one of the most prominent high-value agricultural commodities, coffee.

As a first step toward filling that gap, this paper presents an evaluation of the environmental impacts of organic coffee certification in southeast Colombia. We use original farm-level data and rely on propensity score matching to control for selection bias. We find that certification does have an environmental benefit. It significantly reduces use of two of the three chemical inputs for which we have data (chemical fertilizers and insecticides) and spurs adoption of one of the two environmentally friendly management practices for which we have data (organic fertilizer).

The remainder of this paper is organized as follows. The second section briefly reviews the literature evaluating the environmental effects of coffee certification. The third section presents background on coffee production, organic certification, and our study area. The fourth section discusses our empirical strategy and data. The fifth section presents our results, and the last section discusses their policy implications.

2. LITERATURE

Rigorous evaluations of the environmental impacts of certification are rare, and those that have been conducted often fail to find significant effects. Blackman and Rivera (In Press) reviewed 213 published studies of agricultural commodity and tourism certification and identified only two that both constructed a reasonable counterfactual and focused on environmental (versus socioeconomic) impacts: Rivera and de Leon (2004) and Rivera et al. (2006). Both studies conclude that the environmental effects of certification are negligible. They analyze the Sustainable Slopes Program, a voluntary certification program established by the U.S. ski areas' industry association. Using a Heckman procedure to control for self-selection bias, they compare third-party environmental performance ratings of certified and uncertified ski areas. They find that in the Sustainable Slopes Program's early years, uncertified areas actually had better environmental performance than certified areas, and subsequently, they had equivalent but not superior levels.

As for farm-level quantitative studies of coffee eco-certification, to our knowledge, all existing published studies that construct a reasonable counterfactual focus on socioeconomic impacts.[1,2] Three less rigorous studies analyze environmental impacts by comparing environmental outcomes for certified farms and unmatched uncertified farms and reach mixed conclusions despite the fact that failure to control for self-

[1] See Blackman and Rivera (In Press). Only two of these studies—Arnould et al. (2009) and Bolwig et al. (2009)—find that certification has significant socioeconomic benefits, but in both cases the effects are weak or idiosyncratic: Arnould et al. (2009) find that although certification generates a price premium, it is not consistently correlated with socioeconomic indicators, and Bolwig et al. (2009) argue that in their case, these socioeconomic benefits are mainly due to a design anomaly of the certification scheme.

[2] Three purely qualitative case studies also evaluate the environmental effects of eco-certification: Bray et al. (2002), Utting-Chamorro (2005), and Utting (2008).

selection bias typically generates overly optimistic results. Jaffee et al. (2008) compare social and environmental outcomes for Fair Trade and organic certified growers and unmatched uncertified growers in Oaxaca, Mexico, and find that certified growers adopt more soil conservation practices. Similarly, Martínez-Torres (2008) compares ecological indicators (soil erosion, number of shade species, and leaf litter depth) for certified organic and unmatched uncertified growers in Chiapas, Mexico, and finds that organic growers perform better. However, Philpott et al. (2007) compare ecological indicators for Fair Trade/organic certified growers and unmatched uncertified growers in Chiapas, Mexico, and find no differences between the two subsamples.

To our knowledge, the only existing rigorous quantitative evaluation of the environmental effects of coffee certification is as yet unpublished. Using the same propensity score matching methods as the present paper, Blackman and Naranjo (2010) examine the effect of organic coffee certification on the adoption of various agricultural practices in Costa Rica, a country that has led Latin America in the intensive use of agrochemicals for coffee growing (Rice and Ward 1996). They find that organic certification significantly reduces chemical input use and increases the adoption of some environmentally friendly management practices.

3. BACKGROUND

3.1. Coffee in Colombia

Coffee is Colombia's most important agricultural commodity. In 2009, it was planted on farms spanning 900,000 hectares that produced more than 8 million quintals (100-pound bags) of coffee beans, employed 30% of the rural labor force, and generated more than US$1.7 billion in export revenues. The majority of Colombian coffee farms are smaller than 5 hectares (ICO 2010; FNC 2011).

Coffee growing in Colombia has serious environmental consequences that at least partly offset those economic benefits. Traditionally, Colombian coffee, like most coffee in Latin America, was grown alongside shade trees in an agroforestry system that predated the development of agrochemicals and therefore did not rely on them. However, since the 1970s, more than two-thirds of the country's coffee acreage has been converted to a high-yielding "technified" monocrop in which coffee is grown with minimal shade cover and intensive application of agrochemicals (Rice and Ward 1996; Guhl 2008). Today almost half of Colombian coffee is grown with no shade at all (FNC 2011). The switch to technified coffee has hastened soil erosion and biodiversity loss and contributed to such off-site negative externalities as the contamination and sedimentation of surface and groundwater (Rice and Ward 1996; Pelupessy 2003).

3.2. Organic coffee certification

Organic agriculture certification requires producers to adhere to five broad production principles (Van der Vossen 2005; IFOAM 2010):

- use of composted organic matter instead of chemical fertilizers to maintain soil quality;
- use of natural methods for controlling disease, pests, and weeds instead of synthetic insecticides and herbicides;

- use of soil conservation practices, including contour planting, terracing, planting cover crops, mulching, and planting shade trees;
- minimal use of fossil fuels in the production process; and
- minimal pollution during postharvest handling.

Several international organic certifying bodies, the largest of which is the International Federation of Organic Agriculture Movements (IFOAM), formulate basic organic standards for various commodities. These large organizations accredit smaller national ones, which in turn certify individual producers (not cooperatives) and conduct follow-up monitoring. Organic certifications require growers to complete a transition period of two to three years during which they must discontinue use of chemical inputs and adopt various conservation and pollution prevention practices. Certified producers are monitored at least once a year to ensure they continue to meet organic standards.

From coffee growers' perspective, organic certification has both benefits and costs (Giovanuci and Ponte 2005; Van der Vossen 2005; Calo and Wise 2005). The main benefit is the price premium, which is set in international markets and averages 10% to 20%, depending on coffee quality. In addition, organic production reduces the costs of purchased inputs for growers who formerly depended on chemical inputs. It can also improve coffee quality. On the cost side, organic production typically increases labor costs and reduces yields for growers who formerly depended on chemical inputs. In addition, transaction costs—for initial certification, for subsequent annual monitoring and reporting—are significant. Annual costs can easily amount to 5% of sales and are often borne by the grower. Note that the transition period implies that the grower must pay these costs for two to three years without the principal benefit of certification—a price premium.

3.3. Study area and period

Organic certification of coffee in Colombia began in the 1980s. Today, exports of Colombian organic coffee exceed 95,000 quintals per year, roughly 16% of total global exports (ICO 2009). We examine certification in Cauca, a department (state) in the southeastern part of the country. Cauca is home to 16% of Colombia's coffee farms and 8% of its coffee acreage (FNC 2011). The state is one of Colombia's leading centers of organic coffee production. We focus on five municipalities in Cauca with particularly high rates of organic certification: Inza, Cajibio, Tambo, Timbio, and La Sierra. In all, 331 growers harvesting 587 hectares were certified organic in these municipalities. In addition, 162 farmers harvesting 211 hectares were in transition to certification. The vast majority of these farmers were certified in the decade prior to our survey. Very few were certified prior to 1998. Three organizations certified the organic growers in our study area: Bio-Latina, IMO-CONTROLS, and Organic Crop Improvement Association. Each of these organizations holds multiple accreditations for multiple markets.[3]

[3] Bio-Latina is accredited by the German Deutsches Akkreditierungssystem Prüfwesen (DAP), the U.S. Department of Agriculture National Organic Program (UNDA NOP), Japan Agricultural Standards (JAS), and the Conseil des Appellations Agroalimentaires du Québec (CARTV). IMO-CONTROLS is accredited by the Swiss Accreditation Service (SAS), USDA NOP, and JAS. The Organic Crop Improvement Association is accredited by USDA NOP, the International Organic Accreditation Service, the European Union Equivalent Standards, JAS, and CARTV.

4. EMPIRICAL STRATEGY AND DATA

4.1. Propensity score matching

Our analysis of organic certification's impact on environmental performance confronts the usual program evaluation challenge (Rubin 1974; Holland 1986). Ideally, the impact of a program would be measured by comparing the outcome of interest for each agent both with and without program participation. However, we never actually observe both outcomes. In practice, therefore, a program's impact is typically measured by comparing the average outcome for participants and for a control group of nonparticipants—with the latter average serving as the counterfactual. But as discussed in the introduction, this approach can be undermined if certain types of participants who tend to have certain outcomes select into the program. For example, in our case, small, undercapitalized farms that cannot afford to use chemical inputs may self-select into organic certification because the net benefits are high: they can meet organic standards and obtain price premiums without having to discontinue chemical input. Or farms on steeply sloped land that already use soil conservation measures may self-select into certification because they do not have to adopt them to meet organic standards. An evaluation that failed to control for such selection would conflate the effects of certification on outcomes with the effects of preexisting differences between certified and uncertified farms.

To address this selection problem, we use a matching estimator. That is, following Rosenbaum and Rubin (1983) and more recently Blackman et al. (2010), List et al. (2003), and Dehejia and Wahba (2002), we construct a matched control sample of uncertified farms that are very similar to the certified farms in terms of observable characteristics. We measure program impact as the average treatment effect on the treated (ATT)—the difference between the percentage of certified farms that use a management practice and the percentage of matched uncertified farms that use it.

This approach depends on two identifying assumptions. The first assumption, "ignorability" or "conditional independence," is that conditional only on agents' observed characteristics, the participation decision is ignorable for purposes of measuring outcomes. That is, we are able to observe and control for all variables that simultaneously affect the participation decision and the outcome variables. This first assumption is untestable. The second assumption, "common support" or "overlap," is that the distribution of observed characteristics for nonparticipants is similar to that for participants, such that agents with similar characteristics have a positive probability of being participants and of being nonparticipants.

Creating a large set of matched pairs of farms with the exact same observed characteristics is challenging when, as in our case, these characteristics are numerous. However, Rosenbaum and Rubin (1983) demonstrate that we need to match farms only on the basis of their propensity score—that is, their likelihood of certification as predicted by a regression model—which amounts to an index of farm and grower characteristics weighted by their importance in predicting certification. The propensity score method collapses the difficult problem of matching all observable characteristics to a much simpler one of matching a single summary variable.

Various methods are available to match participants and nonparticipants based on propensity scores (Caliendo and Kopeining 2008; Morgan and Harding 2006). To ensure robustness, we report results from five: (i) nearest neighbor 1-to-1 matching, wherein each certified farm is matched to the uncertified farm with the closest propensity score; (ii) nearest neighbor 1-to-4 matching, wherein each certified farm is matched to the four uncertified farms with the closest propensity scores and the counterfactual outcome is the average across these four; (iii) nearest neighbor 1-to-8 matching; (iv) nearest neighbor 1-to-16 matching; and (v) kernel matching, wherein a weighted average of all uncertified farms is used to construct the counterfactual outcome. For all five models we enforce a common support and allow matching with replacement.

Calculating standard errors for ATT estimated using propensity score matching is not straightforward because these errors should, in principle, account for the fact that propensity scores are estimated and for the imputation of the common support (Heckman et al. 1998). Therefore, following Dehijia and Whaba (2002) and others, we bootstrap standard errors (using 1,000 replications).

4.2. Data

The data used for our analysis come from an original 2007 of survey of 379 coffee growers in five municipalities in the department of Cauca. As noted above, we selected this department and these particular municipalities because they are home to a relatively large concentration of organic farms. We randomly selected our survey sample from a lists of coffee farmers. The survey questionnaire was administered on-site by trained enumerators in face-to-face sessions that typically lasted 40 minutes. The survey solicited information on both grower characteristics (e.g., age, education) and farm characteristics (e.g., eco-certification, hectares cultivated, types of inputs used) for two years, 2007 and 1997.[4]

We asked about 1997 grower and farm characteristics to control for possible endogeneity in our matching analysis. All but seven of the organic growers in our survey sample were organic certified after 2000. Hence, for all but these seven growers, 1997 grower and farm characteristics predate—and are therefore exogenous to—the decision to obtain organic certification. As noted below, we drop these seven growers from our regression sample.

Among the production practices on which we have data are six that are monitored by organic certifiers—our outcome measures. We divide them into four "negative" practices that typically must be discontinued for organic certification, and two "positive" practices that must be adopted. The negative practices are

- the use of chemical fertilizers;
- the use of chemical insecticides;
- the use chemical herbicides; and
- disposing of sewage in open fields.

The positive practices are

- the use of organic fertilizer; and
- the use of shade cover for coffee trees.

[4] Coffee prices in 1997 were the highest in the last two decades, reaching 1.31 US per lb. This price spike may have induced changes in agricultural practices, although we expect that transformation of productive system is slow.

We constructed our matching sample as follows. Starting with the 379 randomly selected growers whom we surveyed, we eliminated 94 farms that could not provide data for 1997 because they did not yet exist or were not yet harvesting coffee in 1997. In addition, we eliminated 35 growers because they had obtained a certification other than (or in addition to) organic.[5] We dropped these growers so that we could disentangle the effect of organic certification from other types of certification. We also dropped four growers who were certified or transitioning to certification prior to 1998 to control for the endogeneity problem noted above. Finally, we dropped five growers who were certified after 1997 but who gave up their certification prior to 2007. We dropped these growers because they would be counted as uncertified in our matching sample but may have had outcomes or characteristics that were affected by having been certified. Having dropped these 147 growers, our regression sample comprises 232 growers, all of whom had been producing coffee since at least 1997, none of whom were organic certified in 1997, and none of whom had ever obtained an eco-certification other than organic. Fifty-six of these 232 growers were organic certified in 2007 and 176 had never been eco-certified at any time.

4.3. Variables

Table 1 lists, defines, and presents summary statistics for the variables used in our matching analysis, including both outcome variables and grower and farm characteristics. For the seven dichotomous outcome variables listed above, mean use rates for the negative practices range from a low of 12% for herbicides to a high of 38% for chemical

[5] 28 farms were Rainforest Alliance certified and seven were Fair Trade certified.

fertilizers. The mean use rates for the positive practices range from 51% for organic fertilizer to 95% for shade cover. Hence, the proportion of growers using practices consistent with organic certification—particularly, eschewing agrochemicals—is quite high. Except in the case of organic fertilizer use, the majority of the uncertified growers in our sample used such practices. Hence, it is reasonable to expect that a disproportionate share of *de facto* organic growers—that is, those already meeting organic standards—self-selected into organic certification, and thus certification only had limited effects on the environmental performance of the average grower in our sample. Our empirical analysis aims at determining whether that in fact was the case.

[Insert Table 1 here]

To match certified and uncertified farms, we used propensity scores generated by regressing an organic certification dummy onto a rich set of grower and farm characteristics. The grower characteristics are FEMALE, a dichotomous dummy equal to one for female growers; AGE, the age of the grower in 2007; EDUCATION, the highest grade completed in 2007; FAMILYSIZE, the size of the family in 2007; and MEMBER, a dichotomous dummy equal to one if the grower was a member of a coffee committee affiliated with the Federación Nacional de Cafeteros de Colombia in 1997.[6]

The farm characteristics, all of which correspond to the year 1997, are OWN_FARM, a dichotomous dummy equal to one if the grower owns (versus rents or leases) the farm; NO_TREES, the number of coffee trees on the farm; FARM_SIZE, the total extent of the farm in hectares; NO_LOTS, the number of geographically distinct lots

[6] A limitation of the survey is that it did not ask about family size for 1997.

on the farm; ORGANIC_MAT an estimate of the number of kilograms of manure produced by all animals on the farm;[7] CAPTIAL_INDEX, a count of the number of common capital items owned (depulper, mill, silo, fumigator, motor, other); PBORBON, PCOLOMBIA, PTIPICA, PCASTILLO, and PVARIEDAD, the percentages of the farm's coffee trees that are the most common varieties in the study area (Borbon, Colombia, Tipica, and Castillo); PCLASICO, PESPECIAL, and PCALIDAD, the percentages of coffee sold in the thee quality grades for Colombian coffee (Clasico, Especial, and Calidad); BUYER_INT, a dichotomous dummy variable equal to one if the grower sells to an intermediary (versus a cooperative, association, or exporter); TRANS_VEH, a dichotomous dummy equal to one if the grower transports the coffee by vehicle (versus by animal or on foot); PHH_F_WORKER and PHH_OF_WORKER, the percentages of household members who work on and off the farm; and MUNICIP_INZ, MUNICIP_CAJ, MUNICIP_TAM, MUNICIP_TIM, and MUNICIP_LAS, dichotomous dummy variables indicating in which of the five study municipalities the farm is located.

5. RESULTS

5.1. Propensity scores and balance tests

Table 2 presents the results from the probit regression (of organic certification on grower and farm characteristics) used to generate propensity scores. The results indicate that compared with average farms in our sample, certified farms tend to consist of fewer distinct lots, to have coffee trees of the Borbon variety, to sell Calidad-grade coffee, and to be located in either the Cajibo or Timbo municipality.

---

[7] These estimates are based on the simple linear model in Muñoz (2001) that relates the number and type of farm animals to the quantity of manure.

[Insert Table 2 here]

Having generated propensity scores and used them to match certified and uncertified farms, we performed balance tests for the five matching estimators (Table 3). The nearest neighbor 1-8, 1-16, and kernel estimators balanced all 28 covariates (i.e., generated a statistically insignificant difference in means for certified and matched uncertified farms for all 28 covariates). The nearest neighbor 1-1 and 1-4 estimators balanced all but one covariate (PPASILLA). Table 3 reports median standardized bias—Rosenbaum and Rubin's (1983) balance statistic—across all covariates for each matching estimator.[8] The highest median standardized bias is 9.513 for the nearest neighbor 1-1 estimator, and the lowest is 4.984 for the kernel estimator. Although a clear threshold for acceptable median standardized bias does not exist, according to Caliendo and Kopeining (2008), a statistic below 3% to 5% is generally viewed as sufficient. These overall encouraging balance statistics are likely due to the fact that even though our probit selection model has 28 covariates, our sample includes more than three uncertified farms for each certified farm. As a result, we are able to find fairly close matches for each certified farm.

[Insert Table 3 here]

---

[8] Standardized bias is the difference of the sample means in the certified and uncertified subsamples as a percentage of the square root of the average of sample variances in both groups.

5.2. Average treatment effect on the treated

Table 4 presents results from the five matching estimators for the negative practices.[9] The results indicate that certification significantly reduces use of three of the four negative practices—chemical fertilizer, chemical insecticides, and sewage disposal. For each of these three negative practices, ATT is negative and significant for at least four of the five matching estimators. For chemical fertilizers, it ranges from 32 to 38 percentage points; that is, the rate of chemical fertilizer use is 32 to 38 percentage points lower among certified growers than among matched uncertified growers representing the counterfactual. For chemical insecticides, ATT ranges from 12 to 14 percentage points, and for herbicides, it ranges from 1 to 6 percentage points.

[Insert Table 4 here]

Table 5 presents results from the five matching estimators for the two positive practices.[10] The results provide strong evidence that organic certification increases use of organic fertilizer. For this practice, ATT is positive and significant for all five matching estimators, and the magnitude of the effect is substantial, ranging from 44 to 60 percentage points. For shade, ATT is not significant for any of the five matching

[9] Note that the mean of the outcome variables for certified farmers is positive, albeit small, implying that a handful of the 56 certified growers in our sample used chemical inputs or disposed of their sewage in their fields in 2007. Organic standards allow the occasional use of chemical inputs when deemed necessary and preauthorized by a certifying agency inspector.

[10] Note that the mean of the outcome variables for certified farmers is less than 1, implying that some of the certified growers in our sample had not adopted the two environmental management practices we consider. Organic inspectors relax certification requirements in certain cases. In general, inspectors enforce prohibitions against negative practices (use of agrochemicals) more stringently than they require the positive ones (soil conservation, etc.).

estimators. This is not surprising, given that 95% of the matching sample, including 94% of the uncertified growers, uses shade.

[Insert Table 5 here]

5.3. Sensitivity analysis

Might endogeneity drive our results? As noted above, the effectiveness of our matching estimators in controlling for selection bias depends on the untestable identifying assumption that we are able to observe confounding variables that simultaneously affect growers' decisions to obtain organic certification and to use (or not use) the production practices that serve as our outcome variables. That is, we essentially assume endogeneity is not a problem. We calculate Rosenbaum bounds to check the sensitivity of our results to the failure of this assumption (Rosenbaum 2002; Aakvik 2001).[11] Rosenbaum bounds indicate how strongly unobserved confounding factors would need to influence growers' decisions to obtain organic certification in order to undermine the matching result. To be more specific, the Rosenbaum procedure generates a probability value for Wilcoxon sign-rank statistic for a series of values of Γ, an index of the strength of the influence unobserved confounding factors have on the selection process. $\Gamma = 1$ implies that such factors have no influence, such that pairs of growers matched on observables do not differ in their odds of obtaining organic certification; $\Gamma = 2$ implies that matched pairs could differ in their odds of certification by as much as a factor of two because of unobserved confounding factors; and so forth. The probability

[11] An example of an unobserved confounder might be environmental consciousness or managerial skill. Each could cause growers to select into organic certification and—independent of certification—to use fewer negative practices and more positive ones.

value on the Wilcoxon sign-rank statistic is a test of the null hypothesis of a zero ATT given unobserved confounding variables that have an effect given by Γ. So, for example, a probability value of 0.01 and a Γ of 1.2 indicate that ATT would still be significant at the 1% level even if matched pairs differed in their odds of certification by a factor of 1.2 because of unobserved confounding factors.

We calculate Γ*, the critical value of Γ at which ATT is no longer significant at the 10% level in each case—that is, for each combination of production practice and matching estimator—where ATT is significant (Tables 4 and 5, last column). Among the negative practices, Γ* is at least 2.2 for both insecticides and chemical fertilizers, and in most cases, it is considerably larger. For chemical fertilizers it is at least 3.4 for all five matching estimators, and for insecticides it is at least 6.6 for four of the five estimators. In the case of sewage, however, Γ* tends to be much lower, ranging from 1.4 to 2.4. Among the positive practices, Γ* is at least 5.6 for organic fertilizer. Hence, our sensitivity tests suggest that unobserved confounders would need to be quite strong to undermine our results in the case of all the practices for which ATT is significant except sewage disposal. In other words, endogeneity is unlikely to drive our results.

6. CONCLUSION

We have used detailed original survey data on 232 coffee farms in southeast Colombia to identify the environmental impacts of organic coffee certification. We have used propensity score matching to control for self-selection bias. Our findings suggest that certification significantly reduces use of two of the three chemical inputs for which we have data (chemical fertilizers and insecticides) and increases adoption of one of the two environmentally friendly management practices for which we have data (organic fertilizer). Hence, overall, we find that organic certification improves coffee growers' environmental performance.

Our findings contrast with those from the only two methodologically rigorous published studies of certification environmental impacts—Rivera and de Leon (2004) and Rivera et al. (2006)—both of which find that eco-certification has no causal effects. They also contrast with findings from the one of the three less rigorous published studies of coffee certification—Philpott et al. (2007). What might explain these differences? First, we have examined a certification scheme that has relatively well defined, stringent standards enforced at the individual farm level by independent third-party monitors. The Sustainable Slopes Program examined by Rivera and de Leon (2004) and Rivera et al. (2006) has relatively lax standards enforced by a trade association, and the Fair Trade certification programs analyzed by Philpott et al. (2007) are at the cooperative level rather than the farm level. Second, our study has looked at the impact of certification on various management practices, not on direct measures of ecological impacts like bird diversity, the focus of Philpott et al. (2007). Presumably, certification can alter management practices more easily than it can generate changes in ecological indicators.

It is interesting to compare our findings with those of Blackman and Naranjo (2010), which to our knowledge is the only other rigorous study of the environmental effects of coffee certification, and which also finds that organic certification has environmental benefits. The authors attribute this effect partly to the fact that in Costa Rica, coffee is highly technified, so most farms must change their management practices to obtain organic certification. Thus, opportunities for certification impacts to be diluted by self-selection—that is, for growers already meeting organic standards to obtain certification—are relatively limited. In Cauca, Colombia, by contrast, coffee is less technified. As noted above, except in the case of organic fertilizer use, the lion's share of the uncertified growers in our sample use practices consistent with organic certification. Hence, our study suggests that the effectiveness of organic certification in spurring environmental benefits has less to do with the preexisting characteristics of coffee growers than with the design characteristics of the certification program.

What are the policy implications of our findings? They suggest that commodity certification schemes that require adherence to well-defined stringent standards and are enforced at the individual farm level by independent third-party monitors can have significant environmental benefits, even in areas where self-selection threatens to dilute these benefits. That said, certification schemes meeting these criteria may have an important disadvantage: they are likely to entail significant costs for producers. Absent high price premiums or other benefits from certification, these costs will discourage certification. As a caveat, we hasten to note that our study is among the first to examine these issues using quantitative methods that should control for sample selection. Much more evidence from other study sites is needed before we can draw general conclusions about whether and under what circumstances eco-certification, or even just organic certification, has environmental benefits.

REFERENCES

Aakvik, A. 2001. Bounding a matching estimator: The case of a Norwegian training program. *Oxford Bulletin of Economics and Statistics* 63(1): 115–43.

Arnould, E., 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.

Blackman, A. and M. Naranjo. Does ecocertification have environmental benefits? Organic Coffee in Costa Rica. Resources for the Future Discussion Paper 10-25.

Blackman, A. and J. Rivera. In Press. Producer-level Benefits of Sustainability Certification. *Conservation Biology*.

Blackman, A., B. Lahiri, B. Pizer, M. Rivera Planter, and C. Muñoz Piña. 2010. Voluntary regulation in developing countries: Mexico's Clean Industry Program. *Journal of Environmental Economics and Management* 60(3): 182–92.

Bolwig, S., P. Gibson, and S. Jones. 2009. The economics of smallholder organic contract farming in tropical Africa. *World Development* 37(6): 1094–104.

Bray, D., J. Sanchez, and E. Murphy. 2002. Social dimensions of organic coffee production in Mexico: Lessons for eco-labeling initiatives. *Society and Natural Resources* 15**:** 429-446.

Caliendo, M., and S. Kopeining. 2008. Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys* 32: 31–72.

Calo, M., and T.A. Wise. 2005. Revaluing peasant coffee production: Organic and Fair Trade markets in Mexico. Global Development and Environment Institute, Tufts University. October.

Dehejia, R., and S. Wahba. 2002. Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics* 84(1): 151.

Ecolabel Index (2011). www.ecolabelindex.org.

Eilperin, J. 2010. Environmental certification becoming increasingly crowded and contested field. *Washington Post* 3 May: A4.

Ferraro, P. 2009. Counterfactual thinking and impact evaluation in environmental policy. In M. Birnbaum and P. Mickwitz (eds.), *Environmental Program and Policy Evaluation: New Directions for Evaluation* 122: 75–84.

Federación Nacional de Cafeteros de Colombia (FNC). 2011. Informacion Económica Cafetera. Estadisticas Históricas.

Frondel, M., and C. Schmidt. 2005. Evaluating environmental programs: The perspective of modern evaluation research. *Ecological Economics* 55(4): 515–26.

Giovannucci, D., and S. Ponte. 2005. Standards as a new form of social contract? Sustainability initiatives in the coffee industry. *Food Policy* 30: 284–301.

Guhl, A. 2008. Coffee production intensification and landscape change in Colombia. In A. Millington and W. Jepson (eds.) *Land Change Science in the Tropics: Changing Agricultural Landscapes.* Springer Science: New York. 93-116.

Heckman, J., H. Ichimura, J. Smith, and P. Todd. 1998. Characterizing selection bias using experimental data. *Econometrica* 66: 1017–98.

Holland, P. 1986. Statistics and causal inference. *Journal of the American Statistical Association* 99(467): 854–66.

International Coffee Organization (ICO) 2011. Colombia: Data for crop/calendar year commencing 2009. Available at http://www.ico.org/countries/colombia.pdf (accessed October 7, 2011).

International Coffee Organization (ICO) 2009. Organic coffee export statistics, coffee year 2007/2008. Available at http://www.ico.org/documents/wp-statistics-140e-organic.pdf (accessed October 11, 2011).

International Federation of Organic Agriculture Movements (IFOAM). 2010. Principles of organic agriculture. Available at http://www.ifoam.org/about_ifoam/principles/index.html (accessed January 13, 2009).

Jaffee, D. 2008. Better, but not great: the social and environmental benefits and limitations of Fair Trade for indigenous coffee producers in Oaxaca, Mexico. Pages 196-222 in R. Ruben (ed.) *The Impact of Fair Trade.* Wageningen Academic Publishers, The Netherlands.

List, J., D. Millimet, P. Fredriksson, and W. McHone. 2003. Effects of environmental regulations on manufacturing plant births: Evidence from a propensity score matching estimator. *Review of Economics and Statistics* 85(4): 944–52.

Martínez-Torres, M. E. 2008. The benefits and sustainability of organic farming by peasant coffee farmers in Chiapas, Mexico. Pages 99-126 in C. Bacon, V. Méndez, S. Gliessman, D. Goodman, and J. Fox (eds.) *Confronting the Coffee Crisis: Fair Trade, Sustainable Livelihoods, and Ecosystems in Mexico and Central America.* MIT Press, Cambridge, Massachusetts.

Morgan, S., and D. Harding. 2006. Matching estimators of causal effects: Prospects and pitfalls in theory and practice. *Sociological Methods and Research* 35(1): 3–60.

Muñoz, Moreno. 2001. Potencial de tres comunidades para adoptar caficultura orgánica. *Cenicafe* 54(4): 289-302

Pelupessy, W. 2003. Environmental issues in the production of beverages: Global coffee chain. In B. Mattsson and U. Sonesson (eds.) *Environmentally Friendly Food Processing*. Woodhead Publishing Limited: Boca Raton, FL. 95-113.

Philpott, S., P. Bichier, R. Rice, and R. Greenberg. 2007. Field testing ecological and economic benefits of coffee certification programs. *Conservation Biology* 21(4): 975–85.

Rice, R., and J. Ward. 1996. Coffee, conservation and commerce in the western hemisphere: How individuals and institutions can promote ecologically sound farming and forest management in northern Latin America. Washington, DC: Smithsonian Migratory Bird Center.

Rivera, J., and de Leon, P. 2004. Is greener whiter? The Sustainable Slopes Program and the voluntary environmental performance of western ski areas. *Policy Studies Journal* 32(3): 417–37.

Rivera, J., P. de Leon, and C. Koerber. 2006. Is greener whiter yet? The Sustainable Slopes Program after five years. *Policy Studies Journal* 34(2): 195–224.

Rosenbaum, P. 2002. *Observational Studies*, 2nd ed. New York: Springer-Verlag.

Rosenbaum, P., and D. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70: 41–55.

Rubin, D.B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66: 688–701.

Soto, G. 2009. Research Professor, Agroecology, Tropical Agricultural Research and Higher Education Center (CATIE), Turrialba, Costa Rica. Personal communication with authors. August 6.

Van der Vossen, H. 2005. A critical analysis of the agronomic and economic sustainability of organic coffee production. *Experimental Agriculture* 41: 449–73.

Utting, K. 2009. Assessing the impact of fair trade coffee: towards an integrative framework. *Journal of Business Ethics* (2009) 86: 127–149.

Utting-Chamorro, K. 2005. Does fair trade make a difference? The case of small coffee producers in Nicaragua. *Development in Practice* 15: 584-599.

Wehrmeyer, W., and Y. Mulugetta (eds.). 1999. *Growing Pains: Environmental Management in Developing Countries*. Sheffield, UK: Greenleaf Publishing.

Table 1. Variables, definitions, and means

Variable	Definition	Mean All (n=232)	Mean Cert. (n=56)	Mean Uncert. (n=176)	Diff. means
TREATMENT					
CERTIFIED	organic certified	0.24	1	0	***
OUTCOMES					
Negative practices					
CHEM_FERT	applies chemical fertilizer 2007 (0/1)	0.38	0.09	0.47	***
INSECTICIDE	applies insecticide 2007 (0/1)	0.13	0.02	0.16	***
HERBICIDE	applies herbicide 2007 (0/1)	0.12	0.09	0.13	
SEWAGE	dispose sewage in field 2007 (0/1)	0.29	0.16	0.34	***
Positive practices					
ORG_FERT	applies organic fertilizer 2007 (0/1)	0.51	0.88	0.40	***
SHADE	uses shade 2007 (0/1)	0.95	0.96	0.94	
CONTROLS					
Grower					
FEMALE	female (0/1)	0.28	0.23	0.29	
AGE	age 2007 (years)	51.53	52.73	51.15	
EDUCATION	highest grade completed 2007	3.70	3.68	3.70	
FAMILYSIZE	total persons in family 2007	4.94	5.20	4.85	
MEMBER	member coffee committee 1997 (0/1)	0.34	0.23	0.37	*
Farm					
OWN_FARM	own (vs. rent/lease) farm 1997 (0/1)	0.96	0.95	0.97	
NO_TREES	number coffee trees 1997	3252.92	3292.32	3240.38	
FARM_SIZE	farm size 1997 (has.)	2.16	2.44	2.07	
NO_LOTS	no. lots comprising farm 1997	1.66	1.88	1.59	*
ORGANIC_MAT	organic material available 1997 (kilograms)	2907.28	3430.93	2740.67	
CAPITAL_INDEX	no. capital items owned 1997 (1-6)[a]	1.08	1.18	1.05	
PBORBON	prop. coffee trees borbon variety 1997	0.04	0.90	0.02	***
PCATURRA	prop. coffee trees caturra variety 1997	0.58	0.46	0.61	**
PCOLOMBIA	prop. coffee trees colombia variety 1997	0.09	0.12	0.09	
PTIPICA	prop. coffee trees tipica variety 1997	0.25	0.31	0.23	
PCASTILLO	prop. coffee trees castillo variety 1997	0.01	0.01	0.01	
PVARIEDAD	prop. coffee trees other variety 1997	0.03	0.01	0.03	
PPASILLA	prop. coffee sales pasilla grade 1997	0.24	0.59	0.13	*
PFEDERACION	prop. coffee sales federacion grade 1997	0.87	0.58	0.96	*
PCLASICO	prop. coffee sales clasico grade 1997	0.04	0.00	0.05	*
PESPECIAL	prop. coffee sales especial grade 1997	0.01	0.01	0.01	
PCALIDAD	prop. coffee sales calidad grade 1997	0.10	0.26	0.05	***
BUYER_INT	sold coffee to intermediary 1997	0.46	0.54	0.44	
TRANS_VEH	transp. coffee to market with vehicle 1997	0.69	0.64	0.70	
PHH_F_WORKER	prop. household works on farm 1997	0.57	0.58	0.57	
PHH_OF_WORKER	prop. household works off farm 1997	0.19	0.20	0.18	
MUNICIP_INZ	municipality Inza	0.38	0.57	0.32	***
MUNICIP_CAJ	municipality Cajibo	0.16	0.05	0.19	*
MUNICIP_TAM	municipality Tambo	0.16	.16	0.15	
MUNICIP_TIM	municipality Timbo	0.13	0.02	0.17	***
MUNICIP_LAS	municipality Sierra	0.17	0.20	0.16	

*, **, *** = difference in mean for certified and uncertified subsamples significant at 10%, 5%, and 1% level

[a]capital items are: depulper, mill, silo, fumigator, motor, other

Table 2. Probit regression results
(dependent variable = organic certification)

Variable	Coefficient	S.E.
Grower		
FEMALE	0.072	0.259
AGE	0.000	0.001
EDUCATION	-0.018	0.055
FAMILYSIZE	0.082	0.057
MEMBER	-0.384	0.250
Farm		
OWN_FARM	-0.159	0.485
NO_TREES	0.000	0.000
FARM_SIZE	-0.001	0.052
NO_LOTS	0.295***	0.116
ORGANIC_MAT	0.000	0.000
CAPITAL_INDEX	0.001	0.110
PBORBON	1.116*	0.607
PCATURRA	0.256	0.334
PCOLOMBIA	0.835*	0.483
PCASTILLO	-0.066	1.451
PVARIEDAD	-0.821	1.005
PPASILLA	0.323	0.273
PFEDERACION	-0.029	0.086
PESPECIAL	0.552	1.169
PCALIDAD	1.396*	0.393
BUYER_INT	-0.124	0.285
TRANS_VEH	-0.220	0.240
PHH_F_WORKER	0.034	0.272
PHH_OF_WORKER	0.081	0.277
MUNICIP_INZ	0.109	0.366
MUNICIP_CAJ	-1.103**	0.497
MUNICIP_TIM	-1.381**	0.598
MUNICIP_LAS	-0.060	0.372
CONSTANT	-1.450	0.939

*, **, *** significant at 10%, 5%, and 1% level

Table 3. Matching quality: Number of covariates achieving balance (N)
and median standardized bias (SB) after matching,
for five propensity score matching methods[a,b,c]

Method	N	SB
(i) Nearest neighbor 1-1	27	9.513
(ii) Nearest neighbor 1-4	27	5.427
(iii) Nearest neighbor 1-8	28	6.693
(iv) Nearest neighbor 1-16	28	7.364
(v) Kernel	28	4.984

[a]The model includes 28 covariates.
[b]For a given covariate, the standardized bias (SB) is the difference of means in the certified and matched uncertified subsamples as a percentage of the square root of the average sample variance in both groups. We report the median SB for all covariates.
[c]Median SB before matching is 15.637.

Table 4. Negative practices: Average treatment effect on treated (ATT) estimates, by outcome variable and matching method; critical value of Rosenbaum's Γ

Propensity score matching method	**Mean certified**	**ATT**	**S.E.[a]**	**P-value**	**Γ*[b]**
CHEM_FERT					
(i) Nearest neighbor 1-1	0.089	-0.382	0.132	**0.004**	3.6
(ii) Nearest neighbor 1-4	0.089	-0.459	0.102	**0.000**	5.6
(iii) Nearest neighbor 1-8	0.089	-0.361	0.087	**0.000**	4.0
(iv) Nearest neighbor 1-16	0.089	-0.383	0.078	**0.000**	3.8
(v) Kernel	0.089	-0.324	0.094	**0.001**	3.4
INSECTICIDE					
(i) Nearest neighbor 1-1	0.018	-0.127	0.078	**0.105**	2.2
(ii) Nearest neighbor 1-4	0.018	-0.141	0.064	**0.028**	6.6
(iii) Nearest neighbor 1-8	0.018	-0.120	0.055	**0.030**	9.0
(iv) Nearest neighbor 1-16	0.018	-0.170	0.058	**0.003**	10.0
(v) Kernel	0.018	-0.117	0.051	**0.023**	9.2
HERBICIDE					
(i) Nearest neighbor 1-1	0.089	-0.055	0.093	0.558	1.2
(ii) Nearest neighbor 1-4	0.089	-0.045	0.077	0.553	2.2
(iii) Nearest neighbor 1-8	0.089	-0.032	0.063	0.615	2.6
(iv) Nearest neighbor 1-16	0.089	-0.033	0.057	0.565	2.8
(v) Kernel	0.089	-0.006	0.058	0.911	2.6
SEWAGE					
(i) Nearest neighbor 1-1	0.161	-0.200	0.135	0.137	1.4
(ii) Nearest neighbor 1-4	0.161	-0.277	0.112	**0.014**	2.4
(iii) Nearest neighbor 1-8	0.161	-0.264	0.100	**0.008**	1.8
(iv) Nearest neighbor 1-16	0.161	-0.269	0.090	**0.003**	1.6
(v) Kernel	0.161	-0.187	0.089	**0.037**	1.4

[a]Computed using bootstrap with 1,000 repetitions.
[b]Critical value of odds of differential assignment to organic certification due to unobserved factors (i.e., value above which ATT is no longer significant).
[c]Not estimated.

Table 5. Positive practices: Average treatment effect on treated (ATT) estimates, by outcome variable and matching method; critical value of Rosenbaum's Γ

Propensity score matching method	**Mean certified**	**ATT**	**S.E.[a]**	**P-value**	**Γ*[b]**
ORG_FERT					
(i) Nearest neighbor 1-1	0.875	0.600	0.120	**0.000**	7.2
(ii) Nearest neighbor 1-4	0.875	0.555	0.101	**0.000**	8.4
(iii) Nearest neighbor 1-8	0.875	0.441	0.092	**0.000**	5.6
(iv) Nearest neighbor 1-16	0.875	0.474	0.082	**0.000**	12.0
(v) Kernel	0.875	0.538	0.090	**0.000**	8.8
SHADE					
(i) Nearest neighbor 1-1	0.964	0.036	0.089	0.683	1.4
(ii) Nearest neighbor 1-4	0.964	0.055	0.066	0.409	3.0
(iii) Nearest neighbor 1-8	0.964	0.030	0.048	0.542	4.2
(iv) Nearest neighbor 1-16	0.964	0.016	0.040	0.689	6.0
(v) Kernel	0.964	0.024	0.052	0.642	4.6

[a]Computed using bootstrap with 1,000 repetitions.
[b]Critical value of odds of differential assignment to organic certification due to unobserved factors (i.e., value above which ATT is no longer significant).