European Review of Agricultural Economics Vol **47** (**5**) (2020) pp. 1920–1962
doi:10.1093/erae/jbaa015
Advance Access Publication 13 September 2020

Sweeping the flies away: evidence from a fruit fly eradication program

Lina Salazar†, Julian Aramburu‡,*, Marcos Agurto§, Alessandro Maffioli and Jossie Fahsbender†**

*†Inter-American Development Bank, Washington DC, USA ‡Yale University, New Haven, USA §Lima School of Economics, Universidad de Piura, Lima, Peru **Inter-American Development Corporation, Washington DC, USA*

Received May 2019; final version accepted May 2020

Abstract

This article evaluates the short-term impacts of a fruit fly integrated pest management program in Peru. Exploiting arbitrary variation in the program's intervention borders, we use a geographical regression discontinuity design to identify the program's effects on agricultural outcomes. Pre-treatment balance tests show that producer and farm-level pre-treatment characteristics evolve smoothly at the intervention border. Results indicate that farmers within treated areas improved pest knowledge and are more likely to implement prevention and control practices. Also, they increased fruit production and sales. Our findings are confirmed by placebo tests and are robust to alternative regression discontinuity bandwidths and polynomials.

Keywords: agricultural productivity, policy evaluation, geographic regression discontinuity, vegetable health, Peru

JEL code: H41, O12, O13, Q12, Q13, Q18

1. Introduction

Agricultural pests and diseases are major causes of economic and food losses throughout the world. Crop losses due to plant diseases have major implications for global food production and therefore, food security (Savary, Ficke and Aubertot, 2012). In fact, it is estimated that pests, diseases and weeds are responsible for yield losses that range between 10 and 40 per cent worldwide (Savary, Ficke and Aubertot, 2012; Oerke, 2006). Savary et al. (2019) estimate yield losses caused by pathogens and pests for several crops that range from 17.2 to 30 per cent in the case of wheat, rice, maize, potato and soybean.

*Corresponding author: E-mail: julian.aramburu@yale.edu

This problem is expected to intensify, as the link between climate change and the incidence of plant pests and diseases is widely documented (Adedayo et al., 2014; Lau et al., 2013). In fact, climate change favours the proliferation of certain pests (Huot et al., 2017; Evans et al., 2014; Ghini et al., 2011), and alterations in their geographic distribution around the world (Hill et al., 2016; den Bossche Van and Coetzer, 2008). In Central America, for example, higher temperatures due to climate change may have contributed to the rust epidemic caused by the Hemileia vastatrix fungus, that caused 25 per cent of coffee harvest losses in the period from 2012 to 2013, with negative impacts on other socio-economic indicators (ECLAC, 2014).

The fruit fly is one of the most harmful agricultural pests affecting horticultural crops in developing countries (Szyniszewska and Tatem, 2014). It is present in all continents, posing a threat to a wide range of commercial and native fruit production as well as farmers' livelihoods. In Africa, for example, the presence of this pest has caused crop losses that reach up to 80 per cent of fruit production and generated costs of about US$2 billion annually due to export bans, which approximately represents 2 per cent of the regional agricultural GDP (FAO, 2014, 2018). In the Pacific region, fruit losses associated with fruit flies ranged between 12 and 97 per cent for seven island countries (Allwood and Leblanc, 1996). In South America, pest-related economic losses for Brazil alone are estimated at US$242 million per year (Oliveira et al., 2013).

In this paper, we will analyse the agricultural and economic effects of a comprehensive integrated pest management (IPM) intervention that aimed to eradicate the fruit fly in the coastal region of Peru. The pest affects approximately 233,000 farmers. To evaluate the project's impact, we implement a geographic regression discontinuity design (RDD) by comparing the outcomes of farmers who reside in the neighbourhood of the intervention border. The intervention border was set arbitrarily and based on the program's budget constraints. Our results suggest that treated farmers improved their knowledge about the pest and are more likely to implement best practices for pest prevention and control. They also experienced increased agricultural sales and productivity, which are driven by higher fruit output and sales. The validity of our findings is confirmed by several placebo tests. Geographic RDD provides a novel and rigorous framework for the assessment of agricultural programs in which coverage area at a given time is restricted by budgetary or logistic constraints.

This article develops as follows: Section 2 presents a literature review on the evaluation of different IPM programs; Section 3 presents in more detail the Peruvian fruit fly IPM program; Section 4 describes the data; Section 5 presents and justifies the empirical methodology; Section 6 shows the main regression results, assess their robustness to alternative bandwidths and higher order polynomials in latitude–longitude and presents falsification tests. Finally, Section 7 concludes and discusses future research.

2. The impacts of IPM programs in the developing world: literature review

IPM refers to the group of practices that emphasise on the healthy growth of crops with the least possible disruption of their agro-ecosystems, encouraging natural pest control mechanisms that are economically viable whilst reducing negative effects on the environment and human health (FAO). The increasing demand for pest control methods that are cost-effective and environmentally friendly has led towards more integral phytosanitary interventions that reduce the level of toxicity linked to pesticide utilisation whilst increasing agricultural profits and productivity. In this context, IPM has evolved rapidly to include a mix of actions that combine the utilisation of organic pesticides, biological control techniques, quarantine facilities and technical assistance, amongst others. However, the complexity of these interventions also challenges the implementation of rigorous impact assessments that measure their causal effects on agricultural outcomes. This section presents a review of the different studies and methodologies implemented to assess the effectiveness of IPM programs.

Hruska and Corriols (2002) and Mancini, Van Bruggen and Jiggins (2007) are amongst the first studies to analyse the impact of IPM programs in Nicaragua and India, respectively. Their findings indicate that providing technical assistance in IPM to resource-poor farmers leads to a significant reduction in the number of pesticide applications and increased economic returns. Other early studies also present evidence suggesting that biological control practices can successfully deal with agricultural pests, not only in terms of suppression (i.e. establishing low pest prevalence areas) but also in terms of pest eradication (i.e. establishing pest-free areas), containment and prevention (Enkerlin, 2005; Vreysen, Hendrichs and Enkerlin, 2006; Payne et al., 2011). Rakshit et al. (2011) show that the adoption of cuelure traps to control the melon fly led to 40–130 per cent higher productivity in Bangladesh, while Pretty and Bharucha (2015) show that a compilation of IPM projects implemented in Asia and Africa have resulted in yield increases of 40.9 per cent and a decline in pesticide use of 30.7 per cent, relative to original application levels. These studies mostly rely on simple difference in means tests, before and after comparisons or case-specific qualitative studies. This limits the scope of their conclusion by issues related to self-selection into treatment as well as to the potential bias that results from unobserved confounding factors.

Several early IPM evaluation studies also use correlation and multiple regression analysis to estimate the effects of IPM interventions. For example, Escalada et al. (2009) apply parametric and non-parametric correlation techniques to analyse the impact of a variety of IPM practices using survey data from Vietnam. They show that the interventions have an immediate effect on the adoption of pest management practices in the short run but are not sustainable over time. In India, Baral et al. (2006) use linear regression and probit estimation to assess the impacts of adopting IPM practices. The authors find that IPM adoption increased yields (4.7 per cent), percentage of pest-free fruit production (34 per cent) and profits (53.8 per cent). Nevertheless, as it has been broadly discussed in the impact evaluation literature, analyses based on correlation and/or multiple regression might not be able to capture the *causal effects* of a specific intervention (Angrist and Pischke, 2009).

In recent years, modern econometric techniques—which aim to address self-selection into treatment and other endogeneity issues—such as differences in differences, propensity score matching (PSM) or a combination of both have also been used to assess the effectiveness of IPM programs. For example, Davis et al. (2012) combine a difference-in-differences approach with PSM to evaluate the effect of programs that provided IPM training through Farmer Field Schools (FFS) in Kenya, Uganda and Tanzania. They find positive impacts on agricultural income and crop productivity, and they show higher impacts for female-headed households. Moreover, the impacts also differed by farm-size and level of education. Specifically, farmers with a lower level of education and medium size land holdings had larger benefits from implementing IPM practices. Similarly, Sanglestsawai, Rejesus and Yorobe (2015) use PSM and regression analysis to evaluate the impact of an IPM program that provided training amongst onion farmers in the Philippines. The authors find a reduction on insecticide expenditures ranging from 34 to 39 per cent. Also, in Ghana, Carlberg, Kostandini and Dankyi (2014) analyse the impact of an IPM program applying a Heckman selection model to obtain consistent estimates. The authors find a positive significant effect on groundnut production for beneficiary farmers.

IPM programs have also been developed and implemented to expand the adoption of optimal management techniques that aim to reduce the incidence or eradicate the fruit fly. However, few evaluation studies have estimated the impact of these interventions. For example, Rakshit et al. (2011) use an economic surplus model to show that the adoption of pheromone traps to control the fruit fly in the sweet gourd crop led to 40–130 per cent higher productivity in Bangladesh. Later, Kibira et al. (2015) and Muriithi et al. (2016) use a difference in differences approach to evaluate the benefits of a fruit fly intervention in mango production areas in Kenya. These studies show that fruit-fly eradication based on IPM practices (e.g. use of male annihilation technique[1], application of protein bait spray, use of biopesticides, amongst others) led to a reduction on export rejections (54.5 per cent), a decrease in insecticide expenditures (46.3 per cent) and an increase in net income (22.4 per cent).

Overall, IPM programs have an important educational component that aims at improving farmer's knowledge and increasing skills to apply specific management practices. However, most of the studies have focused on analyzing agricultural outcomes such as productivity, agricultural income, exports and pesticide utilisation, placing little emphasis on the importance of measuring farmers' knowledge acquisition. This is a critical issue when considering

1 Male annihilation technique consists in the mass trapping of male flies to reduce fruit fly population density.

sustainability of IPM interventions (Peshin and Dhawan, 2009). Addressing this issue, Gautam et al. (2017) use a knowledge test score to evaluate the impact of training vegetable farmers on IPM practices in Bangladesh. Using a PSM methodology, the authors find improved knowledge about insect pests and pesticide use, as well as increased adoption of IPM practices by treated farmers. Similarly, in Peru, Gotland et al. (2004) apply a PSM to measure the impact of an IPM training program using a knowledge test. The results show that IPM training increased potato farmers' knowledge, with potential impacts on productivity.

A few studies have also addressed spillover effects that are expected to take place from social learning[2]. For example, Jørs et al. (2016) find that an IPM training program in Bolivia reduced pesticide use and improved safer handling of pesticides on treated and non-treated neighbouring farmers. The authors attribute this finding to spillover effects caused by knowledge diffusion. Also, Feder, Murgai and Quizon (2004) use a difference in differences model to evaluate the impact of an IPM program that trained farmers in Indonesia. The authors conclude that better knowledge of pest management techniques reduces pesticide use. However, in this case, there is no evidence of spillovers caused from knowledge diffusion from treated to non-treated farmers located in the same villages.

This study aims at complementing the existing literature fourfold. First, this analysis provides an impact assessment of a comprehensive IPM intervention that includes farmers training and technical assistance as well as biological control, monitoring mechanisms, quarantine facilities and application of organic pesticides. Second, program effectiveness is measured by analyzing agricultural and economic outcomes as well as farmers' knowledge. The latter was achieved by administering a knowledge test to farmers, which covered the main information provided during the fruit fly IPM training sessions. Third, this evaluation implements a geographical RDD, an innovative identification method to evaluate the causal effects of the program. This can be promising for future evaluations of phytosanitary programs. Finally, most of the studies have focused on analyzing impacts of IPM programs in Africa and Asia, whilst little evidence is found for similar programs in Latin American countries.

3. The Fruit Fly Program

According to the Peruvian National Service for Agri-food Health and Quality (SENASA), fruit flies (*Ceratitis capitata* and *Anastrepha* spp.) affect nearly 233,000 fruit producers in the coastal valleys, the country's most important horticultural area, and causes fruit production losses of approximately 30 per cent (SENASA, 2009; CENAGRO, 2012). This has important implications for the country's agriculture, as the cultivated areas allocated to fruit crops in Peru increased by 3 per cent each year between 2001 and 2016, and

2 For a comprehensive review of the social learning literature, see Mobius and Rosenblat (2014).

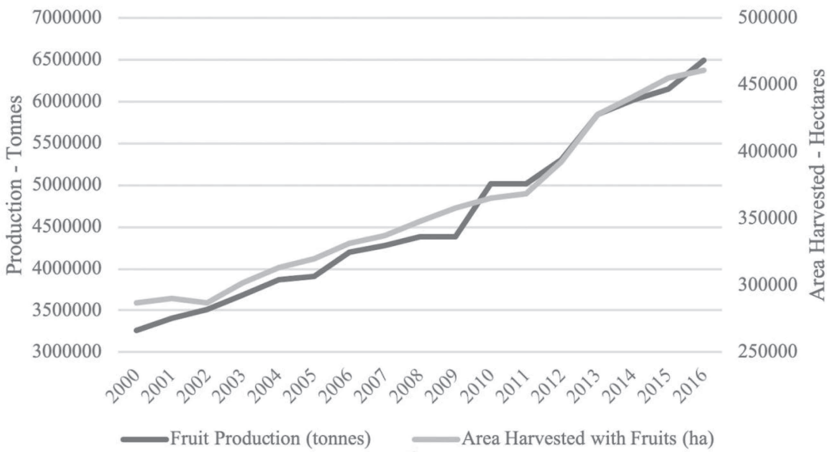


Fig. 1. Fruit production in Peru. *Source*: FAO.

fruit production experienced a 4.5 per cent annual increase over the same period (Figure 1). Moreover, fruit and vegetable exports increased at a faster rate than all exports combined for the period of 1990–2013 (FAO, 2018), representing approximately 50 per cent of all agricultural exports and 5 per cent of all exports in the country (PROMPERÚ, 2018). In this sense, this pest poses a serious risk for the consolidation of the emerging Peruvian fruit and agricultural sector, which has an enormous potential in terms of economic growth and employment opportunities (Morris et al., 2017).

Fruit flies are amongst the most damaging pests affecting fruit production in Peru. The damaging effects of the fruit fly are associated with its reproduction cycle. In particular, the female fly, which can lay up to 500 eggs, deposits its eggs inside the host crop where the larvae develop. Then, the larvae feed on fruit pulp, damaging the infested crop. After the fruit is damaged, the larvae leave the fruit and move into the soil where it is transformed into pupae (a resting stage). Finally, the flies emerge from the pupae and reproduce in approximately 4–5 days[3].

The fruit fly affects farmers by damaging agricultural production, increasing agricultural expenses due to the implementation of control measures, reducing fruit quality and value, and perhaps more critically, restricting access to international markets due to sanitary restrictions imposed to areas with presence of fruit fly. For instance, in September 2019 the European Union tightened its regulation for imports of mangoes produced in countries with the presence of fruit fly. Additional control and treatment measures will be in place, which can reduce market viability and therefore, farmer's profitability[4]. Given that

3 Further details on the reproductive cycle of the fruit fly can be found in thisarticle (SENASA, 2014).

4 "Under this regulation mangoes will only be allowed entry to the EU market if 'they have been subjected to an effective treatment to ensure freedom from Tephritidae (non-European

trade restrictions are being imposed to fruit export countries with the presence of the fruit fly, the establishment of 'pest-free' areas becomes more critical to maintain presence in the international markets and remain profitable.

In sum, the consequences of a high fruit fly prevalence are threefold. First, it restricts access to international markets, as many countries ban imports or impose strict requirements for agricultural products from infested areas. These requirements generally refer to quarantine measures; although in some cases fumigation with methyl bromide and cold treatment is also required, which costs range from US$55 to US$143 per ton of fruit, increasing export barriers in high prevalence areas. Second, pest prevalence reduces fruit productivity and significantly increases production costs. To avoid production damages in host crops, which may vary from 30 to 70 per cent (Kibira et al., 2015), producers rely on phytosanitary treatments, such as the application of insecticides, which translates into lower agricultural profits[5]. Third, the application of toxic or non-authorised pesticides and/or the indiscriminate use of pesticides increases pollution, damages natural resources such as soil and water sources and causes severe health risks for the population (Mahmood et al., 2016). For instance, a recent study by Delgado-Zegarra, Alvarez-Risco, and Yáñez (2018) points to the indiscriminate use of pesticides and non-authorised chemicals amongst farmers in Peru. The study reports the presence of non-authorised pesticides in the case of mandarin, grapes and oranges and non-authorised chemicals for the case of bananas, asparagus, grapes and oranges. Fruit fly management campaigns that lead to 'free areas' of the pest are therefore expected to have important economic, social and environmental benefits[6].

Private schemes for pest eradication, including the fruit fly, face serious limitations mainly due to externalities, coordination failures and information asymmetries (Grimsrud et al., 2008; Lansink, 2011; McKee, 2011; Hennessy and Wolf, 2018). For instance, in the case of the fruit fly, lack of information related to the proper implementation of control and prevention techniques amongst small-scale farmers, as well as the monitoring required to maintain low prevalence areas, make individual efforts unsustainable. Moreover, the externalities associated with fruit fly prevention and control create incentives to decrease individual efforts and to rely on the behaviour of nearby farmers, leading to free-rider problems. As a result, private investment is more than likely to be socially suboptimal and insufficient to achieve definitive eradication in most scenarios (Epanchin-Niell and Wilen, 2014)[7].

Given the potential benefits associated with fruit fly eradication and the difficulties related to private action, the Government of Peru initiated a fruit fly

fruit fly)'. Such effective treatment must be fully documented, including the treatment applied being recorded on the 'plant health certificate that accompanies the consignment'". Source: EPA Monimoring.

5 In Peru, it is estimated that the related costs of pesticides usage averages US$142 per hectare (INEI, 2011).

6 For a detailed discussion of optimal pest management see Epanchin-Niell and Hastings (2010).

7 Nevertheless, the provision of public programs also presents several challenges; such as those related to leakage and undercoverage (Gwatkin et al., 2005).

IPM program in the coastal area of the country. This intervention was implemented by SENASA and comprised a package of complementary activities that included: (i) farmers training in pest prevention and control; (ii) biological control practices through the release of sterile male flies; (iii) application of biopesticides; (iv) implementation of fruit fly traps to monitor the prevalence of the pest; and (v) construction of quarantine centres to monitor, detect and restrict access of contaminated fruit from untreated to treated areas.

The Peruvian fruit fly IPM program has been implemented in three phases from 1998 to 2014, covering more than one million hectares of agricultural land and 150,000 hectares of host crops (SENASA, 2015). The program started in 1998 in the most southern regions of the country (border with Chile), and, as it is shown in Figure 2, has been gradually expanded into the northern regions.

The program has been implemented in phases. In each phase, the program defined an intervention region within which all agricultural valleys were treated, as leaving untreated agricultural lands would impose serious risks in terms of pest prevalence and resurgence. Once a phase was completed in a specific location, the geographically adjacent region was identified to receive the next phase of the program. The intervention is therefore implemented by 'sweeping' adjacent geographical areas, continuously creating borders or frontiers, with treated and untreated agricultural valleys at either side of the border. In practical terms, these intervention borders define a program allocation rule which is mainly determined by budget constraints and geographical continuity.

Thus far, a total of three phases have been implemented. Phase 1 initiated in the agricultural areas adjacent to the border with Chile (green area in Figure 2) and included the agricultural valleys in the Regions of Tacna, Moquegua and Arequipa, covering 19,084 hectares of host crops and 47,015 agricultural hectares. Phase 2 (yellow area in Figure 2), implemented from 2006 to 2009, covered 40,252 hectares of host crops and 249,597 agricultural hectares. Finally, phase 3 (orange area in Figure 2), implemented from 2009 to 2014, extended the program to the northern areas of the country covering 95,381 hectares of host crops and 756,746 agricultural hectares. The purple areas in Figure 2 correspond to untreated agricultural valleys. This impact evaluation assesses the effects from the phase 3 of the program, implemented in the agricultural valleys located in Lima, Ancash and La Libertad.

4. Data

This section describes the data used in the analysis. The dataset is a panel data composed by two time points, which include information for 615 agricultural fruit producers located in 47 villages. Out of these producers, 417 are located in the treated area, and the remaining 198 in the untreated or control area[8].

8 A total of 65 households interviewed in the baseline could not be located during the follow up. This represents an attrition rate of 9.9 per cent. The attrition rate considered during the sampling design was of 10 per cent. Households that were not interviewed in the follow up are equally distributed amongst the treatment and control areas and no systematic pre-treatment

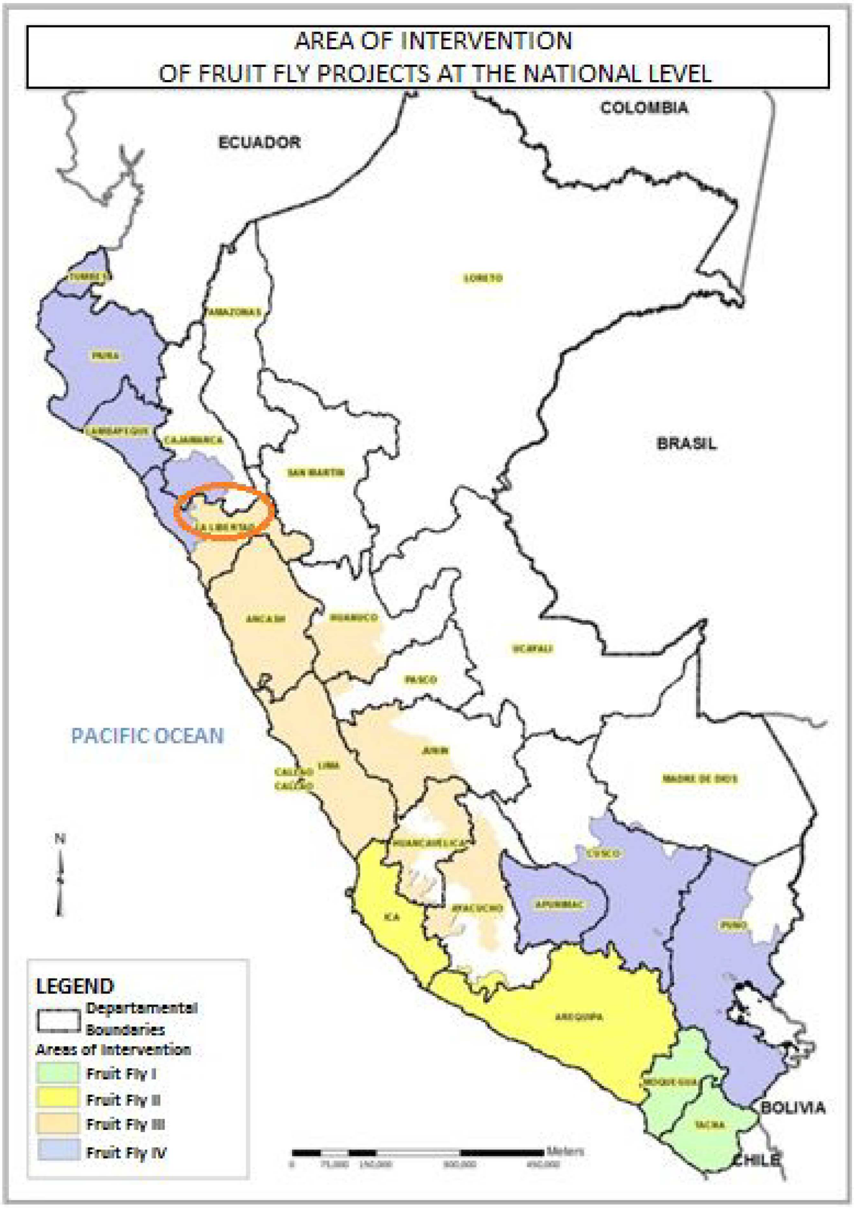


Fig. 2. The fruit fly implementation stages. *Source*: SENASA. *Notes*: Fruit Fly I, II, III and IV refer to phase I, II, III and IV of the program, respectively. During the time of analysis of this paper, phase IV of the program was only *tentatively* planned for future implementation, thus meaning that phase IV is the control area in this analysis. The orange circle represents the area of study in this paper.

differences were found between both groups. This suggests that attrition is randomly distributed and introduces no bias to the study. Attrition analysis are available upon request.

The baseline survey collected information regarding the agricultural year of 2010–2011 and was collected from May to June 2012. The baseline survey contains relevant information related to household's agricultural activities such as land size, portfolio of crops, agricultural output, agricultural inputs, agricultural income, access to credit, etc. It also captured information related to non-agricultural economic activities, household's composition, dwelling characteristics and participation in social organisations.

The follow-up survey captured information about agricultural activities conducted only during the first half of the 2014 agricultural year and was collected between June and July 2014. The questionnaire in the follow-up survey is similar to the baseline. Yet, a fruit fly knowledge test was included and administered to each farmer in the sample. This test contained 13 questions related to specific characteristics of the pest biological cycle, as well as to prevention and control measures.

4.1. Descriptive statistics

Both treatment and control groups are located within the same regional administrative unit (La Libertad). According to SENASA specialists, all the agricultural valleys included in the analysis constitute a relatively uniform geographical area. Moreover, this is the only intervention area for which the program boundaries do not coincide with administrative borders. Columns I–V in Tables 1 and 2 display the descriptive statistics for a set of relevant pre-treatment characteristics for the whole sample, as well as for those units in the treatment and control groups. Table 1 presents socio-demographic and farm characteristics, whilst Table 2 presents variables related to fruit fly prevalence and agricultural outcomes.

Regarding household characteristics, Table 1 shows that 9 per cent of the household heads are illiterate, 55 per cent have completed primary school, whilst only 21 per cent completed secondary education. Households are composed, on average, by four members, 61 per cent who are dependents[9]. Access to electricity and drinking water is relatively widespread in the study area (88 and 54 per cent, respectively), whilst access to a fixed telephone line is limited (16 per cent).

Table 1B shows that whilst almost 91 per cent of the producers have irrigation in at least one plot, only 2 per cent have modern irrigation systems. Also, farmers in the sample have 0.43 hectares (ha) dedicated to fruit crops, which represents about 56 per cent of their total land. On average, the number of fruit trees installed is 507 per household, most of which correspond to avocado (81 per cent) and banana (43 per cent).

Table 2 focuses on pre-treatment characteristics related to agricultural outcomes and fruit fly prevalence. As we can observe, 69 per cent of the farmers

9 The dependency ratio is calculated by dividing the number of household members under 15 years old or over 65 years old (dependents), in relation to the number of members aged between 15 and 65 years old (non-dependents, in working age).

Table 1. Household and Farm Pre-Treatment Characteristics

	I	II	III	IV	V	VI
	N control	*N* treated	Total average	Control average	Treated average	RDD at baseline (balance test)
A. Household Characteristics						
Household head sex (female = 1)	196	417	0.18	0.16	0.19	−0.02
			(0.38)	(0.37)	(0.39)	(0.07)
Household head age (years)	198	417	59.16	59.36	59.06	0.63
			(14.86)	(13.67)	(15.41)	(3.58)
Household size	198	417	4.11	4.14	4.10	−1.12***
			(1.92)	(2.05)	(1.86)	(0.35)
Dependency ratio	198	417	0.61	0.53	0.65	0.08
			(0.42)	(0.43)	(0.42)	(0.11)
Household head is illiterate	198	417	0.09	0.09	0.09	−0.09
			(0.29)	(0.29)	(0.29)	(0.07)
Household head completed primary school	198	417	0.55	0.62	0.52	0.13
			(0.50)	(0.49)	(0.50)	(0.11)
Household head completed high school	198	417	0.21	0.26	0.19	0.04
			(0.41)	(0.44)	(0.39)	(0.10)
Household has electricity	196	416	0.87	0.88	0.86	0.02
			(0.34)	(0.32)	(0.35)	(0.06)
Household has drinking water	198	417	0.54	0.70	0.47	0.03
			(0.50)	(0.46)	(0.50)	(0.09)
Household has phone	196	415	0.16	0.18	0.15	0.16
			(0.37)	(0.38)	(0.36)	(0.10)
Number of rooms	194	412	4.06	4.03	4.08	−0.61
			(1.58)	(1.53)	(1.61)	(0.40)

(Continued)

Table 1. Continued

	I	II	III	IV	V	VI
	N control	*N* treated	Total average	Control average	Treated average	RDD at baseline (balance test)
B. Farm characteristics						
Household has irrigation (yes = 1)	192	394	0.91	0.95	0.88	−0.04
			(0.29)	(0.21)	(0.32)	(0.06)
Household has modern irrigation (yes = 1)	190	376	0.02	0.02	0.03	−0.02
			(0.15)	(0.12)	(0.16)	(0.03)
Number of hectares with fruit crops	193	413	0.43	0.43	0.43	−0.01
			(0.74)	(0.73)	(0.74)	(0.30)
% of hectares with fruit crops	189	391	0.56	0.52	0.58	0.21*
			(0.40)	(0.40)	(0.40)	(0.11)
Total number of fruit trees	192	402	507.06	389.07	563.41	−281.27
			(1199.39)	(968.45)	(1292.57)	(431.86)
Household has avocado trees installed	198	417	0.81	0.80	0.82	0.04
			(0.39)	(0.40)	(0.38)	(0.07)
Household has banana trees installed	198	417	0.42	0.43	0.41	−0.10
			(0.49)	(0.50)	(0.49)	(0.09)
Elevation (MASL)	196	416	471.86	157.59	619.93	−94.34
			(611.06)	(101.54)	(690.12)	(122.11)

Note: Columns III, IV and V present standard deviation in parenthesis. Columns VI and VII present standard errors clustered at the group level in parenthesis.
*** and * indicate statistical significance at the 1 and 10 per cent level, respectively.

Table 2. Fruit Fly and Agricultural Production Pre-Treatment Characteristics

	I	II	III	IV	V	VI
	N control	*N* treated	Total average	Control average	Treated average	RDD at baseline (balance test)
A. Fruit Fly						
Knows or has heard about fruit fly	197	417	0.69	0.74	0.66	0.11
			(0.46)	(0.44)	(0.47)	(0.10)
Received fruit fly-related training in 2011	198	417	0.01	0.01	0.01	0.03
			(0.11)	(0.10)	(0.10)	(0.02)
Affected by fruit fly in 2011	185	333	0.70	0.78	0.65	−0.08
			(0.46)	(0.42)	(0.48)	(0.12)
Adopted fruit fly control measures in 2011	183	330	0.19	0.19	0.19	0.18
			(0.39)	(0.39)	(0.40)	(0.11)
B. Agricultural production						
Total farm output (kg)	186	405	4454.00	3728.32	4787.28	−371.04
			(6693.69)	(6236.39)	(6875.33)	(1491.61)
Total fruit output (kg)	190	409	1516.73	1306.27	1614.49	340.57
			(2682.93)	(2524.30)	(2751.04)	(622.64)
Total farm output (US$)	157	360	1438.33	1293.99	1501.29	−575.71
			(2214.71)	(2048.64)	(2283.24)	(432.50)
Value of non-fruit crops (US$/ha)	161	342	691.14	584.04	741.56	−266.79
			(1296.01)	(1027.60)	(1403.29)	(247.54)
Total farm sales (US$)	192	404	1238.29	1045.46	1329.94	−27.62
			(1895.58)	(1734.23)	(1963.05)	(433.10)
Total fruit sales (US$)	192	408	554.27	482.26	588.15	56.54
			(1176.37)	(1089.36)	(1214.97)	(286.22)
Fruit sales/total sales ratio	192	408	0.46	0.43	0.48	0.18
			(0.45)	(0.44)	(0.46)	(0.13)
Fruit crop losses (kg)	198	417	91.37	93.32	90.45	5.73
			(172.69)	(172.63)	(172.92)	(45.05)
Insecticide expenditure (US$)	198	417	18.11	15.51	19.35	−30.32***
			(50.84)	(46.30)	(52.86)	(10.16)

Note: Columns III, IV and V present standard deviation in parenthesis. Columns VI and VII present standard errors clustered at the group level in parenthesis.
***indicates statistical significance at the 1 per cent level.

Table 3. Program impacts on Fruit Fly knowledge and adoption of control–prevention measures

	I	II	III	IV	V	VI
	Knows or has heard about fruit fly (yes = 1)	Knows or has heard about fruit fly (yes = 1)	Total correct answers in fruit fly test	Total correct answers in fruit fly test	Adopted fruit fly prevention/control measures (yes = 1)	Adopted fruit fly prevention/control measures (yes = 1)
Treated	**0.19***** **(0.07)**	**0.18***** **(0.06)**	**2.22**** **(0.91)**	**2.18**** **(0.84)**	**0.34*** **(0.18)**	**0.30*** **(0.15)**
Household covariates and pre-treatment outcome value included	No	Yes	No	Yes	No	Yes
N	305	293	305	293	304	247
R^2	0.03	0.09	0.03	0.07	0.1	0.11

Note: The regressions corresponding to the odd columns in Table 3 only control for the treatment indicator and the linear polynomial in latitude–longitude. The regressions corresponding to the even columns also control for household characteristics (household size, household head age and education and the proportion of farmland with fruit crops) and the pre-treatment outcome value. Standard errors clustered at the group level appear in parenthesis. ***, ** and * indicate statistical significance at the 1, 5 and 10 per cent levels, respectively.

in the sample report that they have heard about the fruit fly; but only 1 per cent report receiving fruit fly-related training during the year 2011. Also, almost 70 per cent of the farmers report being affected by this pest during the 2011 agricultural year, and approximately 19 per cent report adopting practices for fruit fly prevention and/or control. In terms of agricultural outcomes, the average yield per farm is 4454 kg, from which 1517 kg correspond to fruit production. Households reported agricultural sales for US$1238 on average, and US$554 specifically from fruit sales (46 per cent of total sales). Average fruit crop losses are estimated at 91 kg, and farmers report an average of total expenditure on insecticides of US$18.1.

5. Empirical Methodology

In this paper, we apply a geographic RDD by exploiting the geographical border that determined program participation. As with any impact evaluation that aims to find causal effects, the focus of this analysis is the identification of a proper counterfactual. This is a group of non-treated units that are, on average, comparable to treated units in observable and unobservable characteristics. In this case, using the geographical border that determines program participation as the treatment allocation rule allows us to identify a counterfactual group that is comparable to treated farmers.

Geographic RDD is a special type of the well-known RDD approach, where the allocation variable that determines treatment status is multidimensional (latitude–longitude) (Keele and Titiunik, 2015). As noted by Lee and Lemieux (2010), as long as agents do not have precise control over the treatment allocation rule, inferences based on RDD are comparable to those of randomised experiments, which are considered the golden standard for impact evaluation (Duflo, Glennerster and Kremer, 2007). RDD is also a type of Instrumental Variable (IV) approach where the instrument is the allocation rule. Given that the allocation rule is randomly assigned, the instrument fulfills the exogeneity criteria. In this case, the geographical border that defines treatment assignment is based on the extension of land that can be covered with an assigned budget and therefore, is exogenous to program placement[10].

Geographic RDD has been widely applied in the impact evaluation literature[11]. For instance, Black (1999) exploits school district borders in the United States to assess parental valuation of school quality, whilst Dell (2010) and Dell, Lane and Querubin (2018) uses arbitrary colonial administrative borders to identify the long-term impacts of extractive institutions (*Mitas*) in Peru and state capacity in Vietnam. Mullainathan and Sukhtankar (2011) take advantage of arbitrary zones for sugarcane sales -farmers living within a zone are forced to sell their sugarcane production to the mill designated to a

10 In the case of a sharp RDD, like our case, the instrumental variable approach reduces to the case of perfect compliance.

11 Impact evaluation literature refers to analysis that aim to measure causal effects through the identification of a proper counterfactual that is comparable to treated units.

specific zone- to assess how land ownership structures influence production, credit access and consumption of farmers in India. In addition, geographic RD designs have also been applied in the field of political science where these studies exploit the geographic specificity of laws, regulations, political parties and advertisement campaigns (Gerber, Kessler and Meredith, 2011; Huber and Arceneaux, 2007; Krasno and Green, 2008; Posner, 2004). However, few examples of geographic RDD are found in impact evaluations of agricultural projects. Datar and Del Carpio (2009) exploit a geographic discontinuity comparing two areas in rural Peru affected by an irrigation project. The authors evaluate its impact on farmers' agricultural production and economic welfare. More recently, Pan, Smith and Sulaiman (2018) use the distance-to-branch threshold, which determines program eligibility for an agricultural extension project in Uganda[12], to capture the program effects on technology adoption. To the best of our knowledge, this article is the first to implement geographic RDD to identify the impacts of a comprehensive IPM program that included not only training or extension services but also additional activities that aim at pest eradication[13].

The extension of land covered by the third phase of the program, and therefore the intervention border, was selected based on: (i) geographical proximity to the area treated with the second phase of the program (Fruit Fly II in Figure 2) and (ii) budget constraints that limited the program's geographical scope. This created an artificial border that defined treatment assignment. In other words, financial resources assigned by the government limited the extension of the program as these were not enough to cover all the remaining territory after the second phase was implemented. Figures 3 and 4 show that program implementation, based on geographic continuity, creates an intervention border with treated and untreated producers at both sides of the frontier. In this sense, the program boundary defines a geographical allocation rule where producers located at one side of the border receive the treatment whilst producers located at the opposite side do not receive the program. In other words, the intervention border generates a multidimensional discontinuity in latitude and longitude, which allows us to use a RDD to identify the causal impacts of the program.

The equation that we will estimate to identify the causal impacts of the program is the following:

$$Y_{ij} = \alpha + \rho MF_j + \beta f_{ij}\,(geographic\ location) + \pi X_{ij} + \varepsilon_{ij} \qquad (1)$$

In equation (1), Y_{ij} represents the outcome variable of interest corresponding to producer$i$ in village j. The outcomes to be analysed are producers'

12 Agricultural extension programs refer to those that aim at transferring knowledge or providing training to farmers on new technologies or practices.

13 By 'comprehensive program', we refer to projects that tackle many issues related to pest eradication trough the implementation of multiple activities. These activities not only include training and extension services (as it is usually the case), but also the implementation of quarantine facilities, biological control measures, application of biopesticides, etc.

Table 4 Intervention impacts on agricultural outcomes

	I	II	III	IV	V	VI	VII	VIII	IX	X
	Log total farm output (US$)	Log total farm output (US$)	Log total farm output (kg)	Log total farm output (kg)	Log total farm sales (US$)	Log total farm sales (US$)	Log total fruit output (kg)	Log total fruit output (kg)	Log total fruit sales (US$)	Log total fruit sales (US$)
A. Agricultural production and sales										
Treated	1.65***	1.73***	1.88***	1.68***	2.40**	2.66***	1.82***	1.58***	2.06**	1.97***
	(0.54)	(0.52)	(0.51)	(0.48)	(0.97)	(0.84)	(0.50)	(0.47)	(0.94)	(0.72)
Household covariates and pre-treatment outcome value included	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	271	224	305	281	304	285	305	284	304	285
R^2	0.04	0.06	0.05	0.08	0.05	0.14	0.04	0.12	0.05	0.24
B. Fruit sales and losses										
	XI	XII	XIII	XIV	XV	XVI				
	Fruit sales/total sales ratio	Fruit sales/total sales ratio	% of fruit production losses	% of fruit production losses	Log insecticides expenditure (US$)	Log insecticides expenditure (US$)				
Treated	0.29**	0.28**	0.18	0.10	0.07	0.13				
	(0.13)	(0.11)	(0.19)	(0.18)	(0.40)	(0.36)				

(Continued)

Table 4 Continued

	I	II	III	IV	V	VI	VII	VIII	IX	X
	Log total farm output (US$)	Log total farm output (US$)	Log total farm output (kg)	Log total farm output (kg)	Log total farm sales (US$)	Log total farm sales (US$)	Log total fruit output (kg)	Log total fruit output (kg)	Log total fruit sales (US$)	Log total fruit sales (US$)
Household covariates and pre-treatment outcome value included	No	Yes	No	Yes	No	Yes				
N	304	285	305	294	305	294				
R^2	0.04	0.18	0.02	0.15	0.01	0.08				

Note: The regressions corresponding to the odd columns in Table 4 only control for the treatment indicator and the linear polynomial in lat-long. The regressions corresponding to the even columns also control for household characteristics (household size, household head age and education, and the proportion of farmland with fruit crops) and the pre-treatment outcome value. Standard errors clustered at the group level appear in parenthesis. *** and ** indicate statistical significance at the 1 and 5 per cent levels, respectively.

Table 5. Falsification tests

	I	II	III	IV	V	VI	VII	VIII
	Log total farm output BL (US$)	Log total farm output BL (US$)	Log total farm output BL (kg)	Log total farm output BL (kg)	Log total farm sales BL (US$)	Log total farm sales BL (US$)	Log total fruit output BL (kg)	Log total fruit output BL (kg)
A. Treatment impact on baseline agricultural outcomes and non-fruit agricultural outcomes								
Treated	0.15 (0.60)	0.17 (0.60)	0.65 (0.60)	0.67 (0.58)	0.60 (0.90)	0.58 (0.89)	0.74 (0.53)	0.67 (0.53)
Household covariates included	No	Yes	No	Yes	No	Yes	No	Yes
N	264	255	291	281	296	286	294	284
R^2	0.02	0.05	0.01	0.05	0.02	0.04	0.03	0.04
B. Treatment impact on baseline agricultural outcomes and non-fruit agricultural outcomes								
	IX	X	XI	XII	XIII	XIV	XV	XVI
	Log total fruit sales BL (US$)	Log total fruit sales BL (US$)	Fruit sales/total sales ratio BL	Fruit sales/total sales ratio BL	Log total non-fruit sales at Follow-up (US$)	Log total non-fruit sales at Follow-up (US$)	Log value of non-fruit crops at Follow-up (US$ per hectare)	Log value of non-fruit crops at Follow-up (US$ per hectare)
Treated	1.09 (0.87)	0.97 (0.90)	0.18 (0.13)	0.15 (0.13)	0.37 (0.44)	0.46 (0.46)	0.32 (0.52)	0.44 (0.53)
Household covariates included	No	Yes	No	Yes	No	Yes	No	Yes
N	296	286	296	286	305	294	269	262
R^2	0.05	0.05	0.05	0.08	0.01	0.08	0.01	0.06

Note: The regressions corresponding to the odd columns in Table 5 only control for the treatment indicator and the linear polynomial in latitude–longitude. The regressions corresponding to the even columns also control for household characteristics (household size, household head age and education and the proportion of farmland with fruit crops) and the pre-treatment outcome value. Standard errors clustered at the group level appear in parenthesis.

Table 6. Intervention impacts at alternative distances and polynomials

	25% of sample (8.71 km bandwidth)		75% of sample (34.08 km bandwidth)		Full sample		Excluding 5% of initial sample	
	I	II	III	IV	V	VI	VII	VIII
Score in fruit fly test	0.75	0.83	1.81**	1.41*	1.62**	1.41*	2.93***	2.68***
	(1.15)	(1.32)	(0.70)	(0.74)	(0.63)	(0.73)	(0.95)	(0.88)
Adopts fruit fly control measures	0.03	−0.01	0.37**	0.38***	0.38***	0.39***	0.42**	0.38**
	(0.25)	(0.23)	(0.14)	(0.14)	(0.13)	(0.14)	(0.21)	(0.17)
Log total farm output (US$)	1.09*	0.91	1.30**	1.23***	0.98**	1.26***	2.13***	2.11***
	(0.58)	(0.64)	(0.50)	(0.43)	(0.47)	(0.43)	(0.65)	(0.66)
Log total farm output (kg)	1.15**	1.10**	1.06**	1.10***	0.76*	1.11***	2.22***	1.89***
	(0.42)	(0.43)	(0.47)	(0.41)	(0.43)	(0.41)	(0.63)	(0.62)
Log total farm sales (US$)	0.29	−0.29	1.45**	1.64**	1.17	1.67**	3.33***	3.29***
	(0.98)	(0.89)	(0.71)	(0.74)	(0.71)	(0.74)	(0.98)	(0.85)
Log total fruit output (kg)	1.11***	0.98***	1.23***	1.26***	0.92**	1.28***	2.12***	1.73***
	(0.31)	(0.30)	(0.45)	(0.37)	(0.4)	(0.37)	(0.63)	(0.63)
Log total fruit sales (US$)	−0.26	−0.91	1.37*	1.54**	1.03	1.56**	2.83***	2.30**
	(0.81)	(0.70)	(0.69)	(0.73)	(0.68)	(0.74)	(1.04)	(0.85)
Fruit sales/total sales ratio	0.03	−0.08	0.21**	0.20*	0.16	0.20*	0.38**	0.32**
	(0.16)	(0.11)	(0.10)	(0.10)	(0.1)	(0.1)	(0.15)	(0.13)
Log total non-fruit sales (US$)	0.01	0.12	−0.02	0.14	0.16	0.17	0.22	0.30
	(0.54)	(0.58)	(0.35)	(0.37)	(0.33)	(0.37)	(0.45)	(0.4)
N	152		457		610		286	

Notes: Odd columns present linear polynomial in long-lat plus distance to capital. Even columns present quadratic polynomial in longitude–latitude plus distance to capital. Standard errors clustered at the group level appear in parenthesis. ***, ** and * indicate statistical significance at the 1, 5 and 10 per cent levels, respectively.

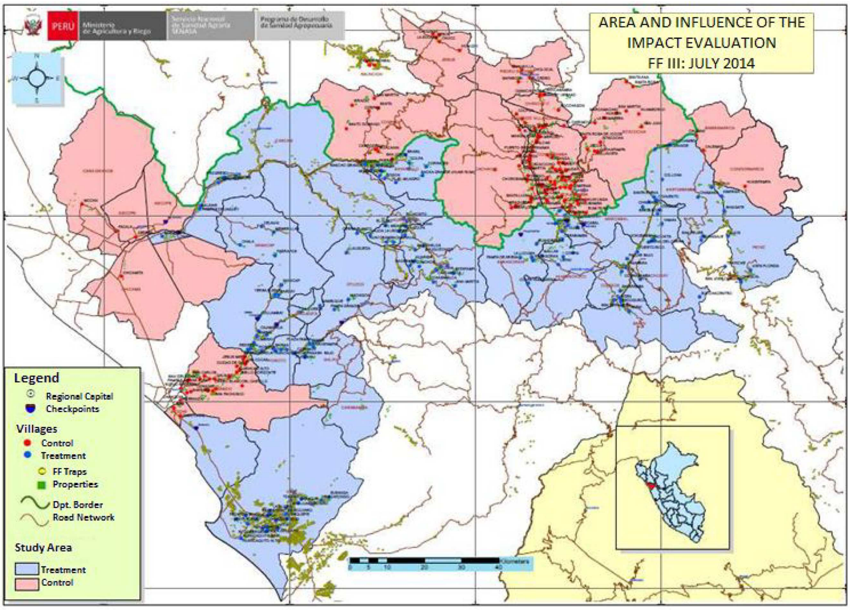


Fig. 3. The fruit fly program border—phase III. This figure shows the intervention border corresponding to the third phase of the program. The blue area depicts the treated districts, whilst the red area shows the untreated districts adjacent to the program border.

Fig. 4. Section of the intervention border corresponding to phase 3 of the fruit fly program (treated area in blue). *Notes*: Red and blue dots represent control and treatment villages, respectively, the green dots represent farmers and the yellow dots correspond to the fruit fly traps installed by SENASA.

knowledge about fruit fly, farmers' adoption of agricultural practices for pest control and prevention, value of agricultural production, value of fruit sales, agricultural income and crop portfolio decisions. The term*MF* is a dummy variable that takes the value of 1 if the producer's village is located within the treatment area and 0 otherwise. Therefore, $\boldsymbol{\rho}$ is the causal parameter of interest and captures the impact of the program. The term $f_{ij}(.)$ is the individual RDD polynomial and is a smooth function of the producer's geographic location. In an RDD design it is critical to properly capture the relation (conditional mean) of the outcome variable as a function of the running variable. The RD polynomial is intended to capture different features of this relation (Gelman and Imbens, 2019). First, since the treatment indicator is a non-linear and discontinuous function of location (longitude and latitude), it is important for the regression to control for the effect that location may have on the outcome variable *independently* of the treatment assignment. Hence, the inclusion of the geographical discontinuity polynomial is crucial to separate the treatment effect from the smooth effects of geographic location (Dell, 2010). Second, proper specification of the discontinuity polynomial avoids interpreting a non-linearity in the data structure as a treatment effect (Angrist and Pischke, 2009). In this article, various functional forms for this term are explored including lineal polynomial in latitude and longitude as well as quadratic polynomial.[14] Finally, the vector X_{ijn} represents farmers' socio-demographic and agricultural characteristics, including the baseline value of the outcomes of interest to control for initial differences at the beginning of the program (ANCOVA approach, McKenzie (2012)).

We selected the geographical RDD approach for various reasons. First, the program was implemented using a sweeping approach, which resulted on the creation of a random border (or frontier) determined by the area of land that could be treated due to budgetary constraints. Hence, the border was set up in as-good-as random fashion as it was determined by the financial resources transferred by the central government to the SENASA authorities. Also, it is worth highlighting that farmers did not intervene in the placement of such border and neither did social, economic, geographic or agricultural characteristics. Therefore, the border is exogenous to pest prevalence, productivity, income, and any other variable that could be affected by program participation. This allocation rule implied that farmers neighbouring the border were more comparable, due to geographical, climatic and socio-demographic characteristics, than farmers away from the border, resembling a randomised experiment (this will be confirmed by the balance tests performed in Section 5.1).

Second, the data were collected using different recall periods within the agricultural cycle, which makes it difficult to implement a difference in difference approach as outcome variables at the baseline and follow up are not

14 Being x and y the latitude and longitude, respectively, a simple lineal polynomial in latitude and longitude is represented by $x + y$, whilst a more complex, higher order quadratic polynomial is represented as by $x^2 + y^2 + x + y + x \bullet y$.

fully comparable. In fact, whilst the baseline collected information about the full agricultural cycle 2010–2011 (1 August 2010 to 31 July 2011), the follow up collected data for the period from January to June of 2014.[15]

Third, the ANCOVA approach, which refers to controlling for the baseline value of the outcome variable in the estimation, allows for a higher level of power in cases when the outcomes of interest have low autocorrelation[16] (McKenzie, 2012).

Finally, visual analysis of the discontinuity at the border is very important when choosing RDD. Figures 6 and 7 clearly show important 'jumps' in the variables of interest at the border of the intervention.

5.1. Validity of the geographic RDD

The presence of a border or frontier by itself does not guarantee the validity of a geographic RD design. In fact, two problems can be of concern when using this methodology: (i) endogeneity of the border placement and (ii) endogeneity of producer's location. The first problem would be troublesome if the location of the border was determined by characteristics which might affect the outcomes of interest such as pest prevalence levels or presence of host crops. However, the borders were set independently of pest incidence levels, presence of vulnerable crops or any other agricultural variable. As mentioned, the area covered by the program was determined by financial restrictions, geographic continuity (following adjacent areas to phase II), and considering natural barriers that could prevent the fruit fly entry intro treated areas, thus facilitating pest control and monitoring.

Despite this, the geography of the Peruvian territory might pose a problem to exogeneity of border placement. For instance, the border may coincide with abrupt geographical changes. Hence, agricultural characteristics, pest incidence levels and market access are unlikely to evolve smoothly at the border, invalidating the RD approach. Following Lee and Lemieux (2010), to corroborate exogeneity of border placement, later in this section we test that relevant factors evolve smoothly at the border using pre-treatment data.

The second problem refers to the possibility that farmers might migrate from untreated to treated regions in order to receive the program. However, land markets in rural areas of Peru are thin and land transactions are difficult to conduct in short-term basis. Besides, there is a strong connection between land tenure and farmers' social networks, particularly in the case of small landholders, which are the main target of the program.[17] Last and foremost,

15 For an interesting discussion on this issue we suggest thid (https://blogs.worldbank.org/impactevaluations/another-reason-prefer-ancova-dealing-changes-measurement-between-baseline-and-follow) post by David McKenzie.

16 The autocorrelation for our main outcomes of interest (i.e. total farm production and sales, fruit production and sales) is very low, between 0.10 and 0.30. According to McKenzie (2012), when there is a single baseline and single follow-up, as it is in our case, the use of difference in differences is only recommended when the autocorrelation is greater than 0.5.

17 All observations in our baseline regressions (bandwidth 17.64) come from small local rural communities denominated '*centros poblados*' (the smallest administrative rural unit in Peru), placed at relatively close distance from each other (as it can be observed in Figure 4). For

post-treatment data confirms that farmers' mobility across treatment areas was not of concern.

Another issue that needs to be addressed is the possibility of geographical spillovers. For instance, spillovers might occur if contaminated fruit is mobilised from untreated to treated areas reducing the effectiveness of the intervention. However, as part of the intervention, SENASA implemented quarantine centres that control and restrict the mobility of infected host crops to treated areas. Also, spillovers might take place through peer-learning effects. Specifically, non-beneficiary farmers located close to the border might be likely to adopt preventative and control measures due to word-of-mouth or learning through observation. Although peer-learning effects are feasible, its impact is limited as it would not be accompanied by other activities included as part of the program (i.e. monitoring traps, biopesticides and quarantine centres). In either case, the presence of spillover effects either through pest contamination or peer-learning, would cause a downward bias in the estimates, and therefore, the effects would provide a lower bound estimate of the true impact. In addition, robustness checks presented in Section 6 confirm the results after we eliminate, from the sample of analysis, producers for whom these spillovers are more likely to take place—those located within 2 km from the program boundary.[18]

This article focuses in the agricultural valleys that are immediately adjacent to the intervention border that corresponds to the third phase of the fruit fly eradication program. As shown in Figures 3 and 4, there is relatively close geographic proximity between agricultural producers located at either side of the border (in terms of distance). This fact is confirmed by GPS data, which indicate that the average distance between a given farmer located in the treatment area and his/her closest neighbour in the opposite area is of approximately 20 km (12.4 miles). The minimum distance is close to 0.4 km (0.25 miles) and the maximum is 39.5 km (24.5 miles). This implies greater degree of comparability between treatment and control. Also, importantly, the border of the intervention transects the Andes mountains, and therefore agricultural valleys at either side of the boarder are located at similar altitude ranges.

Lastly, we restrict our empirical analysis to treated and untreated areas located in the department of La Libertad. This restriction has been applied to the estimations in order to assure that administrative and political factors are similar between treatment and control groups, ruling out the possibility of compound treatments that could bias our estimates (Keele and Titiunik, 2015).

To corroborate the validity of the geographic RDD using our data, we assess the balance between treated and control groups in various pre-treatment characteristics. This follows the recent practice in the literature (Dell, Lane, and

example, whilst the *centro poblado* Pedregal is placed in the treated area with only 445 habitants, the closest *centro poblado* in the control group is Jesus Maria with 381 only habitants. These small *centros poblados*, which population's main economic activity is the agriculture, are fairly representative of the typical rural areas in the agricultural valleys covered by the program.

18 This is the 'doughnut hole' design proposed by Keele et al. (2017).

Querubin, 2018). We estimate linear RDD regressions similar to equation (1), using the baseline value as the outcome variable. Specifically, we regress the baseline values of the variables presented in Tables 1 and 2 using the geographical location and the treatment dummy as independent variables. If the border placement is exogenous, we must expect producers' characteristics to evolve smoothly at the intervention boundary. This is, after controlling for distance to the border, the treatment should *not* be a significant predictor of the outcome at the baseline. Moreover, using pre-treatment data allows us to compare baseline values for the main outcomes of interest. That is, those variables expected to be influenced by the intervention. Then, if outcome variables are balanced at the baseline for treated and control units, post-treatment changes can be causally attributed to program participation.

The RDD estimates for the treatment coefficient, presented in Column VI of Table 1, show that for all variables analysed at the baseline, only two cases present a statistically significant difference between treatment and control. These correspond to household size and the proportion of land with fruit crops (at 10 per cent significance). We address this issue by including them as covariates in our estimations. We also control for the initial value of all outcome variables, including percentage of land planted with fruits. The balancing test in Column VI of Table 2 confirms the comparability of treatment and control groups in regard to outcome variables. In fact, the only variable with a statistically significant difference is insecticide expenses, where it is found that control farmers spend US$30 more in insecticides compared to the treatment group. Reassuringly, no significant differences are found between treatment and control on the total number of fruit trees planted or the number of hectares planted with trees.

Overall, the results from the RDD estimations at the baseline confirm that control and treatment units are highly similar in terms of their household and farm characteristics, as well as in terms of their pre-treatment agricultural outcomes.[19]

6. Results

In the estimations presented in this section, we restrict our sample to units located within 17.64 km from the intervention border. Following Dell, Lane and Querubin (2018), this bandwidth has been chosen as it contains the 50 per cent closest observations (famers) to the program boundary. Later, it is also shown that the results are robust to several alternative bandwidths. Figure 5A shows the full set of observations as well as the control and treatment units that lie within and outside the 17.64 km bandwidth. Our selected bandwidth clearly defines a section of the intervention area in which treatment and control units are relatively close to each other. It is important to highlight that, in the area of intervention, the program was implemented during the second and third

19 Note that, overall, only 3 variables out of 30 pre-treatment variables included in the analysis are statistically significant.

quarters of 2013, and therefore the results should be interpreted as short-term impacts.[20]

The first step in every RDD analysis is to visually inspect the data in the neighbourhood next to the cut-off. If there is a causal impact associated with program participation, we should observe a *jump* or discontinuity in the outcome variable at the cut-off point. Results for the main outcome variables are presented in Figure 6.[21] All panels in Figure 6 present distance to the closest neighbour on the opposite side of the border (measured in km) in the *x*-axis—considering only units located at 17.64 km from the cut-off (50 per cent of the sample)—and the outcome variable in the *y*-axis. Following Calonico, Cattaneo and Titiunik (2015), the dots at the right side of the cut-off (value of zero[22]) correspond to the average value of the outcome for treated units for each binned sample means, and points to the left of the cut-off contain values for the control group.[23,24] Given the narrow bandwidth selected for our baseline analysis and following the recent practice in the RD literature (Zimmerman, 2016), we consider a linear polynomial in distance to approximate the population conditional mean functions for control and treated units.

For all the variables analysed in the panels in Figure 6, the *jump* at the cut-off indicates that the location of the household at either side of the intervention border (treatment or control) generates a discontinuity for the outcome. Indeed, compared to the control group, households located in the treated region are more likely to implement fruit fly prevention and control measures, have higher scores in the knowledge test, greater fruit sales and greater productivity measured as value of fruit production per plant.

The graphical analysis in Figure 6 also suggests that a few units in the control group, which are extremely close to the program boundary, appear to have similar outcomes than those observed in the treatment group. This might have been caused by some sort of spillover effects, as control farmers who are immediately adjacent to the border may have been able to obtain pest-related information from their neighbours or relatives in the treatment area, or managed to attend the program training sessions. Also, it is possible that the lower prevalence of the fruit fly amongst treated neighbours reduced the pest incidence in the control plots located closest to the intervention area. In Figure 7 we implement the same graphical analysis as Figure 6, but excluding those farmers located at less than 2 kilometers from the cut-off (this approach,

20 Notice that the results correspond to self-reported variables as the administered survey requested self-reported information.

21 The graphics for the rest of the variables are available upon request.

22 The variable 'distance to the closest neighbour in the opposite side of the border' has a minimum of 0.39 and a maximum of 39.25 km.

23 To provide the reader with a clear visual approximation of the data variability close to the border, and following Calonico, Cattaneo and Titiunik (2015) we used the integrated mean squared error (IMSE)-optimal evenly spaced method using spacing estimators to select the optimal number of bins. 95 per cent confidence intervals are plotted.

24 In order to construct these figures, distance for the control group is replaced by its negative value.

known as the 'doughnut hole design', has been suggested by Keele et al., 2017). In this case, the graphical analysis shows that the *jumps* in the outcome variables (as observed in the binned values) are stronger, providing some evidence on the presence of spillovers.

Given that the program allocation rule resembles a randomisation in the neighbourhood of the intervention's border (the exogeneity on the location of the program border is confirmed by the balance tests performed in Section 5.1), we can conclude that the assignment to treatment is the only factor that generates this *jump* in the outcome variable. In other words, this *jump* at the cut-off can be interpreted as the causal impact of the program on different outcomes.

We now turn to the econometric estimations of the RDD equation (1). The main results from the econometric estimation are presented in Tables 3 and 4, which confirm impacts on farmers' knowledge on fruit fly control and prevention measures as well as on agricultural outcomes[25].

Table 3 presents the RDD estimates of the program's impact on farmers' pest knowledge, as well as on farmers' adoption of prevention/control measures. One of the main activities implemented by the program is the diffusion of pest information, training and technical assistance provided by SENASA specialists. In line with this, the results confirm that treated farmers are about 18 percentage points more likely to report having knowledge about the fruit fly (24 per cent increase relative to the control group). To objectively assess the effectiveness of the technical assistance provided by the program, we designed a test in collaboration with SENASA, to evaluate the farmers' fruit fly knowledge. The estimates show that, on average, treated farmers correctly answered two more questions (50 per cent increase relative to the control group)[26] than farmers in the control group (columns III and IV).

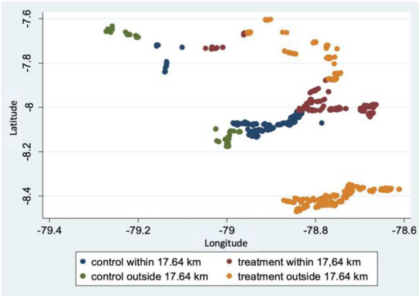
Given that producers in treated areas seem to have improved their fruit fly knowledge, it is expected that they are also more likely to implement adequate prevention and control practices in their farms. The results in Table 3 confirm that treated farmers are 30 percentage points more likely to implement pest prevention and control measures (columns V and VI) (about 157 per cent increase relative to the control group).

Thus far, the empirical results present evidence that treated farmers have improved knowledge and are more likely to implement pest prevention and control practices. In what follows, we explore whether this translated into improvements in agricultural production and income.

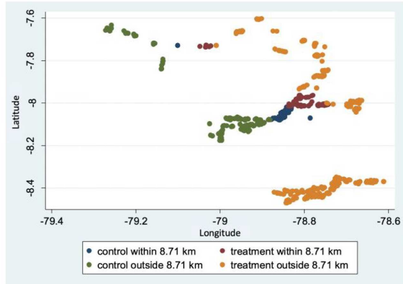
The program impacts on agricultural outcomes are presented in Table 4. Note that the results are relatively similar in the regressions with and without

25 For all the estimations, we present the estimates controlling for distance to the border using a linear latitude-longitude polynomial, household characteristics (i.e. household size, household head age and education, proportion of farmland with fruit crops) and the baseline value of the outcome variable. Similar results are obtained if instead of controlling for the proportion of land allocated to fruit crops, we control for the number of hectares of land with fruit crops and the total number of hectares.

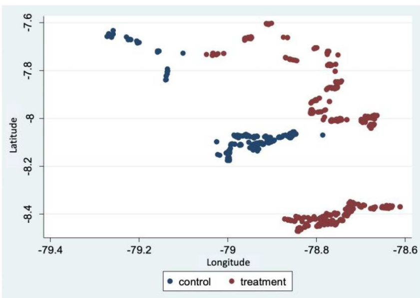
26 Mean for the control group: 4.29 correct answers (out of 13).



5.a Estimation Sample – Bandwidth 17.64 Km.



5.b Estimation Sample – Bandwidth 8.71 Km.



5.c Doughnut Hole Approach (remaining obs.)

Fig. 5. Treatment and control farmer location for different estimation samples.

household-level covariates, displayed in the odd and even columns, respectively. Specifically, relative to the control group, treated farmers experienced almost a fourfold increase (450 per cent) in total production output and total production value and a tenfold increase (1329 per cent) in total agricultural sales.[27]

To confirm that effects on total agricultural production and sales are driven by changes in fruit crops (which are the crops affected by the pest), columns VII–XII in Table 4 present the estimated treatment effects on fruit production, fruit sales and the proportion of fruit sales to total sales ratio. The results show that treated farmers experienced increased fruit production and sales, which are very similar in magnitude to those observed for total agricultural production and total sales. The results also indicate that fruit sales participation in total sales increased by 28 percentage points for treated farmers (65 per cent increase relative to the control group). Finally, no significant impacts are observed either on fruit crop losses (as a proportion of total production) or insecticide use (US$

27 These numbers were calculated using the log-level regression coefficient estimates interpretation that applies to any log-linear regression of the following type: $\log(y) = \alpha + \beta \mathit{Treatment}\, \gamma X + \varepsilon$. For this type of regressions, the formula $\%\Delta y = 100\times\ (e^{\beta} - 1)$ gives the percentage change of the outcome variable without logs.

log value).[28] The lack of effects on insecticide use may be because treated farmers could have purchased the inputs at the beginning of the agricultural cycle and prior to receiving the program (about the third trimester of 2013). Another explanation might be related to farmers' level of risk aversion or lack of information about the appropriate amount of insecticide to be applied post-treatment. Specifically, high risk-aversion and information asymmetries may cause farmers to maintain previous levels of insecticide utilisation based on pre-treatment experience.

Given that beneficiaries lacked control over the assignment variable (the intervention frontier), and that relevant pre-treatment characteristics evolve smoothly at the frontier between treated and control groups, the results obtained from the estimations confirm that the fruit fly program has been successful at improving pest knowledge, adoption of control and prevention practices, fruit production and productivity and agricultural income generated from fruit sales.[29]

6.1. Placebo (Falsification) Tests

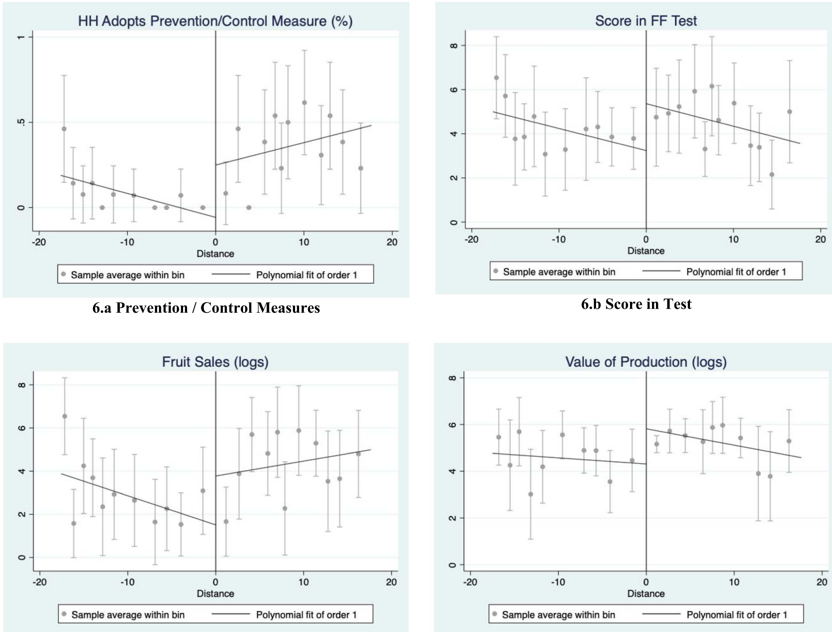
To confirm that the effects estimated capture the true causal impacts of program participation, instead of systematic differences between treatment and control areas, a series of placebo tests were implemented. Results of these tests are presented in Table 5.

First, we assess the program effects on the pre-treatment outcomes, using them as dependent variables in the RDD regressions (columns I–XII in Table 5). These must not be affected by the intervention. The results from these falsification tests confirm that the program has no impact on any of the pre-treatment outcomes included: no coefficient in Table 5 is statistically significant. Second, the RDD model is estimated using the outcomes of non-fruit crops as dependent variables (columns XIII–XVI in Table 5). Since the program only targeted fruit crops, which are the ones affected by the fruit fly, it is expected to have non-significant impacts on agricultural outcomes related to non-fruit crops. Specifically, the variables analysed are the non-fruit crop sales and the value of production of non-fruit crops per hectare. The estimations corroborate that the program had no effects on these outcomes.[30] These placebo tests confirm that the impacts found in fruit production crops are

28 Lastly, the results in Table 4 as well as those who correspond to more complex RDD polynomial specification in latitude and longitude are consistent with the ones obtained by bootstrapping the standard errors (these are available upon request). However, it is worth to acknowledge the fact that, as the comparison bandwidth narrows, the number of observations decreases. Therefore, this can increase the standard errors and decrease the precision of the point estimates. In this case, however, the significance is still present despite the reduction in the sample size.

29 For obvious reasons, the R^2 throughout all these estimations is low. This is a natural feature in the impact evaluation literature, whose main focus lies on building proper counterfactuals rather than explaining overall variability of the outcome variable.

30 The total non-fruit output in kilograms was also analysed and no significant impacts were found.



6.a Prevention / Control Measures

6.b Score in Test

6.c Fruit Sales (log)

6.d Value of Production (log)

Fig. 6. Discontinuity at the cut-off for main outcomes (bandwidth 17.64 km). *Note*: Distance in the *x*-axis refers to the distance to the closest neighbour on the opposite side of the border (in km).

capturing true causal effects instead of systematic or uncontrolled differences between treatment and control units.

6.2. Sensitivity to Alternative Bandwidths and Higher Order RDD Polynomials

As mentioned, the principal results of the RDD estimations consider a linear polynomial in latitude and longitude (Tables 3 and 4) and restrict the estimation sample to those farmers placed at a maximum distance of 17.64 km from an observation in the opposite treatment group (which represent 50 per cent of the famers).[31] In this section, it is shown that these results are robust to the consideration of alternative estimation bandwidths,[32] higher order RD polynomials and the distance to the regional capital.

31 Panel A in figure 5 shows the geographic location of these farmers in terms of latitude-longitude (treatment in red and control in blue), as well the geographic location of the treated and control farmers outside the restricted sample area (treatment in yellow and control in green).

32 Additional to the results presented in this section, an additional bandwidth is taken into consideration: Farmers located at a maximum distance of 8.71 km from a treated unit (the 25 per cent closest farmers to the intervention border). The results are also similar to those of the main specification presented in Tables 3 and 4. Available upon request.

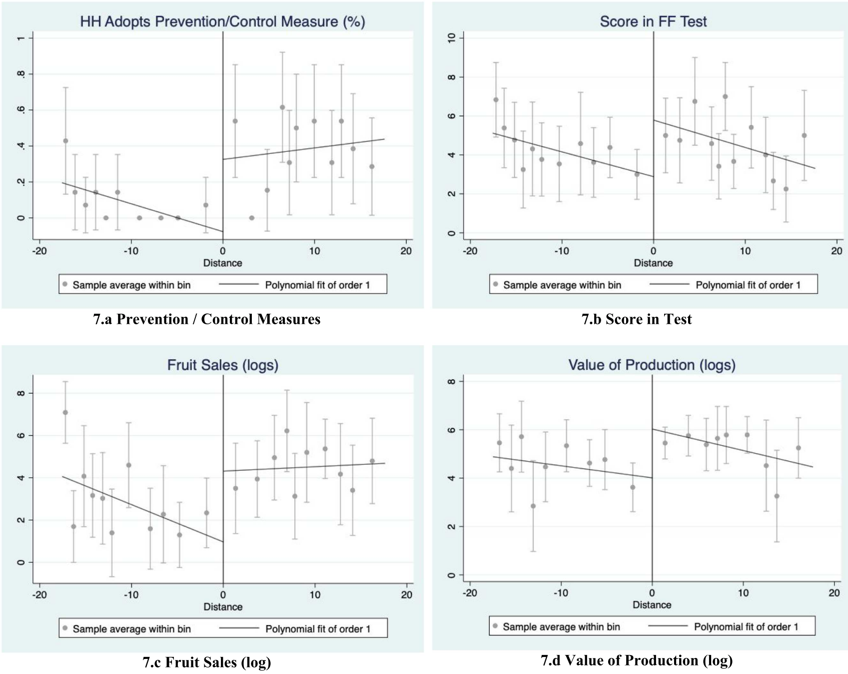


Fig. 7. Discontinuity at the cut-off for main outcomes (Doughnut hole design. Bandwidth 17.64 km). *Note*: Distance in the *x*-axis refers to the distance to the closest neighbour on the opposite side of the border (in km).

Table 6 presents estimation results considering different bandwidths and two alternative polynomial specifications. The bandwidths considered are as follows: 8.71 km. (which represents 25 per cent of the sample), 34.08 km. (75 per cent of the sample) and 39.25 km. (full sample). For the two alternative polynomial specifications, we include first the distance in km to the regional capital, which is the most important commercial and industrial centre in the study area (column I). Second, we present the results of a more flexible specification using a quadratic polynomial in latitude and longitude and also including distance to the regional capital (column II). We show that the treatment estimates obtained for all these specifications present point estimates similar in magnitude and significance to the ones presented in Tables 3 and 4, with two exceptions. First, some estimates become insignificant for the narrowest bandwidth. This might be due to the loss in power that results from considering only 25 per cent of the sample (total of 152 observations). Second, outcomes related to sales loose statistical significance when estimated using the linear polynomial specification in the full sample (column V). Besides this, these sensitivity analyses corroborate that the main results of the paper are robust to multiple bandwidths and different polynomials.

A critical assumption in any RDD is the no-interference component of the Stable Unit Treatment Value Assumption (SUTVA), which implies that '... treated units cannot interfere with control units in a way that causes the treatment to spill over and affect the control units' (Keele et al., 2017). In the case of geographic RDD, this violation is more likely to occur than in the traditional RDD, as the analysis is based on geographical proximity and location. In the intervention area under analysis, the program border was generally defined taking advantage of natural barriers. In this sense, spillover effects caused, for instance, by spraying insecticide are not likely to take place. However, spillovers related to training and pest information diffusion may cause some degree of interference, and contamination issues may arise. To address this problem, Keele et al. (2017) propose a 'doughnut hole' design, where treatment and control units closely located to the border are dropped from the sample of analysis. We follow this approach by eliminating those observations located within a 2 km distance from the intervention boundary (Table 6, Columns VII and VIII). Note that the estimates from these 'doughnut hole' design are slightly higher than those observed in columns I and II. This finding might suggest the presence of spillovers, although this claim would need to be further analysed. Nevertheless, the presence of positive spillovers might underestimate the results and therefore, provide a conservative estimate of the causal impacts.

7. Concluding remarks

In this article, we analyse the impacts of an IPM program, the fruit fly eradication program, implemented in Peru. Specifically, we exploit the geographical borders created by the intervention to identify the causal effects of the program on farmers' knowledge and agricultural outcomes using a geographic RDD approach. Given that intervention borders were set arbitrarily and that pre-treatment characteristics of treated and control units evolve smoothly at the intervention border, the results obtained using this method present the causal impacts of the program.

The findings provide evidence that the Fruit Fly Program in Peru has been successful at reaching its short-term objectives. First of all, producers in treated areas have better knowledge about the fruit fly biological cycle as well as about the measures for its prevention and control. As a result, beneficiary farmers are more likely to adopt pest management practices in their own farms. This confirms that farmers modify their behaviour to implement IPM practices. Moreover, improved knowledge and management combined with other activities implemented by the National Phytosanitary Authority (SENASA), such as trap installation and organic pesticide applications, have significantly improved fruit production and income from fruit sales. In particular, we find positive significant impacts in total fruit production, total fruit sales and proportion of fruit sold with respect to total sales. Despite all this, no significant impacts on insecticide use are found. This might suggest that farmers are risk averse and therefore, complete behavioural change requires further information about the program's benefits in the long run. In addition, further research must be conducted to understand whether program impacts are sustainable in the long run.

Also, preliminary analysis suggests the possibility of spillover effects. This might be downward biasing the main results, indicating that the estimations could present conservative impacts of the program on the variables of interest. However, this will require further analysis as the intervention was not designed to capture spillover effects. This leads us to highlight the importance of considering a careful project design in order to measure direct and indirect impacts from agricultural projects, particularly when peer-to-peer learning are expected.

The results are robust to alternative bandwidths and to the inclusion of higher order polynomials on latitude and longitude variables. They are also not likely to be driven by systematic differences between treated and control areas, as tested using placebo analysis. Also, the falsification tests confirm that impacts are due to improvements in fruit production outcomes, as no significant effects on non-fruit sales or productivity are found.

This article contributes to the literature by identifying the causal impacts of a comprehensive IPM program on small-scale farmers. Commonly, it is believed that phytosanitary programs mainly benefit large scale exporters of agricultural products. This analysis confirms that small-scale farmers, who are not necessarily connected to the international market, can also experience significant benefits from this type of programs.

References

Adedayo, A. M., Olafiranye, O., Smith, D., Hill, A., Zizi, F., Brown, C. and Jean-Louis, G. (2014). Obstructive sleep apnea and dyslipidemia: Evidence and underlying mechanism. *Sleep and Breathing* 18(1): 13–18.

Allwood, A. J. and Leblanc, L. (1996). Mango fruit fly (Bactrocerafrauenfeldi): Why so many in Federated States of Micronesia. In *ACIAR Proceedings of Regional Symposium: Management of Fruit Flies in the Pacific, Brown, Prior & Anderson, Melbourne*, 125–130.

Angrist, J. and Pischke, J. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ, USA: Princeton University Press.

Baral, K., Roy, B. C., Rahim, K. M. B. *et al.* (2006). *Socio Economic Parameters of Pesticide Use and Assessment of Impact of an IPM Strategy for the Control of Eggplant Fruit and Shoot Borer in West Bengal, India (No. 05103)*. Tainan, Taiwan: AVRDC.

Black, S. (1999). Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics* 114(2): 577–599.

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association* 110(512): 1753–1769.

Carlberg, E., Kostandini, G. and Dankyi, A. (2014). The effects of integrated Pest management techniques farmer field schools on groundnut productivity: Evidence from Ghana. *Quarterly Journal of International Agriculture* 53(1): 73–88.

CENAGRO – Censo Nacional Agropecuario. (2012). *Resultados Definitivos: IV Censo Nacional Agropecuario – 2012*. Peru: Lima.

Davis, K., Nkonya, E., Kato, E. *et al.* (2012). Impact of farmer field schools on agricultural productivity and poverty in East Africa. *World Development* 40(2): 402–413.

Datar, G. and Del Carpio, X. (2009). *Are Irrigation Rehabilitation Projects Good for Poor Farmers in Peru?* The World Bank, Policy Research Working Paper 5154.

Dell, M. (2010). The persistent effects of Peru's mining mita. *Econometrica* 78(6): 1863–1903.

Dell, M., Lane, N. and Querubin, P. (2018). The historical state, local collective action, and economic development in Vietnam. *Econometrica* 86(6): 2083–2121.

den Bossche Van, P. and Coetzer, J. A. (2008). Climate change and animal health in Africa. *Revue scientifique et technique (International Office of Epizootics)* 27(2): 551–562.

ECLAC. (2014). *Economic and Social Panorama of the Community of Latin American and Caribbean States*. Chile: United Nations Economic Commission for Latin America and the Caribbean (ECLAC), Santiago, 2013.

Enkerlin, W. R. (2005). Impact of fruit fly control programmes using the sterile insect technique. In *Sterile Insect Technique*. Netherlands: Springer, 651–676.

Escalada, M. M., Heong, K. L., Huan, N. H. *et al.* (2009). *Changes in rice farmers' pest management beliefs and practices in Vietnam: an analytical review of survey data from 1992 to 2007. Planthoppers: new threats to the sustainability of intensive rice production systems in Asia*, 447–456.

Evans, J. P., Ji, F., Lee, C., Smith, P., Argüeso, D. and Fita, L. (2014). Design of a regional climate modelling projection ensemble experiment–NARCliM. *Geoscientific Model Development* 7(2): 621–629.

Feder, G., Murgai, R. and Quizon, J. B. (2004). Sending farmers back to school: The impact of farmer field schools in Indonesia. *Applied Economic Perspectives and Policy* 26: 45–62.

Epanchin-Niell, R. S. and Hastings, A. (2010). Controlling established invaders: Integrating economics and spread dynamics to determine optimal management. *Ecology Letters* 13(4): 528–541.

Epanchin-Niell, R. S. and Wilen, J. E. (2014). Individual and cooperative management of invasive species in human-mediated landscapes. *American Journal of Agricultural Economics* 97(1): 180–198.

Delgado-Zegarra, J., Alvarez-Risco, A. and Yáñez, J. A. (2018). Uso indiscriminado de pesticidas y ausencia de control sanitario Para el mercado interno en Perú. *Revista Panamericana de Salud Pública* 42: e3.

Duflo, E., Glennerster, R. and Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of Development Economics* 4: 3895–3962.

FAO. (2014). *Four in one – new discovery on pest fruit flies, Link to article*.

FAO. (2018). *Global body adopts new measures to stop the spread of plant pests, Link to article*.

Ghini, R., Bettiol, W. and Hamada, E. (2011). Diseases in tropical and plantation crops as affected by climate changes: Current knowledge and perspectives. *Plant pathology* 60(1): 122–132.

Gwatkin, D. R., Wagstaff, A. and Yazbeck, A. S. (2005). *Reaching the poor with health, nutrition and population services: What works, what doesn't, and why*. Washington DC, USA: The World Bank.

Gautam, S., Schreinemachers, P., Uddin, M. N. *et al.* (2017). Impact of training vegetable farmers in Bangladesh in integrated pest management (IPM). *Crop Protection* 102: 161–169.

Gerber, A. S., Kessler, D. P. and Meredith, M. (2011). The persuasive effects of direct mail: A regression discontinuity based approach. *Journal of Politics* 73(01): 140–155.

Grimsrud, K. M., Chermak, J. M., Hansen, J. *et al.* (2008). A two-agent dynamic model with an invasive weed diffusion externality: An application to yellow Starthistle (Centaurea solstitialis L.) in New Mexico. *Journal of Environmental Management* 89(4): 322–335.

Gotland, E. M., Sadoulet, E., de Janvry, A. *et al.* (2004). The impact of farmer filed schools on knowledge and productivity: A study of potato farmers in the Peruvian Andes. *Economic Development and Cultural Change* 53: 63–92.

Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics* 37(3): 447–456.

Hennessy, D. A. and Wolf, C. A. (2018). Asymmetric information, externalities and incentives in animal disease prevention and control. *Journal of Agricultural Economics* 69(1): 226–242.

Hill, M. P., Bertelsmeier, C., Clusella-Trullas, S. *et al.* (2016). Predicted decrease in global climate suitability masks regional complexity of invasive fruit fly species response to climate change. *Biological Invasions* 18(4): 1105–1119.

Hruska, A. J. and Corriols, M. (2002). The impact of training in integrated pest management among Nicaraguan maize farmers: Increased net returns and reduced health risk. *International Journal of Occupational and Environmental Health* 8(3): 191–200.

Huber, G. A. y Arceneaux, K. 2007. Identifying the persuasive effects of presidential advertising. *American Journal of Political Science* 51 (4): 957–977.

Huot, B., Castroverde, C. D. M., Velásquez, A. C., Hubbard, E., Pulman, J. A., Yao, J. *et al.* (2017). Dual impact of elevated temperature on plant defence and bacterial virulence in Arabidopsis. *Nature communications* 8(1): 1–12.

INEI – Instituto Nacional de Estadística e Informática. (2011). *Perú: Anuario de Estadísticas Ambientales 2011*. Peru: Lima.

Jørs, E., Konradsen, F., Huici, O., Morant, R. C., Volk, J. and Lander, F. (2016). Impact of training Bolivian farmers on integrated pest management and diffusion of knowledge to neighboring farmers. *Journal of Agromedicine* 21(2): 200–208.

Keele, L. J. and Titiunik, R. (2015). Geographic boundaries as regression discontinuities. *Political Analysis* 23(1): 127–155.

Keele, L., Lorch, S., Passarella, M. *et al.* (2017). An Overview of Geographically Discontinuous Treatment Assignments with an Application to Children's Health Insurance. In *Regression Discontinuity Designs: Theory and Applications*. Bingley, United Kingdom: Emerald Publishing Limited, 147–194.

Kibira, M., Affognon, H., Njehia, B. *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): 343–353.

Krasno, J. S.y. and Green, D. P. (2008). Do televised presidential ads increase voter turnout? Evidence from a natural experiment. *Journal of Politics* 70(1): 245–261.

Lansink, A. O. (2011). Public and private roles in plant health management. *Food Policy* 36(2): 166–170.

Lau, K. Y., Vaughan, A. S., Chen, G., Hosier, I. L. and Holt, A. F. (2013). On the dielectric response of silica-based polyethylene nanocomposites. *Journal of Physics D: Applied Physics* 46(9): 095303.

Lee, D.y. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature* 48(2): 281–355.

Mahmood, I., Imadi, S. R., Shazadi, K. *et al.* (2016). Effects of pesticides on environment. In *Plant, Soil and Microbes*. Cham: Springer, 253–269.

McKee, G. J. (2011). Coordinated pest management decisions in the presence of management externalities: The case of greenhouse whitefly in California-grown strawberries. *Agricultural Systems* 104(1): 94–103.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of development Economics* 99(2): 210–221.

Mancini, F., Van Bruggen, A. H. and Jiggins, J. L. (2007). Evaluating cotton integrated pest management (IPM) farmer field school outcomes using the sustainable livelihoods approach in India. *Experimental Agriculture* 43(1): 97–112.

Morris, M., Diaz, L., Sebastian, A. *et al.* (2017). *Gaining momentum in Peruvian agriculture: opportunities to increase productivity and enhance competitiveness*. Washington, D.C.: World Bank Group.

Muriithi, B. W., Affognon, H. D., Diiro, G. M. *et al.* (2016). Impact assessment of integrated Pest management (IPM) strategy for suppression of mango-infesting fruit flies in Kenya. *Crop Protection* 81: 20–29.

Oliveira, C. M., Auad, A. M., Mendes, S. M. *et al.* (2013). Economic impact of exotic insect pests in Brazilian agriculture. *Journal of Applied Entomology* 137: 1–15.

Oerke, E. C. (2006). Crop losses to pests. *The Journal of Agricultural Science* 144: 31.

Mullainathan, S. and Sukhtankar, S. (2011). *Ownership Structure and Economic Outcomes: The Case of Sugarcane Mills in India, Working Paper Series*. London, United Kingdom: International Growth Centre.

Mobius, M. and Rosenblat, T. (2014). Social learning in economics. *Annual Review of Economics* 6(1): 827–847.

Pan, Y., Smith, S. C. and Sulaiman, M. (2018). Agricultural extension and technology adoption for food security: Evidence from Uganda. *American Journal of Agricultural Economics* 100(4): 1012–1031.

Payne, W., Tapsoba, H., Baoua, I. B. *et al.* (2011). On-farm biological control of the pearl millet head miner: Realization of 35 years of unsteady progress in Mali, Burkina Faso and Niger. *International Journal of Agricultural Sustainability* 9(1): 186–193.

Peshin, Rajinder, and Dhawan, Ashok K. (eds). (2009). *Integrated Pest Management: Volume 2: Dissemination and Impact*, Vol. 2. Berlin, Germany: Springer Science & Business Media.

Pretty, J. and Bharucha, Z. P. (2015). Integrated pest management for sustainable intensification of agriculture in Asia and Africa. *Insects* 6(1): 152–182.

PROMPERÚ. (2018). *Departamento de Agronegocios de la Sub Dirección de Promoción Internacional de la Oferta Exportable, Link to report*. Lima, Peru.

Posner, D. N. (2004). The political salience of cultural difference: Why Chewas and Tumbukas are allies in Zambia and adversaries in Malawi. *The American Political Science Review* 98(4): 529–545.

Rakshit, A., Karim, A. R., Hristovska, T. *et al.* (2011). Impact assessment of pheromone traps to manage fruit fly on sweet gourd cultivation. *Bangladesh Journal of Agricultural Research* 36(2): 197–203.

Sanglestsawai, S., Rejesus, R. M. and Yorobe, J. M. (2015). Economic impacts of integrated pest management (IPM) farmer field schools (FFS): Evidence from onion farmers in the Philippines. *Agricultural Economics* 46(2): 149–162.

Savary, S., Ficke, A., Aubertot, J. N. and Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Security* 4: 519–537.

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N. and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature ecology & evolution* 3(3): 430–439.

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

SENASA – Servicio Nacional de Sanidad Agropecuaria. (2014). *Biología y comportamiento de la mosca de la fruta*Unpublished Manuscript. Lima, Peru.

SENASA – Servicio Nacional de Sanidad Agropecuaria. (2015). *Memoria Institucional 2015*Unpublished Manuscript. Peru: Lima.

Szyniszewska, A. M. and Tatem, A. J. (2014). Global assessment of seasonal potential distribution of Mediterranean fruit fly, Ceratitis capitata (Diptera: Tephritidae). *PLoS One* 9(11): e111582.

Vreysen, M. J., Hendrichs, J. and Enkerlin, W. R. (2006). The sterile insect technique as a component of sustainable area-wide integrated pest management of selected horticultural insect pests. *Journal of Fruit and Ornamental Plant Research* 14: 107.

Zimmerman, S. D. (2016). *Making the one percent: The role of elite universities and elite peers (no. w22900)*. National Bureau of Economic Research.

Appendix

Annex 1: Impacts in Avocado and Banana Production

In order to obtain more insights on the channels that explain increasing fruit sales value, Table A1 explores the program effects on the two most common fruit crops in the area: avocado and banana. The results analysed correspond to the 17.64 km bandwidth (50 per cent closest observations to the treatment border), where we control for a linear polynomial in distance to the border. As we can observe, treated households have increased their production of avocado and banana and in the case of avocado, total sales also increased.

One possible explanation for this finding is that as a result of the program, local buyers have increased their demand for fruit produced in the treated areas, which have had a positive effect on prices. Another complementary explanation is that treated farmers are able to provide a fruit of better quality, which has a lower risk of being contaminated by fruit fly larvae, and therefore is priced higher.

Annex 2: Variables included in the analysis

B.1. Household characteristics variables

- Household head sex: dummy that equals 1 if the head of the household is a female, 0 otherwise.
- Household head age: age of the head of the household, in years.
- Household size: number of people living in the household.
- Dependency ratio: calculated for each household as the following fraction

$$\frac{\text{Number of people aged } 0-14 + \text{number of people aged 65 and over}}{\text{Total members of the household}}$$

- Household head is illiterate: dummy that takes the value of 1 if head of the household declares to be illiterate, 0 otherwise.
- Household head completed primary school: dummy that takes the value of 1 if the head of the household declares to have an educational attainment of at most primary, 0 otherwise.
- Household head completed high school: dummy that takes the value of 1 if the head of the household declares to have an educational attainment of at most high school, 0 otherwise.
- Household has electricity: dummy that takes the value of 1 if the household has electricity service, 0 otherwise.
- Household has drinking water: dummy that takes the value of 1 if the household has drinking water service, 0 otherwise.

Table A1 Program effect over avocado and banana production (18 km bandwidth—50 per cent closest observations to the border)

	Avocado			Banana		
	I	II	III	IV	V	IV
	Log total output (Kg)	Log total sales (US$)	Price per kilo sold (US$)	Log total output (Kg)	Log total sales (US$)	Price per kilo sold (US$)
Treated	1.85***	2.21***	0.17	1.68**	0.45	0.01
	(0.60)	(0.73)	(0.60)	(0.79)	(0.54)	(0.09)
N	267	267	267	121	121	121
R2	0.04	0.05	0.01	0.05	0.04	0.01

Note: Standard errors clustered at the group level appear in parenthesis. ***, ** and * indicate statistical significance at the 1, 5 and 10 per cent levels, respectively.

- Household has phone: dummy that takes the value of 1 if there is at least 1 phone in the household, 0 otherwise.
- Number of rooms: number of rooms in the dwelling.

B.2. Farm characteristics variables

- Household has irrigation: dummy that takes the value of 1 if the household declares to have any type of irrigation system in at least one of the parcels worked, 0 otherwise.
- Household has modern irrigation: dummy that takes the value of 1 if the household declares to have any type of modern irrigation system (sprinkler or drip) in at least one of the parcels, 0 otherwise.
- Number of hectares with fruit crops: total number of hectares with fruit plants.
- Total number of fruit trees: total number of fruit trees that the head of the household reports to have in their parcels.
- Household has avocado trees installed: dummy that takes the value of 1 if the head of the household declares to have at least one avocado tree installed in their land, 0 otherwise.
- Household has banana trees installed: dummy that takes the value of 1 if the head of the household declares to have at least one banana tree installed in their land, 0 otherwise.
- Elevation (MASL): elevation measured as Meters Above the Sea Level, obtained by GPS measures at the entrance of farmer's main parcel.

B.3. Fruit Fly variables

- Knows or has heard about fruit fly: dummy that takes the value of 1 if the head of the household declares to know or have ever heard about the existence of the fruit fly, 0 otherwise.
- Received fruit fly-related training: dummy that takes the value of 1 if the head of the household declares to have received any type of agricultural training related to the fruit fly, 0 otherwise.
- Affected by fruit fly: dummy that takes the value of 1 if the head of the household declares to have at least one fruit plant/crop affected by the fruit fly, 0 otherwise.
- Adopted fruit fly prevention/control measures: dummy that takes the value of 1 if the head of the household declares to have implemented any type of fruit fly prevention or control measure, 0 otherwise.

- Total correct answers in fruit fly test: sum of the total number of correct answers in the fruit fly knowledge test (min = 0, max = 14).

B.4. Agricultural production variables

- Log total farm output (kg): total farm output is calculated by adding the weight, in kilos, of sold and unsold production for all crops in all parcels that the household works, for the whole agricultural cycle considered. Production reported in a unit different than kilo is converted to its equivalent in kilos. Production is self-reported. The variable is top-coded at the 95th percentile. Log is calculated after adding 1 to the value of farm output.
- Log total farm output (US$): total farm output is calculated by adding the value of sold and unsold production for all crops in all parcels that the household works, for the whole agricultural cycle considered. For production sold, the value is obtained by multiplying the kilos sold by the price obtained. For production unsold (destined to self-consumption, exchange), value is calculated by multiplying production in kilos by the median price per kilo for that crop at the *centro poblado* (town). Production reported in a unit different than kilo is first converted to its equivalent in kilos, and prices are converted to price-per-kilo. Both production and prices are self-reported. The variable is top-coded at the 95th percentile. Log is calculated after adding 1 to the value of farm output.
- Log total farm sales (US$): total farm sales are calculated by adding the value of sold production for all crops in all parcels that the household works, for the whole agricultural cycle considered. The value is obtained by multiplying the kilos sold by the price obtained. Production reported in a unit different than kilo is first converted to its equivalent in kilos, and prices are converted to price-per-kilo. Both production and prices are self-reported. The variable is top-coded at the 95th percentile. Log is calculated after adding 1 to the value of farm sales.
- Log total fruit output (kg): total fruit output is calculated by adding the weight, in kilos, of sold and unsold production for fruit crops in all parcels that the household works, for the whole agricultural cycle considered. Production reported in a unit different than kilo is converted to its equivalent in kilos. Production is self-reported. The variable is top-coded at the 95th percentile. Log is calculated after adding 1 to the value of fruit output.
- Log total fruit sales (US$): total fruit sales are calculated by adding the value of sold production for fruit crops in all parcels that the household works, for the whole agricultural cycle considered. The value is obtained by multiplying the kilos sold by the price obtained. Production reported in a unit different than kilo is first converted to its equivalent in kilos, and prices are converted to price-per-kilo. Both production and prices are self-reported. The variable is top-coded at the 95th percentile. Log is calculated after adding 1 to the value of fruit sales.
- Fruit sales/total sales ratio: this is the quotient between total fruit sales (US$) and total farm sales (US$).

- Fruit crop losses (kg): total fruit losses are calculated by adding the weight, in kilos, of lost production for fruit crops in all parcels that the household works, for the whole agricultural cycle considered. Production lost is self-reported. The variable is top-coded at the 95th percentile. Log is calculated after adding 1 to the value of lost production.
- Percentage of fruit production losses: this is the quotient between fruit crop losses (kg) in the numerator and fruit crop losses (kg) plus total fruit output (kg) in the denominator.
- Log insecticides expenditure (US$): total insecticides expenditure is calculated by adding the value of expenditure in insecticides for all insecticides bought in the whole agricultural cycle. This is obtained by multiplying quantity of all insecticides bought by price paid, all in the same unit of measure (either kilos or liters). The variable is top-coded at the 95th percentile. Log is calculated after adding 1 to the value of insecticides expenditures.
- Log total non-fruit sales (US$): total non-fruit sales are calculated by adding the value of sold production for non-fruit crops in all parcels that the household works, for the whole agricultural cycle considered. The value is obtained by multiplying the kilos sold by the price obtained. Production reported in a unit different than kilo is first converted to its equivalent in kilos, and prices are converted to price-per-kilo. Both production and prices are self-reported. The variable is top-coded at the 95^{th} percentile. Log is calculated after adding 1 to the value of non-fruit sales.
- Log value of non-fruit crops (US$ per hectare): this is calculated first by obtaining the value of production per hectare per crop for non-fruit crops, and then adding this across these crops. Value of production per hectare per crop is obtained first by adding the value of sold and unsold production for that crop in all parcels that the household works, for the whole agricultural cycle considered. For production sold, the value is obtained by multiplying the kilos sold by the price obtained. For production unsold (destined to self-consumption, exchange), value is calculated by multiplying production in kilos by the median price per kilo for that crop at the *centro poblado* (town). Production reported in a unit different than kilo is first converted to its equivalent in kilos, and prices are converted to price-per-kilo. Both production and prices are self-reported. Then, this measure is divided by the total hectares worked with that crop in the whole agricultural cycle. The variable is top-coded at the 95th percentile. Log is calculated after adding 1 to the value of non-fruit crops.

B.5. Geographic variables

- Latitude: latitude, in decimal degrees, obtained by GPS measures at the entrance of farmer’s main parcel.
- Longitude: longitude, in decimal degrees, obtained by GPS measures at the entrance of farmer’s main parcel.

- Distance to the closest neighbour in the opposite group: Euclidean distance, measured in kilometers, to the closest farmer in the opposite treatment group. This distance is calculated taking the latitude and longitude as inputs.