FONDATION POUR LES ÉTUDES ET RECHERCHES SUR LE DÉVELOPPEMENT INTERNATIONAL

Working Paper 43 March 2012 Development Policies

A Second Look at the Pesticides Initiative Program: Evidence from Senegal*

Melise Jaud
Olivier Cadot

Melise Jaud, The World Bank.

Olivier Cadot, University of Lausanne, Ferdi, Cepr and Cepremap..

Abstract

This paper investigates whether the Pesticides Initiative Program has signicantly affected the export performance of Senegal's horticulture industry. We apply two main microeconometric techniques, difference-in-differences and matching difference-in-differences, to identify the effect of the Pesticides Initiative Program on exports of fresh fruits and vegetables. We use a unique firm-level dataset containing data on sales, employment, and exports by product and destination markets, as well as firm enrolment year, over 2000-2008. The results suggest that while the program had no signicant effect on exports pooled over all products and destinations, it had a positive effect when considering fresh fruits and vegetables exports to the European Union.

Keywords: Senegal, EU, Agriculture, Export Promotion, Technical Assistance, Impact Evaluation.
JEL classication codes: F13, L15, L25, O13, C23.

* This paper is part of a joint program of Switzerland's NCCR and the World Bank on the application of impact-evaluation methods to trade-related interventions. Cadot worked on this paper while at the World Bank's Trade Division. Without implicating them, we would like to express our thanks to Celine Carrere, Ana Fernandes, Julien Gourdon, Jaime de Melo, Alberto Portugal, Akiko Suwa, two anonymous referee, and Jo Swinnen, editor of the WTR's special issue, for useful comments. We are also grateful to the European Commission and to Senegal's *Direction Générale des Douanes* and *Agence Nationale de la Statistique et de la Démographie* for access to their data.

1 Introduction

Technical assistance (TA) is, for the E.U., an important component of Aid for Trade (AfT), responding to the need to help the private sector in Southern partners, in particular low-income ones, cope with increasingly stringent SPS and product standards. Among the many forms of assistance counted as AfT (see WTO/OECD 2011), TA may also be, together with export promotion, one of the most genuinely trade-focused ones. Yet, we don't know much about its impact and effectiveness.

In general, rigorous studies on the effectiveness of AfT are few. As for its allocation, Gamberoni and Newfarmer (2009) found that, after controlling for absorption capacity, it correlated with an indicator of 'demand' based on indices of under-trading. As for its impact, the early literature looked for effects on *donor* trade, a natural choice since a good chunk of aid was tied. Wagner (2003) found donor exports to the recipient country were indeed boosted by aid. By contrast, using a gravity equation in first differences on four European donors and 26 African recipients, Osei, Morrissey and Lloyd (2004) found an unstable and, on the whole, insignificant impact of aid on trade flows from donor to recipient. Nelson and Juhasz Silva (2008) used a more conventional gravity equation including bilateral aid flows as a regressor (instrumented by their one-year lagged value) and found a significant although small elasticity, again, for flows from donor to recipient.

Only some of the most recent studies have looked at whether aid raised the export capacity of *recipient* countries. Cali and te Velde (2009) regressed trade costs and the value of exports on control variables and lagged AfT disbursements, using data from the OECD's Credit Reporting System, on a panel of countries covering 1995-2007. Their various specifications together suggest that aid to trade facilitation and infrastructure seems to have a significant effect (the former on trade costs, the latter on export values), while aid to productive capacity is insignificant. The same results hold when looking at sectorally-targeted aid and controlling for country-sector fixed effects. Brenton and von Uexkull (2009) combined mirrored product-level (HS4) export data with a long panel of export-development aid data from the GTZ and from the OECD/WTO Trade Capacity Building Database for 48 countries. They used a difference-in-differences approach regressing exports on lagged exports, country and year-product fixed effects, and contemporaneous and lagged aid coded in binary form. By and large, once outliers were dropped, propensity-score matching yielded insignificant coefficients, suggesting that once selection effects were taken into account, export-development programs provided little significant boost to exports. By contrast, Ferro, Portugal-Perez and Wilson

(2011) found significant effects of aid to services on the export performance of downstream manufacturing industries using an original identification strategy that combined cross-country data on aid flows to services with data on the service intensity of downstream manufacturing industries from input-output tables. All in all, it is fair to say that, as the literature stands, the effect of AfT on the export performance of beneficiary countries has been established only weakly on the basis of aggregate numbers, Ferro et al. (2011) being possibly the strongest exception.

There is even less specifically on technical-assistance programs.[1] A brochure published by the EU Commission (EC 2006) tells the interesting story of a Kenyan fruit and vegetable exporter who got assistance from the EU's Pesticides Initiative program (PIP); the case study presents the program as

> "[...] to provide support to companies like Myner [the Kenyan exporter], to help them get up to speed with European food safety and traceability requirements. [...] Since it began in July 2001, the PIP has had a positive effect on more than 26'000 ACP producers, many of whom are small-scale farmers. Nearly 6.25 million Euros has been committed to the program, with each applicant allocated around 86'000 Euros. In line with the principles of the ACP-EU partnership agreement signed in Cotonou in 2000, the PIP aims to contribute to the development of the ACP's private sector and to promote regional integration."

As for the case study itself, the brochure explains that

> "[w]hen Myner Exports began working with the Pesticides Initiative Program, or PIP in 2002, it was exporting about 300 tonnes of French beans, snow peas, passion fruit, and sugar snaps a year to the European Union. Today, the company exports some 900 tonnes a year. "

Quoting this particular case study in their assessment of the EU's TRA, te Velde et al. (2006) noted that "in an ideal world, one would compare this supported company with a similar one that was not supported". This is precisely what we set up to do in the present paper, although using Senegal instead of Kenya as our sample of study. This paper investigates whether the PIP has significantly affected the export performance of Senegal's horticulture industry. We chose Senegal because the data we had access to provided a unique combination that made it possible to construct a treatment group of firms that got assistance and a control group of firms that did not, and this for a sample period that ran from before the program to its end.

[1] See Marcano and Ruprah (2009).

Using this rich data set, we used a wide array of approaches to estimate the effect of the PIP on firm-level exports of treated products (FFV). We ran DID regressions of the value of exports, by product $\times$ firm $\times$ destination, on control variables as well as a dummy variable marking 'treatment' by the PIP. In order to deal with selection issues, we combined the DID approach with propensity-score matching. We also used a control-function approach similar to Heckman's selection model (Heckman 1979).

In most specifications we tried, we failed to find a significant impact. Only when considering the firm aggregate exports of FFV to the EU, did we find a positive and significant effect of the program. Given the small size and peculiarities of the sample, our results should be treated with caution and we would stop short of concluding that the PIP was useless on the basis of this single impact-evaluation exercise. Additionally, beneficiary firms self-selecting in the program would introduce a bias in the estimated treatment effect. However the direction of the bias is not clear. If self-selected firms are larger and potentially more efficient than others the estimated treatment effect should be biased upwards. Inversely less productive firms may be more likely to rely on financial aid or rent-seeking, biasing the effect downwards. If so, clearly more research is needed on this issue, possibly on other, larger samples to assess whether the PIP had any impact or not.

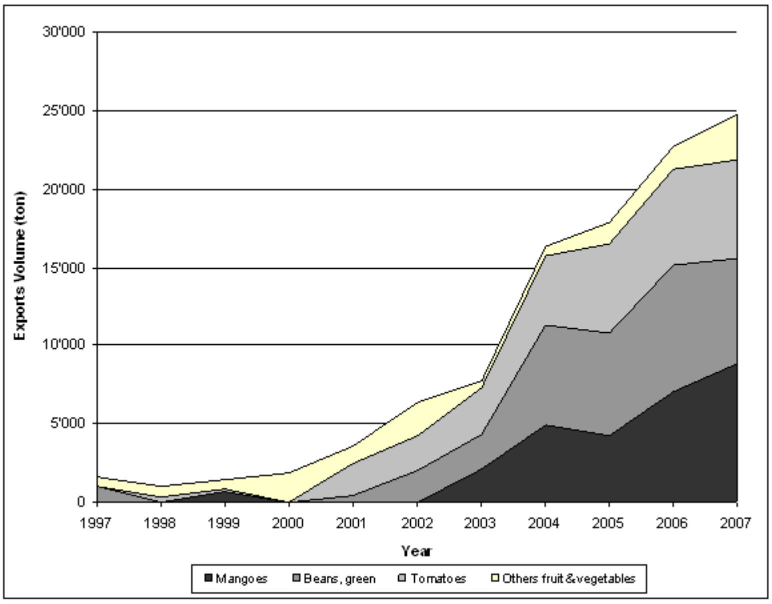
The paper is organized as follows. Section 2 presents some stylized facts on the exporting environment of Senegalese FFV producers and on how the PIP addresses its objective of alleviating some of the constraints that these producers face. Section 3 presents the impact evaluation: data, estimation issues. Section 4 presents the baseline results and some robustness tests. Finally section 5 concludes.

2 Stylized Facts

2.1 Senegal's Exports of Fresh Fruits & Vegetables

According to Senegal's National Horticultural Direction, national exports of fresh fruits and vegetables (FFV) have been rising at a rate of about 15% a year since 2001. French beans account for most of this increase, with volumes up from 652 to almost 9'000 tons between 2001 and 2007 (Figure 1). In 2007, exports volumes of FFV were dominated by French beans (42%), followed by cherry tomatoes (23%), mangoes (16%) and other minor crops including melons, peppers and hibiscus.

Figure 1: Horticulture Exports (in tons) from Senegal, 1997-2007.



Source: FAOstat

Apart from small volumes shipped to neighboring countries, the European Union (EU) remains the main destination market. In 2007, France accounted for 40% of Senegal's FFV export volumes, followed by the Netherlands (35%) and Belgium (16%). In 2008, Senegal ranked fourth among African suppliers of French beans to the EU, after Morocco, Egypt and Kenya (Maertens and Swinnen 2009).

In order to penetrate the EU fruit and vegetable market, Senegal's exporters must comply with strict regulations. Among those, new traceability requirements and recent changes in the EU's pesticides regulation have been of particular concern to the horticultural industry. The EU's pesticide regulation has become notably more stringent since the 1990s. About 350 active substances out of the 823 initially approved for use in the EU have been gradually withdrawn, and Maximum Residue Levels (MRL) and Import Tolerances (IT) are now imposed at levels specific to particular protection chemical-crop combinations. When an IT has not been established, a default value of 0.01 mg/kg corresponding to level of detection in inspection labs are used. As for traceability, the 2002 General Food Law imposes a 'one-step-forward, one-step-back' principle within the EU (with no obligation to keep records in third countries). However in practice EU buyers tend to go beyond the strict legal requirement. Complete traceability all the way up to the overseas producers is part of many private standards like the GlobalGAP.

Interestingly, EU's rising standards have also profoundly altered the structure of the supply chain in Senegal's horticulture sector (Maertens and Swinnen, 2009; Maertens et al, 2011). The cost of compliance with stricter standards have induced consolidation at the intermediation stage, with only larger firms able to cope. Intermediaries exert tighter controls over farming methods, and the structure of upstream farm production has changed, with a sharp decrease in the incidence of contract farming and a rise in that of large-scale estate production.

2.2 The PIP in Senegal

The legislative changes in EU standards and their potential detrimental effects on small growers and exporters were one of the primary drivers behind the launch of the PIP. The program is financed by the European Development Fund and implemented by the COLEACP.[2] It has two main objectives; first to enable ACP exporters of FFV to comply with European traceability and food-safety requirements (in particular as regards pesticide residues), and second to consolidate the position of small-scale producers in the ACP horticultural value chain. The PIP's support activities are organised around five components: (i) good company practises, (ii) training, (iii) capacity building, (iv) regulation & standards, and (v) information & communication. The core of the support (almost 30% of program budget) goes to component (i), which consists of helping producers and exporters to set up internal food-safety management systems in production and marketing operations. The regulation & standards component ensures that all substances recommended in crop protocols are authorized in both the EU and origin country. Additionally, when needed, the program introduces registration of active substances as well as import-tolerance applications.

Senegal ranks fourth, after Kenya, Ghana and Uganda, for the number of PIP protocols signed. In Senegal, PIP beneficiaries produce and export essentially green beans, cherry tomatoes and mangoes to the European market. Their needs relate to in-company training on hygiene and food-safety procedures, development of traceability systems and safe use of pesticides. Four companies also requested support to obtain GlobalGAP certification; for those, pre-audits were conducted. SEPAM obtained its certification in 2004 whereas Soleil Vert, Baniang and AgriConcept are in the process of certification. Except for SEPAM, SAFINA and GDS, Senegal's FFV exporters have outgrower contracts with smallholders rather than own production sites. In 2006, in cooperation with Senegal's ANCAR (*Agence Nationale de Conseil Agricole et Rural*) the PIP launched 'Golden Bean', an awareness campaign directly targeting 1'000 small FFV producers.

[2]Europe-Africa-Caribbean-Pacific Liaison Committee.

2.3 Selection into the Program

Each applicant submits a request for intervention. To qualify the requested intervention must help achieve product compliance with EU traceability and food safety regulations. The request identifies by self-assessment the problems, possible fixes and anticipated results. It also assesses a time line and budget. Applications are considered on a first-come first-served basis. No prioritization or selection criteria apply. Upon acceptance, a protocol detailling actions to be implemented by each party on a cost-sharing basis (50% except for smallholders who are expected to contribute only 20%) is signed between the beneficiary and the PIP. Wages and investment costs are *de facto* excluded. Actions are chosen under each of the five components and each protocol is specific to a beneficiary. Out of the 320 export companies covered under PIP phase one, 219 benefited from the good company practices component, 153 benefited from training under the capacity-building component.

2.4 The PIP's Evaluation

An evaluation of the PIP's phase one was undertaken in June 2008. The PIP detailed objectives, expected results, performance indicators, and outcomes are summarized in the appendix's evaluation matrix. The evaluation relies on trade data reported by the firms themselves and from Eurostat and data from a survey conducted among PIP beneficiaries and EU importers. The impact of the program is evaluated at the aggregate level. Outcomes for beneficiaries are reported without controlling for location, size or experience, with the exception of objective (S1), for which outcomes are reported by ACP country or type of intervention.

Overall, the evaluation report drew up a very positive image of the program, contributing to the launch of a second five-year phase in 2009. While fairly comprehensive, the PIP's evaluation suffers from a typical drawback of this type of exercise—namely, the lack of a counterfactual to benchmark the performance of treated firms and products. Precisely, we set up to address this issue and estimate the impact of the PIP on Senegal's FFV export flows by taking similar, untreated export flows as the counterfactual.

3 Impact Evaluation

3.1 The Data

Our dataset is constructed using three databases which together form a rich and unique combination. First, we have export data at the transaction level (aggregated annually) over 2000–2008. Each record includes the firm's tax ID, the product code, the country of destination, and the export value (in US dollars) and quantity (in tons) for over 500 HS8 products to 90 countries. Second, the PIP's administration in Brussels provided us with a list of the Senegalese firms that got assistance from the program in each year of the sample period. Finally, we obtained data on employment and sales from the C*entre National d'Identification* (CNI). As the CNI also identifies firms by their tax ID, we could merge the three datasets. Among the reporting firms, almost 3% appear only once in the dataset. That is, they export only one product to one destination one year. As these observations are likely to be mis-reports or individuals, we drop them from our sample. We also drop international organizations and embassies, as well as trading and transport companies (about a quarter of all observations). This leaves us with a sample of almost 2'000 observations.

Let i be an exporting firm, k a product, j a destination country, and t a year. As the PIP targets *products* (FFV), some of a firm's products may be covered and some others not. In addition, technical assistance provided under the PIP helps make FFV marketable on EU markets, but does not necessarily help on other markets with different (or no) food-safety requirements. In view of this, we take the (i, k, j, t) vector as our unit of observation. Our dependent variable is the annual flow of exports of product k to destination j by firm i in year t, y_{ikjt}. That is, we take the intensive margin as our baseline measure of performance.

Table 1 gives descriptive statistics of firm-level covariates and performance indicators for treated and non-treated flows. Covariates include firm i's annual (overall) turnover, $sales_{it}$, and employment, $Nemployee_{it}$. Performance indicators at the intensive margin include firm i's exports of product k to destination j in year t, $export_{ijkt}$; its total exports of FFV to country j in year t, $total\_export_{ijt}^{FFV}$; its total exports of product k worldwide, $total\_export_{ikt}$; its total exports of FFV worldwide, $total\_export_{it}^{FFV}$; and its total exports worldwide, $total\_export_{it}$. Export performance indicators at the extensive margin include firm i's number of destinations serviced with product k in year t, $Ndest_{ikt}$; its number of destinations serviced with FFV; $Ndest_{it}^{FFV}$; its total number of destinations, $Ndest_{it}$; its number of FFV products to destination j, $Nprod_{ijt}^{FFV}$; its number of products to destination j, $Nprod_{ijt}$; its number of FFV products worldwide, $Nprod_{it}^{FFV}$; and its total number of products worldwide, $Nprod_{it}$.

Two observations are in point. First, participating firms are larger than non-participating ones. Thus, size must be controlled for in order for non-participating firms to provide a credible counterfactual. Second, these are all small firms as regards to the product and destination margins. The average numbers of products and destinations are small.

Table 1: Descriptive Statistics, non-Treated and Treated Flows.

	Non-Treated Flows		Treated Flows		Mean(Non-Treated) =Mean(Treated)	
	Mean	Std. dev.	Mean	Std. dev.	t-Stat	Signif.
Sample	1370		492			
Firm characteristics						
Sales_{it}	21.7	1.7	21.1	1.9	4.9	***
Nemployee_{it}	4.4	1.1	5.3	2.1	-6.9	***
Intensive margin						
Export_{ijkt}	8.0	2.4	10.1	2.2	-17.6	***
Total $\text{export}_{ijt}^{FFV}$	8.2	2.9	11.9	2.0	-22.8	***
Total export_{ikt}	8.5	2.6	11.1	2.3	-21.0	***
Total export_{it}^{FFV}	8.6	2.8	13.2	1.8	-35.0	***
Total export_{it}	12.4	1.8	13.1	1.9	-7.7	***
Extensive margin						
Ndest_{ikt}	1.5	0.9	2.2	1.2	-11.6	***
Ndest_{it}^{FFV}	1.3	1.7	8.0	7.6	-19.5	***
Ndest_{it}	5.1	4.3	3.4	1.6	12.3	***
Ndest_{ijt}^{FFV}	0.6	1.1	2.7	2.2	-20.1	***
Ndest_{ijt}	15.0	21.3	3.2	2.5	20.0	***
Ndest_{it}^{FFV}	1.3	1.7	8.0	7.6	-19.5	***
Ndest_{it}	38.9	45.2	9.4	8.6	23.1	***

Variables are in logs. Sales are reported in CFA francs and exports in U.S. dollars. *, **, and *** denote statistical significance of the t-statistics at the 10%, 5%, and 1% levels, respectively.

3.2 Estimation Issues

Estimating the effect of the PIP poses a standard missing-data problem –estimating how much smaller would have been the export flows that got assistance, had they not gotten assistance. Formally, let

$$d_{ikjt} = \begin{cases} 1 & \text{if } (i,k,j,t) \text{ is treated at } t \\ 0 & \text{otherwise} \end{cases}$$

and

$$d_{ikj} = \begin{cases} 1 & \text{if } \exists : t \text{ such that } d_{ikjt} = 1 \\ 0 & \text{otherwise.} \end{cases}$$

That is, d_{ikj} marks the treatment group. The basic estimator for the problem at hand is the difference-in-differences (DID):

$$y_{ikjt} = \mathbf{x}_{ikjt}\beta + \gamma d_{ikjt} + \delta_{ikj} + \delta_t + u_{ikjt} \tag{1}$$

where $\mathbf{x}$ is a vector of observable export flow characteristics, δ_{ikj} and δ_t are respectively firm × destination × product and time fixed effects, and u_{ikjt} is an error term. Fixed effects δ_{ikj} control for time-invariant firm characteristics potentially affecting both performance and selection into the program, like managerial ability (see Angrist and Krueger 1999, Smith 2000, or Jaffee 2002).

Next, in order to deal with selection issues, we combine the DID estimator (1) with matching, following Heckman, Ichimura and Todd (1997). From now on, for simplicity of exposition, let us denote by a 'flow' a (firm × product × destination) triplet. Using results by Rosenbaum and Rubin (1983), matching is done on the basis of the estimated propensity score (PS), using a probit of the participation status on a vector $\mathbf{z}$ of observable firm characteristics. Letting v_{ikjt} be an error term orthogonal to u_{ikjt}, the first-stage selection equation can be written as

$$\Pr(d_{ikj} = 1) = f(\mathbf{z}_{ikj0}\alpha + v_{ikj}). \tag{2}$$

In (2), the vector $\mathbf{z}$ is identical to $\mathbf{x}$ as no outside determinant of participation is available. It must be evaluated at the time the participation decision is made – typically its initial value. The estimated propensity score is then retrieved from (2), and the control group is constructed by selecting untreated firms whose propensity scores are "close enough" to those of treated ones. Under nearest-neighbour matching, each treated flow is matched with the untreated flow having the closest PS within a fixed range.[3] Practically, DID-with-matching estimation is done in two steps. In the first, the participation equation is estimated, yielding an estimated PS and a common support; in the second, the DID equation is estimated on the common support. The common support is formed by disregarding unmatched individuals as well as those with estimated PS of zero or one.

We also control for selection using a control-function approach that closely resembles the Heckit procedure (Heckman 1979). The approach proceeds in two steps. The first-step regression is as before, i.e. (2). In the second step, inverse Mills ratios retrieved from the first step are added to (1) as additional regressors.

Note that, besides selection bias, other issues may complicate the estimation of γ in (1). One is serial correlation. Persistence in the process driving the error term may be aggravated by the extreme form of serial correlation in the treatment variable. Bertrand, Duflo and Mullainathan (2004) show that ignoring this

[3]On this, see Smith and Todd (2005). For robustness purposes we also used the kernel and radius matching. Results remained unchanged.

source of serial correlation can lead to an inflated probability of type-I errors (i.e. being over-optimistic in the evaluation of the treatment's impact). They suggest a correction in two-steps. In Step 1, individual performance for both treated and untreated individuals is regressed on all observables except the treatment. In Step 2, residuals from Step 1 for the treated individuals only are retrieved and averaged for (i) the pre-treatment period, (ii) the treatment period. Those average residuals form a two-period panel are then regressed on an indicator variable τ that takes value 1 for the treatment period. The estimated ATT is the coefficient on τ in the second step.

4 Results

4.1 Baseline Results

Balancing properties are addressed by testing for equality of means between treated and matched controls for nearest-neighbour matching (Rosenbaum and Rubin 1983). Table 2 reports results for each covariate included in the probit model determining selection into treatment. Covariates include the natural logarithm of firm i's initial turnover, the natural logarithm of its initial number of employees, the initial number of products exported by firm i to destination j, the initial number of countries served with product k, the initial natural logarithm of total export value of FFV products from firm i to destination j. Results show that, for many covariates, there is a strong bias before matching but matching eliminates it. The null hypothesis of balanced sub-samples is not rejected except for turnover.[4]

[4]Balancing-test failure on sales may bias estimated treatment effects if sales correlate with the treatment effect. We checked if that was the case by interacting the treatment indicator variable with firm size. Results suggest no differential treatment effects by size and are available upon request. Therefore, even though sales levels differ between the treatment and control groups, this is unlikely to drive our results.We are grateful to an anonymous referee for attracting our attention to this issue.

Table 2: Balancing Properties of Covariates in Treated and Control Groups

	Sample	Mean treated "flows"	Mean control "flows"	% bias between treated and controls	% reduction bias	t-test Mean(treated) =Mean(control) t	$p > t$
$ln(sales_{it_0})$	Unmatched	21.39	21.25	11.5		1.3	0.19
	Matched	21.39	21.56	-13.6	-18.6	-1.72	0.08
$ln(Nemployees_{it_0})$	Unmatched	5.02	4.18	55.2		7.05	0.00
	Matched	5.02	4.94	4.7	91.6	0.44	0.66
$Nprod_{ij_0}$	Unmatched	2.38	32.03	-144.2		-14.77	0.00
	Matched	2.38	2.37	0.1	100	0.16	0.87
$Ndest_{ikt}$	Unmatched	2.41	1.37	114.5		15.1	0.00
	Matched	2.41	2.3	9.7	91.6	0.84	0.40
$ln(total\_export^{FFV}_{ijt_0})$	Unmatched	11.93	7.51	154		16.63	0.00
	Matched	11.93	12.13	-7	95.4	-1.21	0.23

Matching is by nearest neighbour. *, **, and *** denote statistical significance of the t statistics at the 10%, 5%, and 1% levels, respectively.

Table 3 reports difference-in-differences (DID) estimates on the treated, for our baseline specification. That is, the average effect of the PIP on assisted firms' export performance. The dependent variable is the natural logarithm of the export value of product k from firm i to country j in time t. Our main variable of interest is the treatment indicator variable $treatment_{ikjt}$ taking value one if firm i exporting product k to destination country j benefited from the PIP program in year t. All regressions are run at the i, k, j, t level and control for (firm $\times$ product $\times$ country) and time fixed effects. Standard errors are clustered at the firm level. Coefficients in all specifications are not significant suggesting no effect of the program. Columns 1-3 report DID estimates without matching. Control variables include the natural logarithm of annual turnover of firm i in year t, $ln(sales_{it})$, the natural logarithm of the number of employees for firm i in year t, $ln(Nemployees_{it})$ and firm i's experience in servicing product k to country j, $experience_{ikjt}$. That is a dummy variable taking value 1 if the firm exported at least two years product k to country j before time t. Column 4 reports DID estimates with matching, i.e. restricting the sample to the common support defined using the Nearest Neigbhour(NN)-PSM procedure. Matching is done at the firm, product destination level and results are reported using the NN estimator with caliper (0.04).[5] Column 5 reports the treatment-effect estimate using Heckman's two-step procedure, i.e. estimate from the second-step regression run with the inverse Mills ratio λ retrieved from the first step. Column 6 reports results from the second step of the procedure suggested by Bertrand, Duflo and Mullainathan (BDM) (2004). Residuals from the first step are retrieved and averaged for the pre-treatment period and the treatment period.

Table 3: Baseline Results, Average Effect of the PIP on Assisted Firms.

	(1) DiD	(2) DiD	(3) DiD	(4) DiD with Matching	(5) Two stage Heckman	(6) BDM Correction
treatment_{ikjt}	0.354	0.216	0.112	-0.013	0.238	0.111
	(0.235)	(0.288)	(0.322)	(0.418)	(0.362)	(0.293)
$\ln(\text{sales}_{it})$		1.062***	1.125***	1.246***	1.263***	
		(0.212)	(0.262)	(0.243)	(0.246)	
$\ln(\text{Nemployee}_{it})$			-0.099	0.017	0.009	
			(0.155)	(0.152)	(0.148)	
experience_{ikjt}	0.183*	-0.047	0.018	-0.143	-0.145	
	(0.104)	(0.168)	(0.172)	(0.169)	(0.169)	
λ					-0.153	
					(0.153)	
constant	7.94***	-14.59***	-15.94***	-18.98***	-19.39***	-0.073
	(0.225)	(4.552)	(5.422)	(5.128)	(5.246)	(0.166)
Observations	1,862	1,193	1,071	698	698	176
R-squared	0.132	0.179	0.188	0.207	0.207	0.007
Number of id	1,134	657	577	369	369	155

Robust standard errors clustered at firm level are in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Results from the first-step participation regression, which is common to the matching DID and Heckman estimations, are available from the authors upon request. The regression is run at the i, k, j level. Results suggest that the probability of treatment correlates positively with sales and FFV exports, employment, and

[5] We also used kernel matching with bandwidth (0.04) and radius matching with caliper(0.04). Results remained unchanged.

initial performance at the destination margin although negatively with product diversification. Thus, treatment seems to have targeted relatively large firms concentrated on a few products. Results for the first step of the BDM procedures are also available from the authors upon request. The regression is run at the i, k, j, t level and the dependent variable is the log value of exports. Individual performance for both treated and untreated individuals is regressed on all observables except the treatment. Only initial sales have a positive significant effect on subsequent growth, suggesting diverging export performances across firms.

4.2 Robustness

In this section we present estimation results of the effect of the PIP on assisted firms, for two alternative export performance indicators. The choice of aggregation of export flows across all EU-15 destinations and across FFV products follows from the fact that PIP helped firms achieve FFV products compliance with EU traceability and food-safety regulations.

Table 4 reports estimation results when considering firm i export of product k to the EU-15 in year t, as the export performance indicator. Balancing tests are satisfied and are available upon request. At the firm-product level there is no problem of unbalanced covariates in our model. Regressions are run at the i, k, t level and the coefficients on the treatment variable $treatment_{ikt}$ are not significant for any of the specifications in Table 4.

Finally, Table 5 reports results from regressions run at the level of the firm instead of the firm × product × destination combination. Estimation at the firm level may drastically reduces sample size and means mixing up exports that are covered by the program with exports that are not (namely, products other than FFV). However, it is advisable, since the decision to participate and some of the covariates are at the firm level. The export performance indicator is the export value of FFV to the EU-15 exported by firm i in year t. Balancing tests results are also satisfied. At the firm level there is no problem of unbalanced covariates in our model. The coefficient on the treatment variable $treatment_{it}$ is significant at the 5% level only in column (1).

All in all, results suggest that while there seem to be an effect of the program when considering total FFV exports to the EU, the effect disappears when looking at a more disaggregated level (Table 3 and 4). These results are in line with findings in the program evaluation report— namely, the share of ACP exporters in EU imports value of FFV increased from 6.1% to 6.2% between 2001 and 2006.

Table 4: Robustness I, Average Effect of the PIP on Assisted Firms.

	(1) DiD	(2) DiD	(3) DiD	(4) DiD with Matching	(5) Two stage Heckman	(6) BDM Correction
treatment_{ikt}	0.100	-0.023	-0.115	-0.114	-0.669	-0.410
	(0.215)	(0.264)	(0.305)	(0.327)	(0.813)	(0.564)
$\ln(\text{sales}_{it})$		0.909***	0.923***	0.948***	0.890***	
		(0.217)	(0.287)	(0.293)	(0.290)	
$\ln(\text{Nemployee}_{it})$			0.063	0.074	0.078	
			(0.105)	(0.107)	(0.103)	
experience_{ikt}	0.299	-0.252	-0.228	-0.253	-0.213	
	(0.248)	(0.373)	(0.473)	(0.504)	(0.485)	
λ					0.333	
					(0.357)	
constant	8.95***	-9.36	-10.01	-18.98***	-8.54	0.17
	(1.008)	(5.834)	(5.975)	(5.128)	(6.054)	(0.311)
Observations	681	368	288	281	286	93
R-squared	0.153	0.211	0.227	0.218	0.230	0.036
Number of id	373	189	142	139	140	78

Robust standard errors clustered at firm level are in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

5 Concluding Remarks

By and large, we find no significant impact of the PIP on Senegal's FFV export flows when taking similar, untreated export flows as the counterfactual. There are two ways of interpreting such a no-impact result.

The naive interpretation is that the PIP simply fails to achieve its objective. That may well be true, but our failure to reject the null of no impact is not sufficient to reach that conclusion. First, the choice of Senegal as a testing ground was driven by data availability. It has no claim to be a representative or random sample. Different conclusions may be reached from other samples, and clearly a full, cross-country impact evaluation should be undertaken. Second and more importantly, it is possible that the PIP affected not only the treated export flows, but also the untreated ones, through spillovers. Participating firms are the largest

Table 5: Robustness II, Average Effect of the PIP on Assisted Firms.

	(1) DiD	(2) DiD	(3) DiD	(4) DiD with Matching	(5) Two stage Heckman	(6) BDM Correction
treatment_{it}	0.533**	0.381	0.666	0.666	0.711	0.477
	(0.239)	(0.128)	(0.421)	(0.426)	(2.898)	(0.292)
$\ln(\text{sales})_{it}$		0.529	0.496	0.520	0.518	
		(0.372)	(0.532)	(0.538)	(0.626)	
$\ln(\text{Nemployee}_{it})$			-0.181	-0.188	-0.187	
			(0.262)	(0.266)	(0.265)	
experience_{it}	-0.007	-0.007	-0.218	-0.237	-0.236	
	(0.233)	(0.233)	(0.460)	(0.468)	(0.530)	
λ					-0.026	
					(1.570)	
constant	10.07***	0.855	2.193	1.992	2.017	-0.278
	(0.704)	(7.943)	(11.724)	(11.795)	(12.850)	(0.214)
Observations	199	119	85	79	79	16
R-squared	0.301	0.380	0.354	0.362	0.362	0.308
Number of id	69	38	26	21	21	9

Robust standard errors clustered at firm level are in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

and more efficient of Senegal's FFV sector. On one hand, this means that they are more susceptible than others to benefit from the program. Thus, what we obtain is an estimate of the average effect of the treatment on the treated (ATT), which may over-state the program's potential effect on the whole population of producers. On the other hand, it may also mean that smaller firms, although left out of the program, can benefit from it through imitation of best practices and even unobserved assistance from larger firms. The argument is even more potent for untreated *products*: when a firm gets PIP assistance for its FFV activities, it is quite possible (indeed, likely) that its other activities benefit as well from improved managerial practices; or *destination countries*: export flows to non-EU destinations may benefit from the program for the same reason. In the presence of such unmeasured spillovers, the PIP's impact would be underestimated by impact-evaluation methods. This is important to keep in mind, as public assistance (whether from local governments or donors) should be justified by market failures, like spillovers, rather than a positive rate of return to beneficiaries (which would simply create a market demand for assistance services without justifying use of public funds). Thus, impact evaluation of technical-assistance programs like the PIP is a double-edged sword and must be interpreted with caution.

The second conclusion that should be *avoided* is that, either because the data are not sufficiently reliable or comprehensive or because of the caveats just discussed, rigorous impact evaluation should not be undertaken. The lack of rigorous impact evaluation undermines the credibility of claims about the program's benefits made on the basis of case studies, because it is impossible to know whether they are representative or not. Indeed, this is the message conveyed by the 2006 assessment of the impact of EU TRA (te Velde et al. 2006). Worse, in a context where taxpayers are asking for accountability and results in development aid, Paul Milgrom's 'unraveling principle' applies. Rational taxpayers are likely to take all the news that is not told to be detrimental. In other words, the bad news that impact evaluation can possibly generate (as in the present case) is probably fully anticipated. The more the development community can provide rigorous evidence that at least some programs do make a difference; or that some components do; or that, when not, failure is part of useful experimentation and action is being taken to remedy the observed ineffectiveness, the more support there will be for development aid.

However, as the present study highlights, it is difficult to 'improvise' impact evaluation *ex-post* when a program was not designed to be evaluated. Far better would be to think seriously about evaluation *ex-ante*, so that TRA programs generate experimental settings out of which useful lessons could be drawn.

References

[1] Angrist, Joshua D., and Alan B. Krueger. "Empirical Strategies in Labor Economics". In Orley Ashenfelter and David Edward Card, eds. *Handbook of Labor Economics*, Vol. 3A, Chapter 23. Amsterdam: Elsevier, 1999.

[2] Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-in-Differences estimates?" *Quarterly Journal of Economics* 119 (2004): 249-275.

[3] Blundell, Robert, and Monica Costa Dias. "Alternative Approaches to Evaluation in Empirical Macroeconomics". CEMMAP Working Paper CWP10/02, Institute for Fiscal Studies, 2002.

[4] Brenton, Paul, and Jan Erik von Uexkull. "Product specific technical assistance for exports - Has it been effective?" *Journal of International Trade & Economic Development* 18 (2009): 235-254.

[5] Cali, Massimiliano, and Dirk W. te Velde. (2009), "Does Aid for Trade Really Improve Trade Performance?" *World Development* 39 (2009): 725-740.

[6] Economic Community, *Making trade work for development. Trade-related assistance: An update*, Brussels: EC Commission, DG Trade. 2006.

[7] Ferro, Esteban, Alberto Portugal-Perez and John S. Wilson. "Aid to the Services Sector: Does It Affect Manufacturing Exports?" World Bank Policy Research Working Paper 5728. Washington, D.C.: The World Bank, 2011.

[8] Gamberoni, Elisa, and Richard Newfarmer "Aid for Trade: Matching Potential Demand and Supply". World Bank Policy Research Working Paper 4991. Washington, D.C.: The World Bank, 2009.

[9] Heckman, James J. "Sample Selection Bias as a Specification Error". *Econometrica* 47 (1979): 153-161.

[10] —, H. Ichimura, and Petra Todd ."Matching as an econometric evaluation estimator: Evidence from evaluating a job training program". *Review of Economic Studies* 64 (1997): 605-654.

[11] —, —, J. A. Smith and Petra Todd. "Characterising selection bias using experimental data". *Econometrica* 66 (1998): 1017-1098.

[12] Jaffe, Adam B. "Building Program Evaluation into the Design of Public Research Support Programs". *Oxford Review of Economic Policy* 18 (2002): 22-34.

[13] Marcano, Luis, and Inder J. Ruprah. *Does Technical Assistance Matter? An Impact Evaluation Approach to Estimate its Value Added.* Office of Evaluation and Oversight. Washington D.C.: Inter-American Development Bank. 2009.

[14] Maertens, Miet, and Johan F.M. Swinnen. "Trade, Standards, and Poverty: Evidence from Senegal". *Journal of World Development* 37 (2009): 161-178.

[15] —, Liesbeth Colen, and Johan F. M. Swinnen."Globalization and Poverty in Senegal: A Worst Case Scenario?" *European Review of Agricultural Economics* 38 (2011): 31-54.

[16] Nelson, Douglas, and Simone J. Silva."Does Aid Cause Trade? Evidence from an Asymmetric Gravity Model". University of Nottingham Research Paper 2008/21 (2008).

[17] OECD. *How to Evaluate Aid for Trade: Approaches, Methodologies, and Processes.* OECD, COM/DCD/TAD (2010) 2. Paris: OECD. 2010.

[18] Osei, Robert, Oliver Morrissey and Tim Lloyd. "The Nature of Aid and Trade Relationships". *The European Journal of Development Research* 16 (2004): 354-374.

[19] Rosenbaum, Peter, and Donald B. Rubin (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects". *Biometrika* 70 (1983):41-55.

[20] Smith, Jeffrey. "A Critical Survey of Empirical Methods for Evaluating Active Labor Market Policies". *Swiss Journal of Economics and Statistics* 136 (2000): 247-268.

[21] —, and Petra Todd. "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics* 125 (2005), 305-353.

[22] Te Velde, Dirk Willem, and Adrian Hewitt, Massimiliano Cal and Sheila Page. *A Critical Assessment of the EU's Trade-related Assistance to Third Countries. Lessons from the Past, Policy Options for the Future.* London: Overseas Development Institute, Study for the International Trade Committee of the European Parliament. 2006.

[23] Wagner, Don. "Aid and Trade—An Empirical Study". *Journal of the Japanese and International Economies* 17 (2003): 153-173.

[24] WTO/OECD. *Aid for Trade at a Glance: Showing results*; Geneva/Paris: WTO and OECD. 2011.

Appendix

Figure 2: The PIP's Evaluation Matrix

Table 1: PIP Evaluation Matrix

Program component	Objective or Expected results	Performance indicator	Outcome	Conclusion of the report
Global	ACP countries maintain their share (in value terms) in EU imports of fresh fruits and vegetables.	**Indicator G1**: The share in EU Extra-EU imports of FFV remains stable at 7% in value.	**O1**: Share of ACP exporters in EU imports value of FFV increased from 6.1% (371'116 thousands euros) in 2001 to 6.2% (534'681 thousands euros) in 2006.: Share with regards to volumes decreased from 7% (335'647 tons) to 6% (384'437 tons).	Achieved
Specifique	ACP FFV suppliers meet european Maximum Residues Limit (MRL) and traceability regulations requirements.	**Indicator S1**: Exporters with established internal systems of food safety management in the production and marketing process account for 80% of exports of FFV to the EU by the end of the program.	**O1**: The 219 firms that signed a protocle with PIP account for 74% of ACP FFV exports to the EU in 2006 (276'000 tons*). **O2**: The 145 firms that set up a traceability system account for 60% of ACP FFV exports to the EU in 2006 (229'084 tons*). **O3**: The 191 firms that benefited from staff training in food safety procedures account for 70% of ACP FFV exports to the EU in 2006 (265'395 tons*). Out of these, 56 obtained certification for their food safety management system, and account for 37% of exports in 2006 (142'233	Achieved
		Indicator S2: Thoses exporters suffer less shipments rejections due to food safety issues on the EU market.	**NA**. Data on rejections and removals (once the product entered in the EU market) was not available.	NA
		Indicator S3: Improvement in the degree of satisfaction of european importers regarding the level of conformity of ACP exports with public and private standards requirements.	**O1**: 47% of importers (34 respondents) believe the impact of PIP was determinant in bringing ACP exports in conformity with regulations requirements. **O2**: 15% believe its was not sufficient. **O3**: With regards to private standards requirements 35% believe PIP intervention was determinant. While 21% estimate it as not sufficient. **O4**: 55% fear that the end of the program will negatively impact the quality of products.	Achieved
Information and communication	Producers and exporters in ACP countries are informed in time of destination markets requirements with a focus on MRL and Import Tolerances (IT). Therfore, they can adapt their production and production pratices to changes in importers requirements. Importers are informed of exporters compliance efforts.	**Indicator R1**: Is the information system** effective and efficient in providing the needed information to both importers and exporters.	**O1**: Among the 176 beneficiaries that answered, 80% are satisfied with the information/communication toolkit developped by the program. **O2**: Among inporters (34 respondents), 59% know about the PIP, and 50% are aware that their suppliers benefited from the PIP.	Achieved
Regulation and standards	Crop protocols (technical itineraries) are drawn up by the PIP for the main crops exported to the EU market. Applications for Tolerances Import are submited for combinations of product and active substance for which the pesticide use led to residue levels above the	**Indicator R2**: Products targeted by the component account for 90% of export flows by the end of the project	**O1**: Products for which technical itineraries were elaborated account for 91.4% of ACP FFV exports to the EU in 2006 (384'437 thousands euros).	Achieved
Good company practices	Producers adopt Good Agricultural and Production Practices (GAP and GPP). Exporters set up contol and traceability systems to insure food safety.	**Indicator R4**: **Indicator S1**	See above	Achieved
	Export flows from small scale producers are maintained (in value terms) as must as possible.	**Indicator R5**: The decrease in small scale producers export values of FFV to the EU is bounded to 20%.	NA	NA
Capacity building	ACP actors of the FFV industry participate in at least ten national task forces. These task forces act as coordination platforms between the producers/exporters, the private professional organisations, the relevant public services and the authorities in charge of pesticide controls.	**Indicator R6**: At least ten task forces are created and are financially viable by the end of the program.	NA	NA
	ACP exporters and professional organisations in the FFV industry develop an european network of technical structures which serve as	**Indicator R7**: An european network is established	NA	NA

Notes
EU is the EU-15 EU imports correspond to extra -EU imports.
FFV = Fresh fruits and vegetables excluding bananas and citrus fruits that are not covered by PIP
* Based on firms' own declarations.
** Include website, technical documents, newsletter INFOPIP, technical itineraries.
Source: Evaluation finale du Programme Initiative Pesticides, Rapport Final, 2008.