WORKING PAPER Nº IDB-WP-01504

Estimating the Long-Term Effects of a Fruit Fly Eradication Program Using Satellite Imagery

Lina Salazar
Marcos Agurto
Luis Alvarez

Inter-American Development Bank
Environment, Rural Development and Disaster Risk Management Division

August, 2023

Estimating the Long-Term Effects of a Fruit Fly Eradication Program Using Satellite Imagery

Lina Salazar
Marcos Agurto
Luis Alvarez

Inter-American Development Bank
Environment, Rural Development and Disaster Risk Management Division

August 2023

Cataloging-in-Publication data provided by
the Inter-American Development Bank
Felipe Herrera Library

Salazar, Lina
Estimating the long-term effects of a fruit fly eradication program using satellite imagery / Lina Salazar, Marcos Agurto, Luis Alvarez.
p. cm. (IDB Working Paper Series ; 1504)
Includes bibliographical references.
1. Agricultural productivity-Peru-Remote sensing. 2. Pest-Control-Peru-Remote sensing. 3. Remote-sensing images-Peru. 4. Artificial satellites in agriculture-Peru. 5. Fruit-flies-Control-Peru. I. Agurto, Marcos. II. Alvarez, Luis. III. Inter-American Development Bank. Environment, Rural Development and Risk Management Division. IV. Title. V. Series.
IDB-WP-1504

http://www.iadb.org

Estimating the Long-Term Effects of a Fruit Fly Eradication Program Using Satellite Imagery

Lina Salazar
Inter-American Development Bank
lsalazar@iadb.org

Marcos Agurto
Universidad de Piura
marcos.agurto@udep.pe

Luis Alvarez
Inter-American Development Bank
luisalv@iadb.org

Abstract

This analysis applies a regression discontinuity approach combined with remote sensing data to measure the productivity impacts linked to a fruit-fly eradication program, implemented in Peru. For this purpose, satellite imagery was used to estimate a vegetation index over a 10-year span for a sample of 305 producers -155 treated and 150 controls-. The results confirmed that program participation increased agricultural productivity in the short and long terms, in a range from 12% to 49%. However, quantile regression methods suggest that most productive farmers were able to obtain greater impacts.

Keywords: Agricultural Productivity, Impact Evaluation, Remote Sensing, Satellite Images, Peru

JEL Codes: Q12, Q16, O13

Table of Contents

1 Introduction

The fruit fly plague has a significant impact on horticultural production in many regions around the world, especially in developing countries. The most significant impact of fruit fly infestation is the damage it causes to fruit crops, leading to a reduced agricultural income due to increased losses. In Africa, economic losses can reach US $2 billion annually, while in Brazil they could amount to US $ 242 million (Oliveira, et al., 2013). Similarly, in the Middle East, annual losses in Israel, Palestine, Lebanon, and Jordan accounted for US $165 million each year (Enkerlin & Mumford, 1997). In North Africa, the Mediterranean fruit fly also causes significant economic losses. In a review, Boulahia-Kheder (2021) found that could account up to between US $ 60-90 million per year (Ekesi, et al., 2016). However, the estimated damage varies across countries. For instance, while losses from citrus and summer fruits could reach US $ 6 million in Morocco (Aboussaid, et al., 2009), in Egypt this number can ascend to US $ 177 million (Mahmoud, et al., 2017). However, fruit fly plagues also cause significant damages in the developed countries. For instance, in Australia, during mid 1990s losses ascended to AUS $100 million, and their posterior eradication costed an additional AUS $ 34 million. In the case of the United States, recent estimations show that losses due to failure in the eradication of Mediterranean fruit fly from California could surpass US $ 1.8 billion dollars per year (Papadopoulos, 2014). In Peru, the National Services for Agricultural Health (SENASA) estimates that the fruit fly plague causes damages in at least 30% of the production and affects 233, 000 fruit producers in coastal valleys (SENASA, 2009; CENAGRO, 2012). The Integrated Pest Management (IPM) refers a set of agricultural practices to combat pests (including herbivores, pathogens, and weeds) using a combination of sustainable methods to diminish the rely on synthetic pesticides (Karlsson Green, et al., 2020). According to the FAO, the demand for cost-effective and environmentally friendly pest control has led to the development of phytosanitary interventions such as the IPM, that utilizes a range of techniques which includes organic pesticides, biological control, quarantine facilities and technical facilities. (Salazar, et al., 2020). Among IPM, several interventions have been implemented worldwide to eradicate fruit fly. (Enkerlin, 2021) conducts a cost-benefit analysis of programs implemented in several countries. For instance, e Fruit Fly Suppression Program in South Africa, an intervention consisting of aerial and ground releases of sterile male flies at hot spots, reported between 2001 and 2002 direct benefits of US $ 370 000 per year, resulting in a benefit-cost ratio of 2.8:1. In the case of Chile, being declared a fruit fly-free country as a result of its National Fruit Fly Program supposed fruit net exports valued in US $4 000 million in 2016. Similarly, in Mexico MOSCAMED program generated an estimated net revenue of US $39 300 million over 31 years, obtaining a return of US $112 in crops and control costs saved for each dollar that was invested.

In the literature we also find evidence of causal impacts of programs that aim to reduce the incidence or eradicate the fruit fly, even though they are focused on interventions mainly conducted in Africa and Southeast Asia. For example, a fruit fly IPM package developed under the Africa Fruit Fly Program (AFFP) for managing areas of mango production in Kenya reported reductions on export rejections (55%) and insecticide expenditures (46%), as well as a decrease in mango yield losses by 19% on average among IPM users (Kibira, et al., 2015; Muriithi, et al., 2016). Furthermore, when examining the differentiated impact of distinct IPM practices. Midingoyi et al (2019) found that adoption of one, two, or three or more IPM practices provide yields of 6%, 27% and 95%, respectively. In Bangladesh, the Integrated Pest Management Collaborative Research Support Program consisted in the use of a synthetic pheromone called Cuelure to bait trap fruit fly in cucurbits, reported an internal rate of return of the investment between 140% and 151% over 15 years. (Rakshit, et al., 2011).

Considering the potential benefits of eradicating the fruit fly plague in Peruvian coastal valleys, SENASA with the support of the Inter-American Development Bank, launched a program that included the application of bioinsecticides, training farmers in pest prevention and control, biological control measures through the release of sterile male flies, among others. Between 1998 and 2014, the program was implemented in three phases, covering over one million hectares of agricultural land and 150,000 hectares of host crops. Each phase of the program designated a specific region for intervention. Phase 1 of the program covered the southernmost parts, and subsequent phases covered adjacent areas expanding coverage are towards the north of the country, along the coastal area. The implementation of this strategy results in the creation of "intervention borders" that separates treated and untreated areas. This geographical discontinuity sets an allocation rule that resembles a randomized control trial (RCT) in the vicinity of the intervention border associated with each phase. Thus, farmers at each side of the border are expected to be similar in their observable and unobservable features.

With the purpose of estimating the short-term effects of the third phase of the program, (Salazar, et al., 2020) conducted an impact evaluation using a geographical regression discontinuity design. The findings suggest that farmers in treated areas increased their pest knowledge and were more likely to implement best practices for plague prevention and control, which allowed them to experience increases in fruit crops productivity and sales. In this paper, we aim to estimate the short and long-term effects of the third phase of the fruit fly eradication program, using remote sensing data to calculate the Normalized Difference Vegetation Index (NDVI) as a proxy of crop yield.

The use of remote sensing technologies and satellite imagery to construct vegetation indices has applications such as optimization of crop management, crop production forecasting, crop and land monitoring, measurement of the spatial variation of productivity and estimation of agricultural yields, among others (Martos, et al., 2021; Aguilar Rivera, 2015). The NDVI is one of the most popular and effective of these indices used to express vegetation status and quantify vegetation attributes. As other indices created to summarize complex data, the attractivity of NDVI lies in its capacity to identify vegetation and vegetative stress, which is very useful in agriculture and land-use studies (Huang, et al., 2021).

Even though we can find several uses of NDVI in the literature, either to quantify the evolution in forest regeneration (Alphan & Ali Derse, 2013; Meroni, et al., 2017) or to measure the effects of floods and droughts on crops (Shrestha, et al., 2013; Sholihah, et al., 2016), its relationship with agricultural productivity and crop yields has also been explored in the literature. Studies in the early 1990s found that high NDVI values are correlated with better plant health (Chuvieco, 1991) and can be used to estimate accurately crop yields (Quarmby, et al., 1993). More recently, studies conducted in Japan and China show that crop yields in rice and wheat are highly correlated with NDVI values (Huang, et al., 2014; Guan, et al., 2019). For the case of Latin America, Selvaraj et al. (2020) found that a set of vegetation indices (including NDVI) can be used to estimate the growth dynamics and yields of cassava in Colombia. In Salazar, et al., (2021), the authors use the NDVI to estimate the long-term impact of an agricultural program in the Dominican Republic, finding that program beneficiaries reported higher NDVI values during the post-treatment period, suggesting that farmers in this group experienced higher productivity.

The paper proceeds as follows: Section 2 describes the NDVI data; Section 3 presents the SENASA fruit fly program and discusses the empirical methodoly and Section 4 shows the estimations from the regression discontinuity and difference-in-differences models. Lastly, Section 6 concludes.

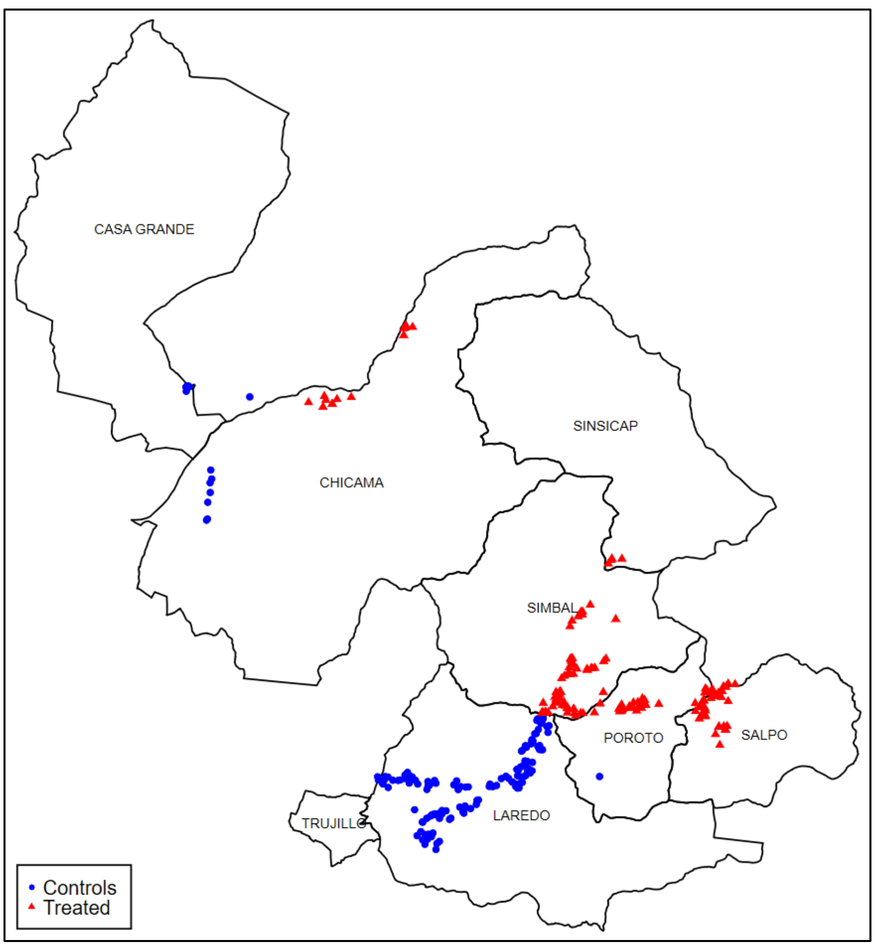
2 Empirical Methodology

In 1998, SENASA initiated a program to eradicate fruit flies from the Peruvian coastal zone. The program includes several activities, such as providing farmers with technical assistance and training on pest prevention and control, installing fruit fly traps, applying insecticides, releasing male sterile flies, and implementing quarantine centers. It was implemented in phases, with each phase targeting a specific region on the coast of Peru. The third phase of the program, implemented between 2012 and 2014, covered 95, 381 hectares of host crops and 756, 746 agricultural hectares in Lima, Ancash, and La Libertad.

In this paper, we will use a regression discontinuity design to examine the effects of the third phase of the program. Following the impact evaluation literature, in order to estimate the causal effect of the program, our empirical methodology must identify a proper counterfactual group, that is to say, a group that consists of untreated units similar in observable and unobservable characteristics to the treated units. In this case, the program implementation strategy created an intervention border with treated and untreated at both sides of the frontier. Thus, the existence of this intervention border creates an allocation rule, as only farmers on the north-east side of the border (see Figure 2) are exposed to the treatment. This geographic discontinuity in the allocation to the program allows to identify a comparable counterfactual group of farmers who did not participate in the program, and use the regression discontinuity approach to identify a local average treatment effect (LATE) of the program.

As producers are unable to manipulate their assignment to the treatment status (e. g.: they cannot change their location from untreated to treated areas), the variation in the treatment in the vicinity of the threshold resembles the randomization obtained in an RCT (Lee & Lemieux, 2010). Furthermore, as suggested by Chaplin et al. (2018) and Gleason et al. (2018), comparisons of RCTs and RDDs indicate that both approaches tend to be similarly reliable in identifying causal effects. For this reason, regression discontinuity designs have seen an increase in their use in agricultural and environmental economics (Wuepper & Finger, 2023).

Figure 1. Location of control and treated units.



As it is highlighted in Salazar et al (2020), the validity of the geographic discontinuity approach in context of this program relies on the exogeneity in the location of the intervention border. In other words, the validity of this assumption holds if the location of the border is not influenced by factors related to agricultural outcomes and is not affected by the interests of potential beneficiaries. In this regard, SENASA affirmed that the delimitation of the coverage area was based only on financial and logistic constraints and all agricultural valleys located in the study area are part of a relatively uniform geographical area. Furthermore, in this phase of the intervention the intervention border does not coincide with an administrative border, as all control and treated areas considered for this analysis are located in La Libertad region.

However, the Peruvian geography itself might pose a threat to this assumption, as, due to the features of the Andean highlands, the program border might coincide with abrupt geographic changes. Salazar et al (2020) addresses this issue by comparing pre-treatment characteristics of treatment and control groups at the household and farm level. Table A.1 shows that most of features are "balanced and very similar among control and treatments units", which suggests that agricultural features vary smoothly at the intervention border.

Still, the program design leaves room for the existence of potential geographical spillovers that might reduce effectiveness of the treatment, such as peer-learning effects. Neighboring farmers in control areas may adopt preventive and control measures through word-of-mouth or observational learning. However, peer-effects might be limited, since the intervention package also includes activities, such as monitoring traps, insecticide use and quarantine centers for contaminated crops. Furthermore, a field knowledge test has demonstrated that beneficiary farmers have greater knowledge about plague characteristics and control measures than control farmers. Nonetheless, it is worth noting that the presence of spillover effects would bias the LATE downwards, and thus the coefficient estimated would represent a lower bound estimate.

Assuming an exogenous determination of the border and that relevant features correlated with agricultural outcomes gradually vary across the border, we can estimate the LATE of the fruit fly eradication program by estimating the following regression discontinuity model:

$$ndvi_{it} = \beta_0 + \tau\, T + h(dist_i, dist_i \times T_i\,) + \varepsilon_{it}$$

Where $ndvi_{it}$ is the NDVI average value registered in plot *i* at agricultural year *t* and T_i is a dummy variable that indicates the treatment status of the unit. Thus, the parameter τ captures the program LATE. $dist_i$ is a continuous variable that measures the distance in kilometers from unit i's location to the intervention border[1] and works as the running (assignment) variable of the regression discontinuity model.[2] Finally, $h(dist_i, dist_i \times T_i)$ is a flexible functional form of $dist_i$ and its interaction with T_i. This interaction term allows the estimation of the conditional mean function with different functional forms at each side of the threshold. As it is suggested by Lee & Lemieux (2010) when a panel data structure is available, given that including individual and time fixed effects is unnecesary for the identification in an regresession discontinuity design, we will conduct the analysis for the entire pooled-cross-section dataset for post-treatment periods.

Alternatively, since we have data available for control and treatment groups before and after the treatment, we can estimate a Difference-in-Differences model to estimate the average treatment effect of the program on agricultural productivity. The following model is estimated:

$$ndvi_{it} = \alpha + \rho T_i + \gamma post_t + \delta\, T_i \times post_i + \varphi yield_{it} + agryear_t + district_i + \epsilon_{it}$$

Where $ndvi_{it}$ is again the NDVI value of plot i at agricultural year t and T_i indicates treatment status. In this specification, $post_t$ is a dummy variable that takes a value of zero in pre-treatment periods (agricultural years 2010-11 and 2011-12) and is equal to one in post-treatment periods. The remaining regressors include $yield_{it}$, a variable that accounts for the trends in avocado crops at the province level, and agricultural year and district fixed effects.

3 Data

Remotely sensed multi-spectral imagery uses a combination of bands to produce a composite image. It is possible to transform these individual bands in the band composite to identify specific features and patterns more clearly. (Kriegler, et al., 1969) proposed a simple band transformation: the difference between near-infrared radiation (NIR) and red radiation, divided by their sum, which resulted in the NDVI (Huang, et al., 2021):

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

By construction, NDVI values range between -1 and 1. Typically, negative values indicate the presence of water bodies, while values close to zero are related to rocks, sand or concrete surfaces. Positive values suggest the presence of vegetation, such as crops, shrubs, grasses and forests. (Huang et al., 2021).

[1] Instead of using a Euclidian distance to the border, we define the distance to the border as the distance between a beneficiary farmer and its closest control located at the other side of the border. See Salazar et al (2020) for further details.
[2] For convenience, we defined the variable such that units located on the control side of the border have negative values, while units on the treated side have positive values. This implies the cutoff point is located at zero.

For this study, the NDVI data was obtained using satellite imagery from the Landsat-7 and Landsat-8 satellites. Landsat images were collected with a bi-monthly frequency over the period 2010-2020, for 305 plots -155 treated and 150 controls- located in districts adjacent to the intervention border that corresponds to the third phase of the program.

The treatment intervention started implementation in mid-2012, for this reason we define three different periods in our sample: a pre-treatment period before the implementation of the program (agricultural years 2010-11 and 2011-12), the short-term post-treatment period composed of four agricultural cycles (2012-13, 2013-14,2014-15 and 2015-16), and the long-term post treatment period, which begins 5 years after the program implementation, composed of four agricultural cycles (2016-15, 2017-18, 2018-19 and 2019-20). Table 1 presents the NDVI values[3] for treatment and control groups over the period analyzed. As we can see, treated observations report a higher value of NDVI than controls, for periods before and after the implementation of the program.

As Table 1 suggests, there are significative differences in NDVI values between both groups prior to the treatment. However, as it is argued in (Salazar, et al., 2020), treatment and control groups possess similar household and farm features in the pre-treatment period (see Table A.1. in the appendix). Further discussion on this is on the next section.

Table 1. NDVI descriptive statistics over the sample period.

	General				Treatment Group		Control Group		Diff
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mean	Min	Max	N	Mean	N	Mean	N	Mean
2010-11	0.18	0.04	0.41	305	0.20	155	0.16	150	0.05***
	(0.00)				(0.08)		(0.06)		
2011-12	0.18	0.02	0.36	305	0.19	155	0.16	150	0.03***
	(0.00)				(0.07)		(0.06)		
Pre-treatment	**0.18**	**0.02**	**0.41**	**610**	**0.20**	**310**	**0.16**	**300**	**0.04*** **
	(0.07)				**(0.07)**		**(0.06)**		
2012-13	0.23	0.04	0.48	305	0.27	155	0.18	150	0.09***
	(0.01)				(0.11)		(0.07)		(0.01)
2013-14	0.26	0.04	0.48	305	0.27	155	0.26	150	0.01
	(0.01)				(0.11)		(0.11)		(0.01)
2014-15	0.22	0.05	0.41	305	0.23	155	0.2	150	0.02**
	(0.00)				(0.09)		(0.07)		(0.01)
2015-16	0.24	0.05	0.5	305	0.28	155	0.19	150	0.08***
	(0.01)				(0.11)		(0.07)		(0.01)
Short-term post-treatment	**0.24**	**0.04**	**0.5**	**1220**	**0.26**	**620**	**0.21**	**600**	**0.05*** **

[3] In this analysis, our outcome variable is constructed by calculating 12-month averages of the NDVI values. Thus, the resulting panel is composed by 3050 observations in 10 periods, each one corresponding to an agricultural year that starts in August and ends in July next year.

	(0.10)				**(0.11)**		**(0.09)**		
2016-17	0.28	0.06	0.53	305	0.31	155	0.26	150	0.01***
	(0.01)				(0.1)		(0.09)		(0.01)
2017-18	0.21	0.06	0.43	305	0.25	155	0.17	150	0.08***
	(0.00)				(0.09)		(0.06)		(0.01)
2018-19	0.28	0.03	0.52	305	0.29	155	0.27	150	0.03**
	(0.01)				(0.12)		(0.11)		(0.01)
2019-20	0.24	0.04	0.49	305	0.25	155	0.22	150	0.02*
	(0.01)				(0.11)		(0.09)		(0.01)
Long-term post-treatment	**0.25**	**0.03**	**0.53**	**1220**	**0.27**	**620**	**0.23**	**600**	**0.04*** **
	(0.10)				**(0.10)**		**(0.10)**		

Note: *** p<0.01, ** p<0.05, * p<0. 1. Standard errors in parentheses.

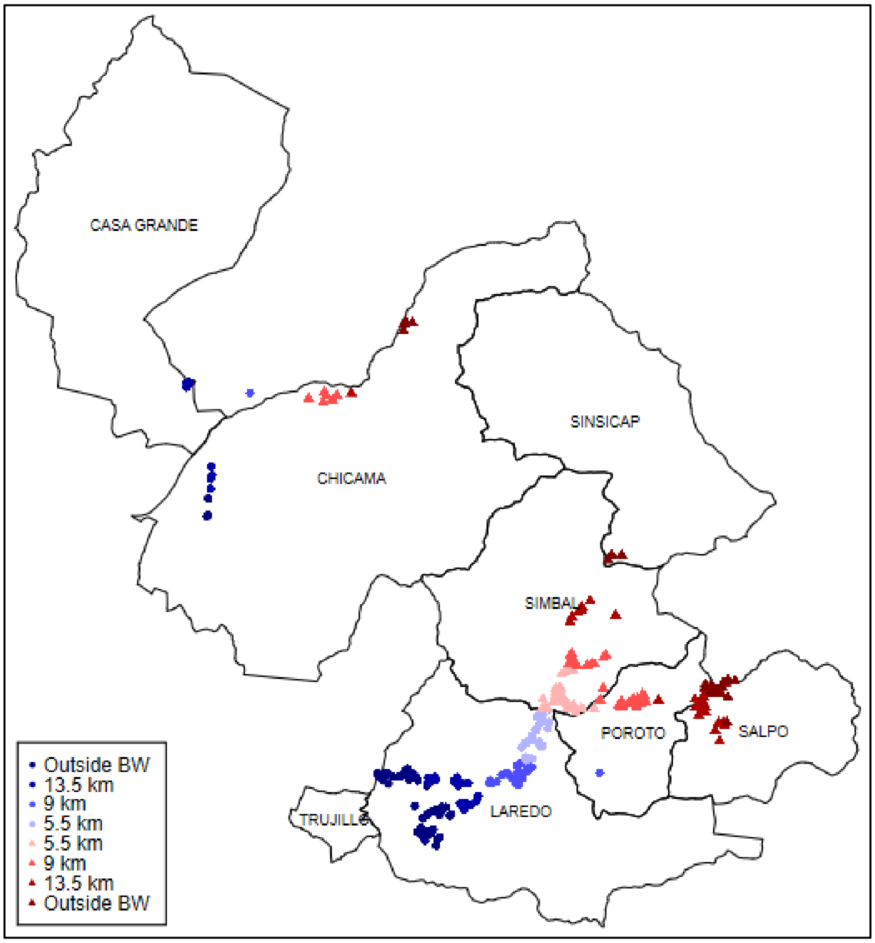
To corroborate the validity of our approach, we analyze the NDVI pre-treatment values of both treated and untreated producers along the intervention border. For our assumptions to hold, our outcome of interest must be similar for farmers located at both sides of the border in the baseline, with these similarities being stronger for producers close to the border. Table 2 compares the NDVI outcome in treated and control groups. The difference in the average NDVI value is statistically indistinguishable from zero at the 5.5 km bandwidth. However, we find significant differences when we expand the bandwidth. When we consider the 50% of observations closest to the intervention border, NDVI average values are higher in the treated area, although this difference is only statistically significant at the 10% level. Similarly, observations in treated and control groups within the 13.5 km bandwidth also report significant baseline differences. For this reason, in order to estimate the average treatment effect of the program, we will also develop a difference-in-differences approach, which will account for pre-treatment time invariant differences.

Table 2. Pre-treatment NDVI values by bandwidth.

	Treatment Group	Control Group	Diff in means
	(1)	(2)	(3)
Bandwidth 5.5 km (25% of obs closest to the border)	0.1687	0.1621	0.0066
N	86	62	
Bandwidth 9 km (50% of obs closest to the border)	0.1743	0.1601	0.0141*
N	188	126	
Bandwidth 13 km (75% of obs closest to the border)	0.1845	0.1605	0.0239***
N	256	192	

Note: *** p<0.01, ** p<0.05, * p<0. 1. Standard errors in parentheses.

Figure 2. Location of control and treatment units according to their distance to intervention border



Note: Blue-colored dots represent control units. Red-colored triangles represent treatment units.

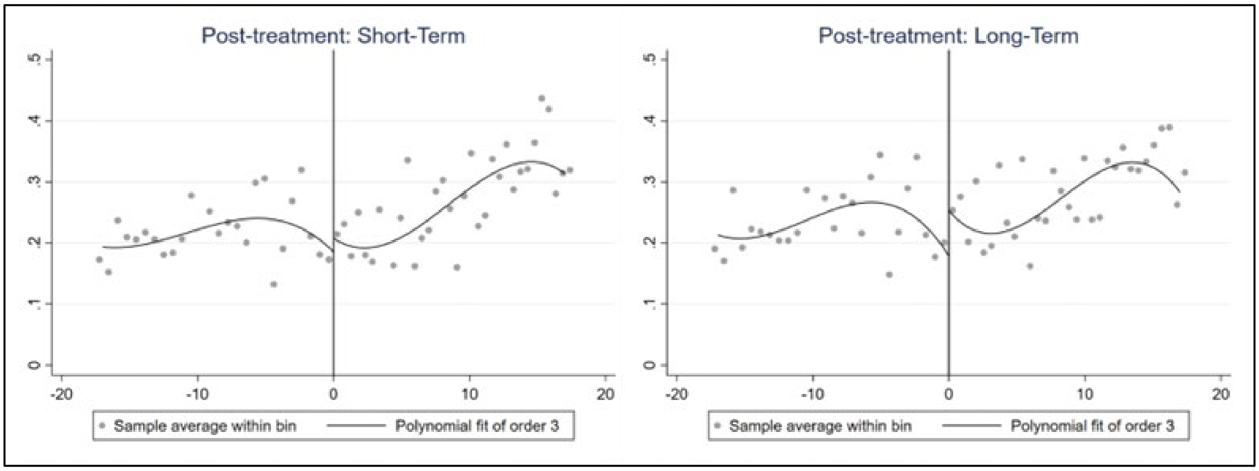
4 Results

4.1 Regression Discontinuity results

The Regression Discontinuity approach is based on the existence of a discontinuity in the outcome variable at or near the cut-off point. The graphics in Figure 3 show this discontinuity after the implementation of the treatment: by plotting distance to the border measured in the X-axis, we can see this “jump “in the NDVI values (plotted in the Y-axis).

If the assumptions of an exogenous placement of the intervention border and the location of control and treatment units discussed in the previous section holds, we could consider that the program allocation rule resembles a randomized process at the vicinity of the intervention border. Thus, this “jump” observed in the control variable at the cut-off point constitutes a LATE.

Figure 3. Discontinuity of NDVI at the cut-off point in post-treatment periods.



Note: distance in the x-axis refers to the distance to the closest neighbor on the opposite side of the border. Treated observations are located to the right of the threshold (*dist>0*).

Having data available for multiple post-treatment periods allow us to implement a regression discontinuity design using observations from those periods to see the evolution of the LATE and estimate both the short-term and long-term impacts of the program. The results of this estimations are presented in Table 3. The regressions estimated in columns 1 and 2 are estimated for observations between three comparison bandwidths: 5.5 km, 9 km and 13.5 km from the border. These arbitrary distances encompass 25%, 50% and 75% of the producers, respectively. See Figure 2 to see the location of the control and treatment units within the comparison bandwidths defined.

In order to mitigate the concerns of overfitting the model, for the estimation of the models with observations only within the 5.5 km (296 observations) and 9 km bands (628 observations), we specified a quadratic polynomial for the $h(dist_i, dist_i \times T_i)$ functional form. For the case of the model using observations from the 13.5 km bandwidth, given that we have 896 observations available, we specified a cubic functional form.[4]

As we can see from Table 3, the first row shows that the LATE estimates obtained for each bandwidth are statistically significantly different from zero with a level of confidence of 95% in both time horizons. The reported short-term estimated LATE coefficients in column 1 range between 0.02 and 0.04. For the case of

[4] These models were estimated using the command *rdrobust* in Stata, which is equivalent to estimate the following equation for the quadratic specification:

$$ndvi_{it} = \beta_0 + \tau\, T + \beta_{11} dist_i + \beta_{12} dist_i \times T_i + \beta_{21} dist_i^2 + \beta_{22} dist_i^2 \times T_i + \varepsilon_{it}.$$

In the cubic case, the equation is the following:

$$ndvi_{it} = \beta_0 + \tau\, T + \beta_{11} dist_i + \beta_{12} dist_i \times T_i + \beta_{21} dist_i^2 + \beta_{22} dist_i^2 \times T_i + \beta_{31} dist_i^3 + \beta_{32} dist_i^3 \times T_i + \varepsilon_{it}$$

the 5.5 km bandwidth, the short-term effect is 0.029, which accounts for an increase of 16% directly attributable to the program, with respect to the pre-treatment average. For the 9 km and 13.5 km bandwidths, the short-treatment effects account for increases of 12% and 20%, respectively. These results are in line with the short-treatment effects found in Salazar et al. (2016) and Salazar et al. (2020) where the authors find a positive significant increase of the program on fruit output and value of fruit production.

Table 3 also suggests that the effects of the program grow over time. Considering that the program began implementation by the end of 2012, the long-term estimated LATE reported in column 2 shows that the program have significant and larger effects three to seven years after the implementation of the program. The average increase in NDVI values during the period between agricultural years 2016-2017 and 2019-2020 that is directly attributable to the program according to our regression discontinuity model ranges between 0.066 (37% increase) and .088 (49% increase).

Table 3. LATE of Fruit Fly Eradication Program on NDVI: RD Estimation

	(1) Short-term effect	(2) Long-term effect
5.5 km to the border (25% obs. closest to the border)	0.0288**	0.0658***
	(0.0124)	(0.0165)
N	296	296
9 km to the border (50% obs. closest to the border)	0.0211**	0.0672***
	(0.0105)	(0.0131)
N	628	628
13.5 km to the border (75% obs. closest to the border)	0.0368***	0.0883***
	(0.0123)	(0.0152)
N	896	896

Note: *** p<0.01, ** p<0.05, * p<0. 1. Standard errors in parentheses.

4.2 Difference-in-differences results

Alternatively, the data panel structure with allows us to estimate a Difference-in-Differences model that would be able to capture an average treatment effect for the whole sample. The results of this methodology are presented in Table 4. The OLS difference-in-differences estimator (column 6) is positive both in the short and long terms, however only in the short term it is statistically different from zero with a 95% confidence level (the p-value obtained for the long-term effect estimator is 0.148).

When considering quantile regressions, however, is it worth noting that both the short- and long-term treatment effects are concentrated in the upper tails of the NDVI distributions. In the short term, most of the quantile treatment effects for quantiles below 0.3 are not statistically different from zero. The treatment effect at the median is 0.022 with a 95% confidence level, which is equivalent to an increase of 12% in the NDVI value. Similarly, the reported treatment effects for at quantiles 0.75 and 0.9 represent increases of 14% and 15%, respectively. In the long run, the most productive workers (from quantiles .45 and above) also seem to be the most benefited from the intervention, with estimated quantile effects between 0.02 and 0.035, that account for increases of roughly 12%-13%. These results suggests that the most more productive farmers benefit more from this intervention more from interventions. to increase productivity, as they may have the resources, know-how and abilities to take advantage of the program. Figure 4 plots the Diff-in-Diff quantile treatment effect in the short and long-term, for quantiles 5 to 95 (gray line), and the Diff-in-diff OLS estimator (dashed straight line). As we can see from the Figure 4, quantile regressions effects estimated are positive both in the short and long run, however, for low quantiles the effects found are not statistically significant, as the confidence interval contains the zero value.

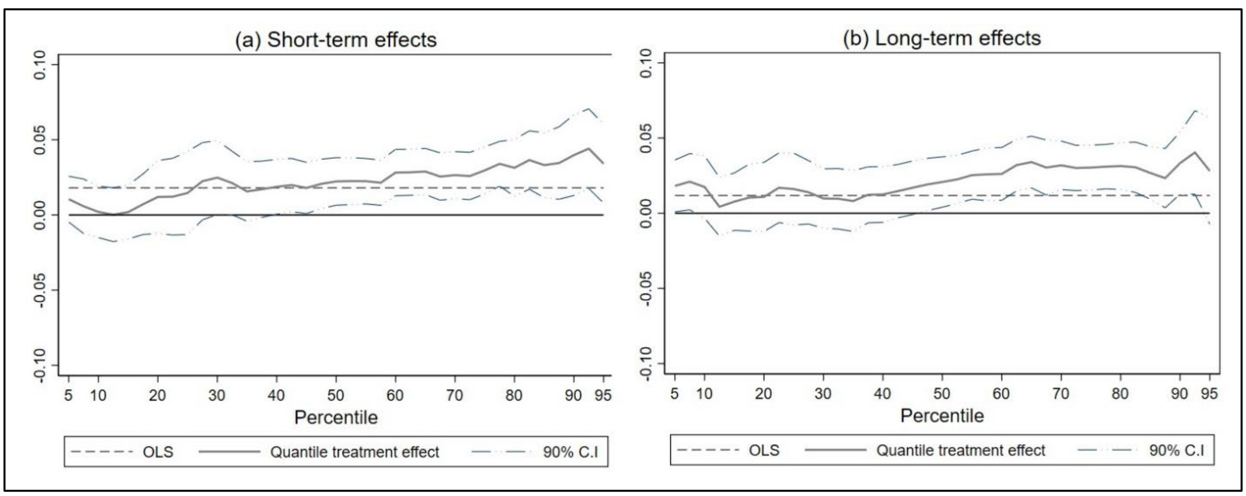
Table 4. ATE and QTE of Fruit Fly Eradication Program on NDVI: Diff-in-Diff Estimation

Table 5. Quantile treatment effects on the treated

	Quantile Regression Estimates					OLS Estimates
	(1)	(2)	(3)	(4)	(5)	(6)
	0.1	0.25	0.5	0.75	0.9	Coeff.
Short-term effect	0.0020	0.0150	0.0223**	0.0319**	0.0398**	0.0180**
	(0.0104)	(0.0168)	(0.0096)	(0.0095)	(0.0163)	(0.0082)
Pseudo R2 / R2	0.1412	0.1339	0.1911	0.2359	0.2434	0.2991
Long-term effect	0.0174	0.016	0.0208**	0.0304***	0.0333***	0.0118
	(0.0126)	(0.0145)	(0.0102)	(0.009)	(0.0125)	(0.0081)
Pseudo R2 / R2	0.1439	0.1599	0.2243	0.2716	0.269	0.3393

Note: *** p<0.01, ** p<0.05, * p<0. 1. Standard errors in parentheses. The pseudo-R2 is reported for quantile regressions (columns 1-5). For the OLS diff-in-diff regression, the adjusted R2 is reported.

Figure 4. Diff-in-diff quantile treatment effects.



4.3 Placebo (Falsification) tests

To confirm that the LATE, ATE and QTE estimations capture the impact of the program, instead of systematic differences between treatment and control groups, we run placebo tests for the regression discontinuity and difference-in-differences approaches. Results are presented in Tables 5 and 6.

We first measure the LATE impact of the fruit fly program on pre-treatment NDVI values (agricultural years 2010-11 and 2011-12). Table 5 confirms that our outcome variable was not affected by the intervention during this period, as no coefficient is found to be significant.

Table 6. Falsification test: Pre-treatment LATE estimations.

	(1) Pre-treatment effect
5.5 km to the border (25% obs. closest to the border)	0.0244
	(0.0336)
N	148
9 km to the border (50% obs. closest to the border)	0.0127
	(0.0272)
N	314
13.5 km to the border (75% obs. closest to the border)	0.0122
	(0.0298)
N	448

Note: *** p<0.01, ** p<0.05, * p<0. 1. Standard errors in parentheses.

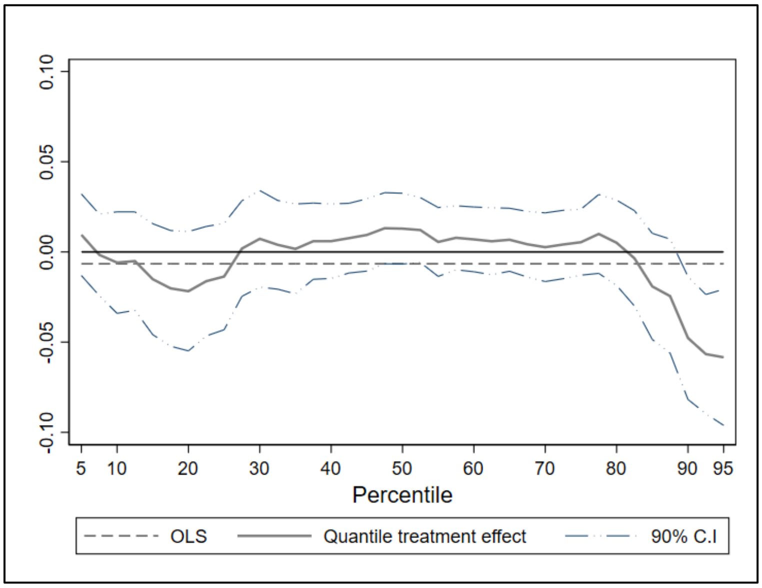
We also estimate the ATE and QTE for the pre-treatment period. As we can see in Figure 5, the quantile treatment effects obtained using the difference-in-differences approach were not statistically different from zero for most quantiles below 0.9. We only find a negative and significant QTE in the higher upper tail of the distribution (see column 5 in Table 5 for the QTE reported for quantile 0.9). Similarly, we do not find a significant average treatment effect using the difference-in-differences estimation during the pre-treatment period (see column 6 in Table 5). Therefore, these placebo tests confirm that the impacts found in NDVI values are capturing true causal effects instead of systematic or uncontrolled differences between the comparison groups.

Table 7. Falsification test: Pre-treatment ATE and QTE estimations.

	Quantile Regression Estimates					OLS Estimates	
	(1)	(2)	(3)	(4)	(5)	(6)	
	0.1	0.25	0.5	0.75	0.9	Coeff.	N
Pre-treatment effect	-0.0059	-0.0137	0.0129	0.0054	-0.0477**	-0.0066	610
	(0.0171)	(0.0179)	(0.0119)	(0.0111)	(0.0207)	(0.0101)	
Pseudo R2 / R2	0.1466	0.1411	0.1723	0.2166	0.2121	0.2858	

Note: *** p<0.01, ** p<0.05, * p<0. 1. Standard errors in parentheses. The pseudo-R2 is reported for quantile regressions (columns 1-5). For the OLS diff-in-diff regression, the adjusted R2 is reported.

Figure 5. Placebo test: Pre-treatment quantile treatment effects



5 Conclusions

This study uses remote sensing data collected over a 10-year time span to assess the short and long-term effects of a fruit fly eradication program implemented in the coastal regions of Peru. Using the NDVI vegetation index obtained from satellite imagery, as a proxy of agricultural productivity, we were able to track the evolution of crops health in 305 plots -155 beneficiaries and 150 controls. Given that intervention borders were set independently of agricultural outcomes, we exploited the program design to estimate the LATE of the SENASA fruit fly eradication program on agricultural productivity through a regression discontinuity model. We find that during the first four years after the implementation of the program, the increases in productivity associated with the eradication program among producers located close to the program border on average ranged between 12% and 20%. This effect seems to increase over time, as during the next four years, the local average treatment effect ranged between 37% and 49%.

Additionally, in order to capture the overall effect of the treatment and the effects on specific parts of the NDVI distribution, we estimated a difference-in-differences model that would allow us to estimate the average and the quantile treatment effects (ATE and QTE, respectively). When find that the overall treatment effect is lower than the local effects found in the regression discontinuity model (10% increase in the short-term, while not statistically different from zero in the long-term). However, when we examine the NDVI distribution, we can see that the positive treatment effect is concentrated in the upper tails, both in the short and long term, which suggests that more productive farmers are able to take advantage of the program and increase their productivity even further.

The potential policy implications could be useful for policymakers in designing and targeting agricultural programs, as our findings suggest that implementing pest eradication and disease control interventions could boost agricultural productivity.

This work, as one of the firsts that combine regression discontinuity experiments with satellite images to evaluate the impact of agricultural programs in the long run, also highlights the potential of remote sensing as a cost-effective alternative to collect data on crops over time and monitor the effects of policies and programs in the long term, years after the project's implementation.

References

Aboussaid, H., El Messoussi, S. & Oufdou, K., 2009. Activité insecticide d'une souche marocaine de Bacillus thuringiensis sur la mouche m´editerran´eenne : Ceratitis capitata (Wied.) (Diptera: tephritidae)). *Afrique Science,* Volume 5, pp. 160-172.

Aguilar Rivera, N., 2015. Percepción remota como herramienta de competitividad de la agricultura. *Revista Mexicana de Ciencias Agrícolas,* Volume 2, pp. 399-405.

Alphan, H. & Ali Derse, M., 2013. Change detection in Southern Turkey using normalized difference vegetation index (NDVI). *Journal of Environmental Engineering and Landscape Management,* 21(1), pp. 12-18.

Boulahia-Kheder, S., 2021. Review on major fruit flies (Diptera: Tephritidae) in North Africa: Bio-ecological traits and future trends. *Crop Protection,* Volumen 140.

Chuvieco, E., 1991. Fundamentos para la teledetección espacial. *Estudios Geográficos,* 52(203), p. 371.

Ekesi, S. y otros, 2016. Taxonomy, Ecology, and Management of Native and Exotic Fruit Fly Species in Africa. *Annual Review of Entomology,* Volumen 61, pp. 2019-238.

Enkerlin, W., 2021. Impact of Fruit Fly Control Programmes Using the Sterile Insect Technique. In: *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated pest Management.* s.l.:s.n., p. 28.

Enkerlin, W. & Mumford, J., 1997. Economic Evaluation of Three Alternative Methods for Control of the Mediterranean Fruit Fly (Diptera: Tephritidae) in Israel, Palestinian Territories, and Jorda. *Journal of Economic Entomology,* 90(5), pp. 1066-1072.

Guan, S. et al., 2019. Assessing Correlation of High-Resolution NDVI with Fertilizer Application Level and Yield of Rice and Wheat Crops using Small UAVs. *Remote Sensing,* 11(112).

Huang, J., Wang, H., Dai, Q. & Han, D., 2014. Analysis of NDVI Data for Crop Identification and Yield Estimation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 7(11), pp. 4374 - 4384.

Huang, S. et al., 2021. A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research,* Volume 32, pp. 1-6.

Karlsson Green, K., Stenberg, J. & Lankinen, A., 2020. Making sense of Integrated Pest Management (IPM) in the light of evolution. *Evolutionary Applications,* Volumen 13, pp. 1791-1805.

Kibira, M. et al., 2015. Economic evaluation of integrated management of fruit fly in mango production in Embu County, Kenya. *African Journal of Agricultural and Resource Economics,* 10(4), pp. 343-353.

Kriegler, F., Malila, W., Nalepka, R. & Richardson, W., 1969. Preprocessing Transformations and Their Effects on Multispectral Recognition. En: *Preprocessing Transformations and Their Effects on Multispectral Recognition.* s.l.:s.n., p. 97.

Mahmoud, M. et al., 2017. Low environmental impact method for controlling the peach fruit fly, Bactrocera zonata (Saunders) and the Mediterranean fruit fly Ceratitis capitata (Wied.) in mango orchards in Egypt. *Journal of Applied Life Sciences and Environment.*

Martos, V., Ahmad, A., Cartujo, P. & Ordoñez, J., 2021. Ensuring agricultural sustainability through remote sensing in the era of agriculture 5.0. *Applied Sciences,* 11(13), p. 5911.

Meroni, M. et al., 2017. Remote sensing monitoring of land restoration interventions in semi-arid environments with a before–after control-impact statistical design. *International Journal of Applied Earth Observation and Geoinformation,* Volume 59, pp. 42-52.

Midingoyi, S. y otros, 2019. Do Farmers and the Environment Benefit from Adopting Integrated Pest Management Practices? Evidence from Kenya. *Journal of Agricultural Economics,* 70(2), pp. 452-470.

Muriithi, B. et al., 2016. Impact assessment of Integrated Pest Management (IPM) strategy for suppression of mango-infesting fruit flies in Kenya. *Crop Protection,* Volume 81, pp. 20-29.

Quarmby, N., Milnes, M., Hindle, T. & Silleos, N., 1993. The use of multi-temporal NDVI measurements from AVHRR data for crop yield estimation and prediction. *International Journal of Remote Sensing,* 14(2), pp. 199-210.

Rakshit, A., Karim, A., Hristovska, T. & Norton, G., 2011. Impact Assessment Of Pheromone Traps To Manage Fruit Fly On Sweet Gourd Cultivation. *Bangladesh Journal of Agricultural Research,* 36(2), pp. 197-203.

Salazar, L. et al., 2020. Sweeping the flies away: evidence from a fruit fly eradication program. *European Review of Agricultural Economics,* 47(5), pp. 1920-1962.

Salazar, L., Maffioli, A., Aramburú, J. & Agurto, M., 2016. Estimating the Impacts of a Fruit fly Eradication Program in Peru: A Geographical Regression Discontinuity Approach. *IDB Working Paper Series N° IDB-WP-677*.

Salazar, L., Palacios, A., Selvaraj, M. & Montenegro, F., 2021. Using Satellite Images to Measure Crop Productivity: Long-Term Impact Assessment of a Randomized Technology Adoption Program in the Dominican Republic. *IDB Working Paper Series N° IDB-WP-01234*.

Selvaraj, M. et al., 2020. Machine learning for high-throughput field phenotyping and image processing provides insight into the association of above and below-ground traits in cassava (Manihot esculenta Crantz). *Plant Methods,* 16(87).

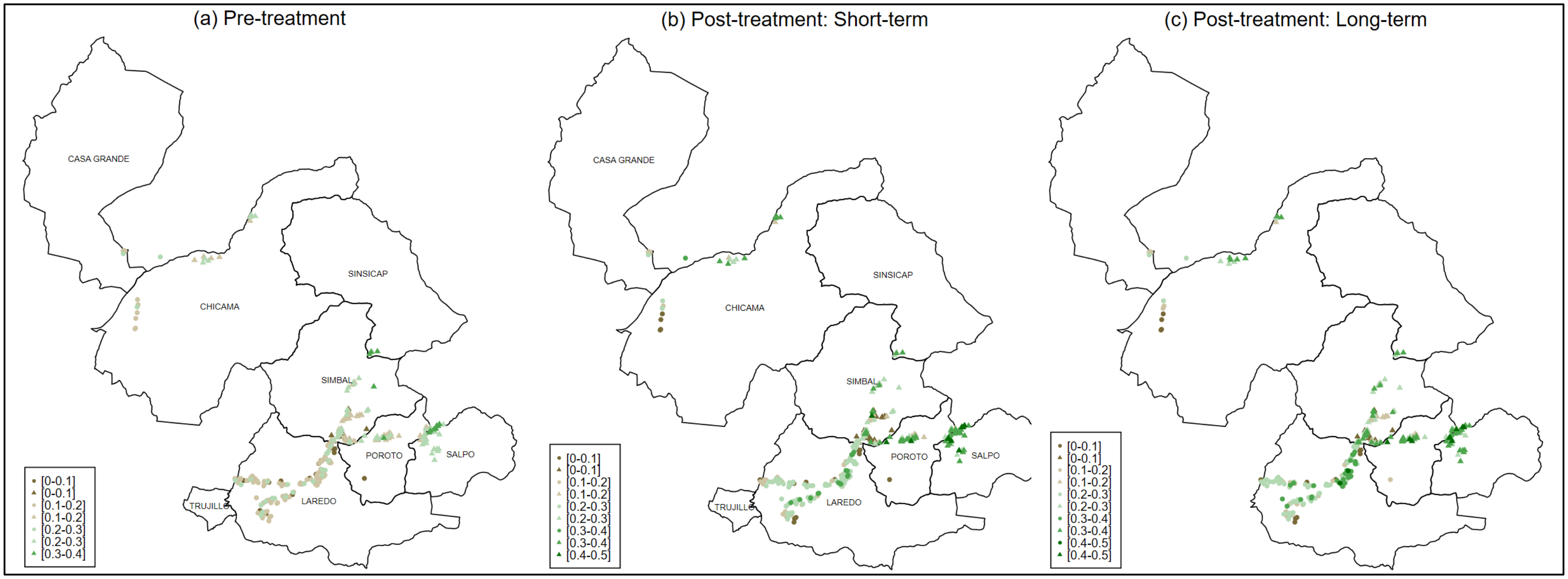
SENASA, 2009. *Erradicación de moscas de la fruta (Ceratitis capitata y Anasthrepa spp.) en las regiones de Lima, Ancash y La Libertad,* Lima: Servicio Nacional de Sanidad Agropecuaria - SENASA .

Sholihah, R. et al., 2016. Identification of Agricultural Drought Extent Based on Vegetation Health Indices of Landsat Data: Case of Subang and Karawang, Indonesia. *Procedia Environmental Sciences,* Volume 33, pp. 14-20.

Shrestha, R. et al., 2013. *Detection of flood and its impact on crops using NDVI - Corn case.* s.l., s.n., pp. 200-204.

Appendix

Figure A. 1



Note: Each dot represents a control unit. Each triangle represents a treated unit. Greener values indicate a higher NDVI value registered at that unit.

Table A. 1. Household and Farm Pre-Treatment Characteristics.

	25% obs closest to the border			50% obs closest to the border			75% obs closest to the border		
	Control	Treatment	Diff	Control	Treatment	Diff	Control	Treatment	Diff
Household characteristics									
Household head sex (male=1)	0.16	0.09	0.07	0.17	0.13	0.04	0.16	0.14	0.02
Household head age (years)	59.57	58.42	1.15	59.36	58.82	0.54	59.47	58.33	1.14
Household size	4.04	3.45	0.59*	4.14	3.82	0.32	4.14	3.96	0.18
Dependency ratio	0.49	0.60	-0.11	0.52	0.64	-0.12**	0.53	0.63	-0.10
% household heads that at most completed primary ed	0.58	0.70	-0.12	0.59	0.60	-0.01	0.61	0.55	0.06
% household heads that at most completed secondary ed	0.26	0.23	0.03	0.24	0.24	0.00	0.26	0.21	0.05
% have electricity at home	0.93	0.96	-0.03	0.90	0.94	-0.04	0.88	0.88	0.00
% have drinking water at home	0.98	0.96	0.02	0.96	0.96	0.00	0.96	0.88	0.08
% have telephone at home	0.15	0.17	-0.02	0.17	0.23	-0.06	0.18	0.18	0.00
Number of rooms	4.07	3.79	0.28	4.07	4.15	-0.08	4.04	4.11	-0.07
Farm characteristics									
% of households that have modern irrigation	0.01	0.00	0.01	0.01	0.01	0.00	0.02	0.00	0.02
Number of ha with fruit crops	0.47	0.37	0.10	0.44	0.63	-0.19*	0.45	0.54	-0.09
% of ha with fruit crops	0.58	0.56	0.02	0.54	0.65	-0.11**	0.53	0.61	-0.08**
Total number of fruit plants	161.61	176.16	-14.55	178.98	282.17	-103.19**	205.59	275.04	-69.45
Household has avocado plants installed	0.86	0.89	-0.03	0.82	0.89	-0.07*	0.80	0.86	-0.06*
Household has banana plants installed	0.45	0.53	-0.08	0.44	0.46	-0.02	0.43	0.40	0.03
Number of observations	98	53		174	125		185	246	

Source: (Salazar, et al., 2016)