Do firms benefit from quality-related training activities?

Antonina Popova[1] Valentina Rollo[2] Jasmeer Virdee[3]

ITC ITC ITC

Abstract

This paper assesses the impact of quality-related business training activities on firm certification and export status, using panel data on 14 Latin American countries for 2006 and 2010. Using a rich set of firm level controls, we apply a difference-in-difference regression specification and propensity score matching to check the robustness of the results. Findings indicate that quality-related trainings help firms gain and retain internationally recognised quality certificates (IRQC). Furthermore, our results show that these trainings help firms transition from non-exporter to exporter status as well as to retain their exporter status in cases where they already export. Training also resulted in annual sales and employment growth. Interestingly, the magnitude of these benefits increases with the size of the firm. This may be related to the lower absorptive capacities of smaller firms, or simply that larger firms selected higher quality training activities.

Key words: Export Promotion, Quality Certification, Impact Evaluation, Training, Difference-in-difference

JEL codes: F13; F14; L25 ; C31

[1] Antonina Popova, Trade and Development Analyst, Office of the Chief Economist, International Trade Centre, Palais des Nations, 1211 Geneva 10, Switzerland; and; Department for European Studies, Primakov National Research Institute of World Economy and International Relations, Russian Academy of Science, Moscow, Russia; e-mail: popova.tonja@gmail.com

[2] Valentina Rollo, Economist, Office of the Chief Economist, International Trade Centre, Palais des Nations, 1211 Geneva 10, Switzerland; e-mail: rollo@intracen.org.

[3] Corresponding author: Jasmeer Virdee, Economist, Office of the Chief Economist, International Trade Centre, Palais des Nations, 1211 Geneva 10, Switzerland; e-mail: jasmeervirdee@gmail.com

I. Introduction

Standards are a set of agreed-upon specifications related to a product's features or a firm's production processes. They are essential to international trade and value chains, because they determine not only whether inputs are compatible with the next stage of production, but also if final products are safe for consumption.

Nevertheless, navigating the maze of standards is no easy job, especially for small and medium-sized enterprises (SMEs). Enterprises must first learn which standards are relevant to their business, select one or more which add value to its offering, implement the necessary changes to become compliant, and then demonstrate compliance (ITC, 2016). To be compliant with any given standard, enterprises must acquire that standard's certificate, which can be costly. Furthermore, these costs usually come in the form of fixed costs. This places SMEs at a disadvantage, as certification costs make up a larger share of their unit costs compared to larger firms.

This issue is particularly relevant in developing countries, where firms tend to be smaller, less productive, and have lower capacities to comply with standards and regulations than in developed countries. In addition, firms in developing countries may also face a more challenging business ecosystem, due to poor governance or lack of testing facilities and logistics infrastructure (ITC, 2016).

As a result, an industry has developed around helping firms acquire quality certificates for their products or their procedures. According to a survey of Spanish firms, 30% reported spending less than €6,000 annually on quality management, while 15% spent over €60,000 (Viadiu & Fransi, 2005). Attending total quality management (TQM) trainings, paying consultants, and paying for the procedural costs of certification all add up. The global ISO certification market was valued at $11.8 billion in 2017 (PMR, 2017).

However, are quality management trainings effective in helping firms acquire certification and eventually exporting? If they are, which types of firms are most likely to benefit? This paper tries to provide an answer to these questions using World Bank Enterprise Surveys panel data on 14 Latin American countries for 2006 and 2010. We assess the impact of attending quality-related trainings on firms' likelihood to acquire or maintain an internationally recognised quality certificate (IRQC) as well as to become or remain exporters.

Using a rich set of firm level controls, we apply a difference-in-difference regression specification and propensity score matching to check the robustness of the results. Our identification strategy takes advantage of the fact that not all firms in the sample received these trainings, so that we have a treatment and control group. Also, we have a wide range of firm-level information before and after the treatment, in both 2006 and 2010. Finally, we account for various sources of unobservable heterogeneity by including sector and country fixed effects.

The results from this difference-in-differences approach indicate that quality-related trainings help firms gain and retain IRQCs. These trainings also help firms transition from non-exporter to exporter status as well as to retain their exporter status. Interestingly, the magnitude of the benefit increases with the size of the firm. The lower absorptive capacities of smaller firms may explain why treatment on smaller firms is less effective. On the other hand, larger firms may simply be selecting higher quality training programmes.

Our paper is related to several strands of the economic literature. First, our focus on the impact of quality-related business training activities on firm certification and export status builds on the assumption that quality certifications have positive effects on firms' performance and export status. As such our paper builds on the literature that shows how certification raises not only firms' growth (Terlaak & King, 2006), but also their likelihood of exporting and export values (Martincus, Carballo & Gallo, 2010; Otsuki, 2011; Goedhuys & Sleuwaegen, 2016).

Second, to become certified, firms need to make considerable investments in personnel and equipment, both in terms of time and resources. Governments and national trade associations have a role to play in supporting firms' efforts to become certified, and provide specific training activities (Kotler & Gertner, 2002). As such, our paper builds on the literature that assess the impact of governmental export promotion initiatives on firms' export competitiveness (Ahmed, Julian & Majar, 2006; Gençtürk & Kotabe, 2001; Lederman, Olarreaga & Payton, 2009; Martincus, Carballo & Gallo, 2010). Some evidence shows the positive impact of export promotion programmes on firms export values (Görg, Henry & Strobl, 2008; Martincus & Carballo, 2010; Freixanet, 2012; Geldres-Weiss & Carrasco-Roa, 2016), as well as on firm's ability to enter and survive in exports markets (Lederman, Olarreaga, and Zavala 2016), and export-related resources and capabilities of small exporters (Leonidou, 2011).[4]

Finally, by focusing on trainings, our paper also contributes to the broad literature that investigates on the impact of different types of training activities on firms' performance. Several papers confirm that training significantly boosts productivity, at the firm level (Zwick, 2006), or at the sector level (Dearden, L., Reed, H. & Reenen, J. V., 2000). This effect is reflected in the reduction of the cost of production (Dearden, L., Reed, H. & Reenen, J. V., 2005; Moretti, 2004; Zwick, 2006; Bilanakos *et al.*, 2014).

To our knowledge there is no empirical evidence showing that quality-related training activities or programmes increase firms' probability to gain an IRQC or to start exporting. This is partly due to the lack of enterprise surveys capturing the relevant information. As such, our paper fills a gap in this strand of the literature by confirming that this form of support to firms is effective in the sample of Latin American firms analysed.

Our findings suggest that, in general, investments in quality control and quality-related trainings in Latin American led to a variety of beneficial outcomes for these firms, including gaining and retaining both quality certifications and exporter status, as well as higher sales and employment. However, larger firms saw better outcomes than smaller firms, which may point to the limited absorptive capacities of smaller firms.

The remainder of the paper is organised as follows. In Section II we present the dataset and describe some of its main features. Section III presents the statistical framework for the difference-in-difference estimation, while Section IV reports the main findings. Section V presents robustness

[4] However, these positive results are counterbalanced by further evidence suggesting that export promotion programmes have no impact on the intensive margins (Lederman, Olarreaga & Zavala, 2016), or the probability to export (Bernard & Jensen, J. B., 2004; Görg, Henry & Strobl, 2008)

checks, based on the propensity score matching method. Section VI discusses the policy implications of the results and Section VII concludes.

II. Data and descriptive statistics

This paper uses firm-level data collected by the World Bank's Enterprise Surveys to determine whether firms which received quality-related trainings were more likely to (a) subsequently acquire an IRQC or (b) become exporters. The question asked in the 2010 World Bank Enterprise Surveys, which determines whether a firm was treated or not is, '*Over the last three years, did this establishment use any services or programs to improve quality control or training to obtain quality certification?*' This means that the treatments were implemented sometime between 2007 and 2009. The discerning reader will have noticed that the question stated above is actually two questions in one. This muddies the interpretation of our results somewhat, but depending on the construction of different specifications one half of the question may be more relevant than the other.

The dataset contains no information on why some firms received (or indeed paid for) 'services or programmes' (hereafter simply referred to as 'treatments' or 'quality trainings'). However, we do know that the surveys conducted in these countries attempted to construct representative samples of the firm population using industry, establishment size, and region strata (World Bank, 2014). Thus, there is little reason to believe an explicit sampling bias related to treatment status was introduced as a result of the sampling process.

We also have no information on the content of the training (i.e. its quality), its duration, or its cost. Given this lack of information, combined with the relatively large numbers of firms surveyed, we must assume that the treatment received by firms represents the average quality of such services in Latin America. Therefore, in this paper, we assess the impact of the average quality-related training activity, and not in relation to specific programmes that may have been active during the period under study.

The panel comprises firm level surveys conducted in 2006 and 2010 in 14 Latin American countries. These were: Argentina, Bolivia, Chile, Colombia, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, and Uruguay. In total, the initial dataset contains 19,646 observations, with 10,430 from the 2006 surveys, and 9,216 from the 2010 surveys.

In order to retain only relevant observations for the analysis, we drop firms that were not surveyed in both years. This condition reduces the sample to 6,226 observations (or 3,113 firms), evenly split between both survey years. Table 1 describes all the variables used in this paper. Table 2 contains a range of summary statistics on the properties of the sample.

The right-most columns of Table 2 show the difference between the treated and non-treated groups for each indicator. With respect both to the export status[5] and the certification status of firms, it is clear from these two columns that the difference between the treated and non-treated groups increased over the time period under investigation. This is an early indication that treatment may have a positive effect.

[5] A firm is considered an exporter in our sample if 1% or more of its sales come from direct export.

Table 2 also gives summaries of several firm characteristics, namely the number of employees, the age of the firm, and the sales. Treated firms were systematically bigger than non-treated firms (by number of employees) and this difference grew over time. There are also significant differences between the age of the firms and the sales for both groups.

III. Statistical framework

This study aims to assess the impact of quality-related business trainings on firms' certification and export status. For this, we apply statistical methods from the program evaluation literature, where we define a quality-related business training as the "treatment". $Treat_i = 1$ denotes the treatment state, that is, the firm participated in quality-related business training, and $Treat_i = 0$ denotes the non-treatment state, that is, the firm did not participate in a quality-related business training. Y_{it} is the outcome associated with each state, for example being certified or being an exporter.

Our identification strategy takes advantage of the fact that not all firms in the sample received these training activities, so that we have a treatment and control group. Also, we have a wide range of firm-level information on firm characteristics before and after the treatment, in both 2006 and 2010, and we know that the treatment happened between 2007 and 2009.

i. Regression Model

To estimate whether quality-related trainings had an impact on firm performance, we adopt a difference-in-difference (DiD) OLS regression model.[6] The specification chosen empirically determines the effect of a treatment on an outcome, and can be written as follows:

Equation (1) $$Y_{it} = \alpha + \beta_1 Treat_i + \beta_2 Time_t + \beta_3 (Treat_i * Time_t)_{it} + \beta_4 X_{it} + \gamma_j + \gamma_S + \varepsilon_{it}$$

Here, subscript *i* represents individual firms and subscript *t* the year of the survey. The outcome variable, Y_{it}, is a binary variable which determines whether a firm holds an IRQC or is an exporter. $Treat_i$ is a treatment dummy which controls for time-invariant differences between the treated and non-treated groups. $Time_t$ is a time dummy which controls for time-variant differences between 2006 and 2010, and the interaction between $Treat_i$ and $Time_t$ isolates the effect of being treated over time on outcome variable Y_{it}. The coefficient of this interaction term is the main statistic of interest in this paper.

Finally, X_{it} is a vector of firm level controls, the composition of which varies according to the specific regression, while γ_j and γ_S are country and sector fixed effects, controlling for country and sector characteristics that remain constant over time and that apply to all firms in each country or sector. All control variables are set to their 2006 values, and not allowed to evolve. A brief description of the construction of the control variables is provided next.

[6] We prefer to rely on a simple linear probability model rather than a non-linear probit (or logit) model for two reasons. The first is that non-linear models suffer from the incidental parameter problem when fixed effects are included in the regressions, like ours. The second is that, both the outcome and independent variables of interest are dummy variables, and thus our coefficients of interest encounter no bounded probability problems, making the use of a probit (or logit) model unnecessary.

Firm size is defined by the natural logarithm of the number of full-time employees. Firm age and the experience of the top manager are both continuous variables, and are expressed in years. We also add the size of the locality the firm is situated in, as a control. If the firm is situated in the capital city, or a city of 1 million people or more, locality is set to 1; for all other localities it is set to 0.

Traditional standard errors are calculated on the assumption that all observations are independent of one another. While this assumption is mostly valid for our dataset, we might expect the error terms to cluster based on country, the surveys were run on a country-by-country basis. We cluster our errors according to a country dummy, and find that on the whole, clustering lowers the error term.

The DiD approach (developed by Marie L. Obenauer & Bertha von der Nienburg, 1915), has been greatly used in economics to assess changes in state laws and regulations (Lechner, 2010), the effect of training and other labour market programmes on several labour market outcomes (Ashenfelter, 1978; Ashenfelter & Card, 1985; Heckman, J. J. & Robb, R., 1986; Heckman & Hotz, 1989; Heckman, J. *et al.*, 1998; Blundell *et al.*, 2004), as well as the impact of immigration on the local labour market (Card, 1990), to name just a few. Over the last decade the DiD method has been significantly improved by Hansen, (2007b), (2007a) on issues of inference, focusing on what can be learned with various group, time period dimensions and serial independence in group-level shocks, by Athey & Imbens, (2006) on nonparametric approaches to difference-in-differences, and by Alberto Abadie, Alexis Diamond & Jens Hainmueller, (2007) on constructing synthetic control groups.

The DiD method relies on the parallel paths assumption, which posits that for the outcome variable under examination, the treated and non-treated groups evolve in parallel, prior to the treatment (Mora & Reggio, 2012). If the parallel paths assumption can be shown to be true, then the addition of control variables into the DiD regression are not needed. However, controls may still be useful to help explain which characteristics are associated with firms that selected treatment.

Unfortunately, the dataset does not contain information on the certification or exporter status of the firms surveyed prior to 2006. This makes it difficult to test the parallel paths assumption directly. We address this problem in more detail in Section V.ii.

IV. Results

i. The impact of treatment on certification status

Results from the difference-in-difference specification (Equation 1) using certification status as the outcome variable are reported in Table 3, model 1 (see Annex). Before executing the regression, firms which held an IRQC in 2006 were dropped. This is done to facilitate the interpretation of the DiD coefficient, which now exclusively represents how effective treatment was for firms transitioning from uncertified to certified status.

The DiD term is the interaction of treatment and time, and the coefficient for this variable is positive and significant. The result is robust to the iterative inclusion of fixed effects and controls. The value of the DiD coefficient remains stable at around 0.25(±0.019)[7]. In other words, 25.0% of treated firms received an IRQC due to treatment. But how does this compare to the untreated firms? The coefficient on the time term, 0.031(±0.009), indicates that 3.1% of untreated firms gained an IRQC between 2006 and 2010. To calculate the fraction of treated firms which had an IRQC in 2010, one simply adds the time and DiD coefficients, which gives 0.281. Calculating the ratio of treated to non-treated, we find that treated firms were 8.1 times more likely to gain an IRQC compared to untreated firms. In terms of controls, the firm's exporter status and (marginally) the manager's experience were significant. This suggests that firms which exported in 2006 with experienced managers are more likely to gain an IRQC than firms without these characteristics.

It is also possible to test whether treatment increases the probability that firms retain their IRQC (Table 3, model 2). One can do this by dropping all firms which did not hold an IRQC in 2006. The time term, -0.612(±0.066), indicates that only 38.8% of untreated firms retained their IRQC over the period of investigation. On the other hand, the DiD coefficient is 0.380(±0.060), meaning that 76.8% of treated firms retained their IRQC. Calculating the ratio of treated to untreated retention rates, we find that treated firms were 2.0 times more likely to retain their IRQC compared to untreated firms, or 38.0% of treated firms retained their IRQC due to treatment. In terms of controls, only the exporter status is significant, suggesting that firms who exported in 2006 were more likely to retain their IRQC.

We recall that the definition of treatment is *any services or programs to improve quality control or training to obtain quality certification.* It seems reasonable to assume that in the case where firms are uncertified in 2006, treatment is more likely to refer to *training to obtain quality certification*, whereas, in the case where firms already held an IRQC in 2006, treatment refers to *any services or programs to improve quality control.* This may help explain why treatment seems to be beneficial under both specifications.

It must be noted, however, that we do not have any information on the number of IRQCs firms in our sample held. As such, it is likely that some of them may have been pursuing a second IRQC, rather than trying to retain an existing IRQC.

ii. The impact of treatment on exporter status

Firms with competent quality management systems and the necessary IRQCs have a higher chance of exporting their goods and services (e.g. Goedhuys & Sleuwaegen, 2016). Thus, we investigate whether treatment increases the chances of firms becoming exporters. Before executing the difference-in-difference regression, firms which were exporters in 2006 were dropped. This facilitates the interpretation of the DiD coefficient, which now represents how effective treatment was for firms transitioning from non-exporter to exporter status.

The time term in Table 3 (model 3) shows that 6.0%(±0.1%) of untreated firms became exporters over the period of investigation. The DiD coefficient is 0.057(±0.018), indicating that 11.7% of treated firms became exporters over the same period. This means that treated firms were 2.0 times more likely to

[7] For the reader's convenience, we include statistical error on the coefficient derived from the linear regressions. Note, however, that this error does not capture all potential sources of error, such survey sampling bias, and as such, is likely to be an underestimate.

become exporters compared to untreated firms, or 5.7% of treated firms became exporters due to treatment.[8] In terms of controls, firm size and certification status in 2006 were significant. This suggests that larger firms which held an IRQC in 2006 were more likely to become exporters.

Similarly to what was done for certification, if one drops all the firms which were non-exporters in 2006, one can calculate the probability that treatment results in current exporters retaining their exporter status. From the time term in Table 3 (model 4), -0.428(±0.039), one can deduce that 57.2% of untreated firms retained their export status. This compared to 78.3% for the treated sample. Calculating the ratio, one finds that treated firms were 37% more likely to retain their exporter status as a result of treatment compared to untreated firms. This suggests that treatment does indeed help current exporters maintain their exporter status. Put another way, 21.1%(±5.2%) of treated firms retained their export status due to treatment. In terms of controls, none were significant.

iii. Does the effectiveness of treatment depend on firm size?

The richness of the dataset allows us to test additional hypotheses, such as whether DiD coefficient associated with gaining an IRQC for the first time varies according to firm-size. To test this, we split the sample by the number of full-time employees into 8 bins of equal size. The results are shown in Figure 1, where the DiD coefficients from the regressions are plotted against firm size.

From Figure 1 it is clear that as firm size increases, the average effect of the treatment on the treated rises from about 10% to 40%.[9] Fitting a trendline to the data results in linear fit with an R^2 value of 0.80. The p-values for the DiD coefficients are also shown, and they are below 0.01. The coefficient on the time term (i.e. the control group) for each of these regressions rises steadily from about 1% for the smaller firms to 8% for the largest.

A similar relationship is observed when splitting the sample by firm sales. The average effect of the treatment on the treated was more effective for firms with higher sales: the DiD coefficient was 16% for firms with sales of about $46,000, but rises to 45% for firms with sales of about $20,000,000 (see Figure 2). Repeating the analysis using export status as the outcome variable yielded no evidence of a dependence on firm size, although the larger errors (due to the smaller samples) make it more challenging to tease out a relationship. In summary, smaller firms do seem to find it harder to translate treatment into gaining a quality certificate. The reasons for this are discussed in Section VI.

iv. The impact of treatment on sales and exports

In the previous subsections, we investigated the impact of treatment on two outcomes; whether firms were more likely to gain/retain an IRQC or become/remain an exporter. However, we can also test the impact of treatment on a host of performance variables. The World Bank Enterprise Surveys

[8] We test multiple definitions for 'exporter', including varying the threshold from 1% of direct exports to 10%, and including indirect exports. We found that the values of the time and DiD terms did vary according to the specific definition, but the ratio of the two remained stable at around 1.87. All specifications produced significant coefficients.

[9] Testing on retaining certification (e.g. dropping firms with did not hold an IRQC in 2006) yielded no results of significance.

recorded the annual sales for enterprises surveyed in 2006 and 2010. Placing the natural logarithm of sales as the dependent variable, we can investigate the impact of treatment on sales.[10]

The results are shown in Table 4, model 1. For this case, no filtering by certification or exporter status was conducted. The coefficient on the DiD term is 0.197 (±0.064), meaning that treatment increased the sales of firms in the treatment group by an average of 0.197 in natural logs. This term is somewhat tricky to interpret for two reasons. The first is that an increase of 0.197 in natural logs translates to different dollar amounts depending on the initial log value. The second is that OLS regressions give information about the effects of the regressors at the conditional mean of the dependent variable only, meaning that the coefficient is only valid at one point in the log scale.

To aid with the interpretation of the impact of treatment on sales, we perform quantile regressions. Quantile regressions allow one to test the behaviour of the coefficients on the conditional distribution of the dependent variable (e.g. logged sales). This allows one to translate the resulting DiD coefficient into a specific dollar amount, and thereby calculate the percentage sales increase in dollar terms.

The resulting coefficients are transformed into percentage sales increases using the value of sales at the specified percentile, and plotted against sales (see Figure 3). The results show that treatment has a large and positive effect for firms on the lower end of the sales distribution, dropping steadily as sales volume increases. However, above a value of roughly $10 million, the DiD coefficient begins to climb again. In summary, treatment boosted the sales of firms with annual sales below $200,000 and above $10 million, with the region in between seeing only small boosts to sales. The sales boost observed for firms with high sales is an unexpected feature. However, it is hard to explain without more information on the type of treatment these firms underwent, compared to the other firms in the sample. We tested the impact of treatment on export sales, but the results were not significant.

v. The impact of treatment on job growth and productivity

If treatment led to an increase in sales, it stands to reason that firms may have simultaneously increased the number of employees. To investigate any potential effect, we follow the same analytical structure as for sales. Table 4, model 2 shows the impact of treatment on the natural logarithm of the number of full-time employees. The DiD coefficient is positive and significant, indicating that employment did rise as a result of treatment. To aid with the interpretation of this effect, we perform quantile regressions along the employment axis for quantiles between percentiles of 0.1 and 0.9, in steps of 0.05.[11] The resulting coefficients are transformed into percentage employment increases using the number of employees at the specified percentile, and plotted against the number of full-time employees (see Figure 4).

At first glance, the results are somewhat surprising. For firms with, for example 11 employees, the results indicate that one should expect a 5.1%(±0.9%) rise in the number of employees. This translates

[10] Unlike in previous subsections, we do not drop firms according to their certification or exporter status, as we are interested in the general effect of treatment.
[11] Larger steps are required as the natural log of small discreet numbers rises in steps.

into an increase of 0.5 employees. The percentage increase in the number of employees falls as the size of the firm increases, but in terms of absolute numbers, the employee count rises.

For firms with 50 employees, treatment results in an average increase of 1 employee. At around 100 employees, the percentage increase in the number of employees starts to rise again, causing the absolute number of additional employee count to rise rapidly. For firms with 150 employees, our calculations imply an average of four extra employees were hired due to treatment. The behaviour of the DiD coefficient for firms of large sizes is consistent with what is observed for firms with high annual sales, namely that the DiD coefficient rises after falling somewhat for medium sized companies. Again, it is hard to explain the source of this rise without more information on the treatments selected by these larger firms. We also tested the impact of treatment on productivity[12] and capacity utilization, but the results were not significant.

V. Robustness checks

i. Propensity score matching

Matching methods are based on the idea that directly matching groups of treatment and control group observations, based on a combination of their observable characteristics, allows one to calculate the counterfactual change in the treatment group if there were no treatment. The propensity score matching method (Rosenbaum, P.R. & Rubin, D. B., 1983), collapses a vector of pre-treatment characteristics, *X*, into a single variable (i.e. the propensity score), and uses this as the matching estimator.

The central advantage of propensity score matching (PSM) is that, by combining covariates into a single score, it can balance treatment and non-treatment groups without losing a large number of observations. The alternative is to balance across a host of covariates simultaneously, which in turn needs large numbers of observations to avoid the “dimensionality problem”.[13] However, PSM also has several disadvantages, namely that it accounts only for observable covariates, and requires a good model which explains why units were selected into treatment. For this part of the paper, we follow the approach used by Huttunen, (2007).

The first step is to calculate the propensity scores. In this study, the propensity score is the conditional probability of an enterprise investing in quality control or attending a quality-related training. This is a binary choice model, and is described below:

Equation (4) $$T_{it} = \begin{cases} 1 & if \beta X_{it}+\gamma_j+\gamma_s+\varepsilon_{it}>0 \\ 0 & otherwise \end{cases}$$

where T_{it} is a binary variable that defines whether firm *i* received a quality-related training at time t, and X_{it} is a vector of controls. To control for unobservable country and sector effects, γ_j and

[12] In this case, productivity was simply defined as the sales per full-time employee.

[13] The dimensionality problem is a sparsity problem. As one increases dimensionality, the parameter space increases so fast that for a given level of statistical significance per volume, the number of observations would have to increase exponentially.

γ_S are included as fixed effects. The propensity score is estimated by a parametric logit model where the controls are fixed to the values they had in 2006 (see Table 5).

The next step is to compare the difference in the outcome before and after the treatment of treated firms with the difference in the outcome of the non-treated firms over the same period. This controls for the possible bias that is due to the effect of unobservable factors (Heckman, Ichimura & Todd, 1997). The formal specification for the Average Effect of Treatment on Treated (ATT) for all types of difference-in-differences matching estimators can be written as:

Equation (5) $$\widehat{ATT}(S) = \sum_{i \epsilon T \cap S_P} \frac{1}{N^T} \left[\left(Y(1)_{it} - Y(0)_{it-1}\right) - \sum_{j \epsilon C \cap S_P} \omega_{ij} \left(Y(0)_{jt} - Y(0)_{jt-1}\right) \right]$$

where $Y(1)_i$ is the treatment outcome for firm *i*, $Y(0)_j$ is the non-treatment outcome for firm *j* (comparison group outcome), *t* is time, N^T is the number of firms in the treatment group, *T*, *C* denotes the set of control units, S_P denotes the region of 'common support',[14] and ω_{ij} is the weight used to match control firms with each treatment firms.

We tried several matching estimators, including nearest neighbour (NN) and kernel matching. While NN matching uses only those control group observations that are closest to treated units, kernel matching uses all control group observations, but weights each observation according to its distance from the treated unit. For kernel matching, we use an epanechnikov weighting function. To select the bandwidth, we follow the method set out by Loader (1999). In brief, this entails plotting the outcome variable against the propensity score and using the epanechnikov function to fit a curve. The fit is altered by adjusting the bandwidth, and the curve with the lowest mean square error is chosen as the 'best-fit' bandwidth.

Results from the propensity score matching method are listed in Table 6. Four outcomes are tested as robustness checks on the results reported in Section IV: the effectiveness of treatment for firms acquiring IRQC, retaining IRQC, becoming exporters, and retaining exporter status. When matching, we test whether the balancing property holds, that is, whether observations in the matched and unmatched groups have the same distribution of observable characteristics independent of treatment status. We find that our samples are well balanced.[15]

For the first outcome, the coefficients derived from using nearest neighbour and kernel matching are all remarkably similar, varying between 0.23 – 0.35. Furthermore, these coefficients are consistent with the coefficient derived from the DiD regression model presented in Section IV (0.25±0.019). For the second outcome, the coefficients derived using our three matching estimators vary between 0.39-0.41. Taking the standard errors into account, which range from 0.05 to 0.07, these results are again consistent with the DiD regression model presented in Section IV (0.38±0.06).

For the third outcome, transitioning from non-exporter to exporter, the coefficients vary from 0.02-0.04, somewhat lower than the coefficients resulting from the regression method (0.057±0.06). However, the errors on the coefficients generated via the PSM method are large, and consequently

[14] This is the region where the treated and non-treated units have the same propensity score values.
[15] Balance tables may be furnished upon request.

the t-statistic is low. As a result, the coefficients are in line with the DiD regression model, but do not provide independent confirmation of the result.

For the fourth outcome tested, the coefficients derived from propensity score matching (which range from 0.06-0.008) are not significant for the nearest neighbour methods or using the kernel estimator. These coefficients are also lower than the one derived from the DiD regression model (0.211±0.034). In conclusion, for two of the four specifications checked, we found that propensity score matching provides complementary evidence for the results presented in Section IV.

ii. Revisiting the parallel paths assumption

In section III, we noted that we have no data for our sample prior to 2006. This makes it difficult to test the parallel paths assumption directly. This is a limitation of the analysis presented in this paper, however, under some specific assumptions, we may be able to assess the likelihood that the parallel paths assumption is satisfied.

In the DiD regression specification, the time term captures the time evolution of the untreated group with respect to the outcome variable under examination. Assuming that this behaviour does not change significantly with time, one can use the coefficient which results from the time term as an indicator of whether the parallel paths assumption is satisfied.

For example, under the specification where we assessed the likelihood that treatment would result in firms gaining an IRQC, the coefficient on the time term was 0.031, or put differently, over a four year period 3.1% of the control group gained an IRQC. This means that, on a per year basis, approximately 0.8% of firms gained an IRQC. Therefore, provided this is a good indicator of the churn rate at which firms in this sample acquire IRQCs, it is likely that in the years preceding 2006 the overwhelming majority of the sample did not hold an IRQC, thereby satisfying the parallel paths assumption.

This approach may be repeated for the three other headline results presented in this paper. For firms retaining their IRQC, gaining exporter status, and retaining exporter status, the churn rates are 15.3%, 1.5%, and 10.7%. The churn rates for retaining certification and exporter status are quite high, and may mean that the parallel paths assumption has not been met in these cases, but is likely to have been met in the gaining exporter status case.

VI. Discussion

In this paper, we conduct a difference-in-difference estimation to investigate whether using quality control services or receiving a quality-related training results in a higher chance of acquiring or retaining an internationally recognised quality certificate, or becoming or remaining an exporter. We also investigate whether such trainings resulted in increased sales, employment, and other firm performance metrics.

Results from Section IV.i show that firms which underwent treatment were 8.1 times more likely to acquire an IRQC compared to untreated firms, and 2.1 times more likely to retain their certification status, if they held a IRQC in 2006, compared to untreated firms. Treatment also had an impact on firm's export status, with those that underwent treatment being 2.0 times more likely to transition from non-exporter to exporter status. For firms which were already exporters in 2006, treatment led to a 37% increased chance that they remained exporters in 2010.

Although the treatment question and the outcome variable are relatively well matched (i.e. one would expect a '*training to obtain quality certification*' to result in increased propensity to hold an IRQC and consequently to export), there are some limitations. The first of these is that the treatment question also includes the concept of '*improve[d] quality control*' which one would not necessarily expect to lead to the acquisition of a quality certificate. Indeed, it is more in line with what one might expect from firms which continually invest in quality. It is unfortunate that the treatment question contains two concepts, but this latter concept may help explain the variety of result found in this paper.

Treatment seems to have a strong positive effect on firms which already held an IRQC in 2006. However, a closer look at this sub-sample reveals that 75% of these firms were treated. As such, being untreated is the exception, and may in-fact be a signal that the firm is about to lose its IRQC as it is unable or unwilling to invest in '*improve[d] quality control*'. The same is true of the 'retention of exporter status' result. Here, around 59% of these firms were treated. This highlights the importance of designing survey questions with only one clear goal in mind.

Results from Section IV.iii show that firm size, defined in terms of number of full-time employees, matters. Splitting the sample into 8 bins according to firm-size, we find that the larger the firm, the more likely it is that treatment results in firms gaining an IRQC. This also applies to size if defined by sales. This is in line with the economic literature, indicating that, apart from access to knowledge, firms also need to be able to absorb knowledge (Cohen and Levinthal 1990; Daniel H. Kim 1998).

In fact, while quality certification has beneficial returns - such as improved legitimacy, cost reduction and increased trade opportunities – adopting certification is costly and highly demanding in terms of personnel and time resources. For these reasons SMEs might find it difficult to get certified (Bansal & Hunter, 2003; Raines, 2002; Prakash & Potoski, 2007). Larger firms have more resources to make use of the training received and get a quality certificate than small firms.

In fact, even in cases where SMEs acquire the knowledge about standards and quality certification, and even get to the stage of adapting to the market's requirements, they still need to demonstrate compliance. Many developing countries do not have domestic certification bodies recognised in the export market. SMEs are then compelled to use foreign certification bodies and the costs rise accordingly. In such cases, only the larger firms have the financial resources to complete the certification process.

Related to absorptive capacities are financial constraints, cited by Graffham, Karehu & MacGregor, (2007) as the main obstacle to establish and maintain quality standards in Kenya, rather than the lack of technical knowledge. In line with this, standards are found to be more beneficial for middle-income farmers, who can bear the related costs of compliance, compared to their low-income counterparts (Hansen & Trifkovic, 2014).

Another explanation for the finding that larger firms benefit more from treatment may be down to the quality of the trainings. If higher quality trainings are in turn more expensive, then they should be both more accessible to larger firms and more likely to be selected by larger firms. Unfortunately, there is no information in the dataset to control for the quality of trainings.

Section IV.iv and Section IV.vi investigate whether treatment led to a positive outcomes on a variety of firm performance metrics. The results show that firm sales increased after treatment, although the rise was statistically significant for only the firms with the lowest and highest annual sales. We tested the impact of treatment on export sales, but the results were not significant, owing mainly to the larger errors as a result of exporters making up only one-third of the sample. We also tested the impact of treatment on employment, and found treatment resulted in firms of all sizes hiring more full-time employees. The size of the effect follows a U-shape pattern, with the smallest firms increasing employment by 4-7%, medium sized firms 2-3%, and large firms by 2-4%. We also tested the impact of treatment on productivity and capacity utilization, but found no results of significance.

The evidence supporting positive firm performance outcomes has two important implications. The first is that, firm managers who choose to invest in quality management are likely to see sales rise, as well as the size of their company expand. The lack of evidence for improvements in capacity utilization and productivity may indicate that treatment mostly helped firms attract new customers or secure larger orders from existing ones. Presumably, the acquisition or retention of an IRQC, which signals competence and quality to buyers, and in turn instils confidence, is the main mechanism behind the increase in sales. Interestingly, treatment led to strong sales increases for those firms which were gaining IRQCs for the first time, whereas those firms which merely retained their IRQC saw no increase in sales (see Table 4, model 3 and 4). This suggests that it is the first time acquisition of an IRQC which is driving the boost seen in sales. Unfortunately, we cannot assess the impact of treatment on profits, nor the return on investment of treatment, as Enterprise Surveys did not record such information.

The second implication of these results regards policymakers. Policymakers seeking to improve the quality management of firms will be content to learn that the average treatment used by firms in Latin America is effective, although whether such treatment represents value for money is another question. Nevertheless, it is true that treatment is much less effective for smaller firms. The reasons for this are not clear, but as described earlier, lower absorptive capacities and limited financial resources are likely to play an important role. Policymakers may want to direct resources towards these firms with tailored quality management programmes and implementation support. Doing so would contribute to the overall health of the SME sector, and in turn lower certification-related barriers to entry. Finally, policymakers will also be content with the finding that treatment increased the number of exporters even if direct exports themselves did not increase.

VII. Conclusions

In this paper we present the results of a difference-in-difference estimation aimed at investigating the effects of quality-related trainings in 14 Latin American countries. Our results indicate that, compared to control groups, programs to improve quality control or quality-related training activities help firms acquire IRQC (8.1 times more likely), retain their IRQC (2.0 times more likely), transition from non-exporter to exporter status (2.0 times more likely), and to maintain their exporter status (37% more likely).

Further regressions also indicate that larger firms, defined either in terms of number of full-time employees or in terms of annual sales, seem to benefit more from quality-related training than smaller firms. The lower absorptive capacities of smaller firms may explain why treatment on smaller firms is less effective. At the same time, smaller firms may be selected lower quality treatments. Finally, treatment led to an increase in sales and employment, but no significant increase in exports, productivity or capacity utilization.

What can policymakers draw from the above results? The first is that, in general, investments in quality control and quality-related trainings in Latin American led to a variety of beneficial outcomes for these firms. Value for money, however, is another question, and unfortunately we have no way of assessing this dimension. For any given firm in the treated group, the probability that they would gain an IRQC was about 28.1% compared to 3.1% in the control group; a large difference. Offering training to small firms may be less effective than doing so for larger firms. However, because we cannot control for the quality of the training, it may be the case that smaller firms simply select inferior trainings. Furthermore, policymakers may wish to implement specific programmes to help SMEs improve quality, which in turn should help reduce certification related barriers to entry for these firms.

VIII. References

Ahmed, Z.U., Julian, C.C. & Majar, A.J. (2006) Export Incentives and International Entrepreneurship in Malaysian Firms. *The International Journal of Entrepreneurship and Innovation*. [Online] 7 (1), 49–57. Available from: doi:10.5367/000000006775870497.

Alberto Abadie, Alexis Diamond & Jens Hainmueller (2007) Synthetic control methods for comparative case studies: estimating the effect of California's Tobacco Control Program. *National Bureau of Economic Research*.

Ashenfelter, O. (1978) Estimating the Effect of Training Programs on Earnings. *The Review of Economics and Statistics*. [Online] 60 (1), 47. Available from: doi:10.2307/1924332.

Ashenfelter, O. & Card, D. (1985) Using the Longitudinal Structure of Earnings to estimate the Effect of Training Programs. *The Review of Economics and Statistics*. 67 (4), 648–660.

Athey, S. & Imbens, G. (2006) Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica*. 74 (2), 431–497.

Bansal, P. & Hunter, T. (2003) Strategic Explanations for the Early Adoption of ISO 14001. *Journal of Business Ethics*. 46 (3), 289–299.

Bernard, A.B. & Jensen, J. B. (2004) Exporting and Productivity in the USA. *Oxford Review of Economic Policy*. [Online] 20 (3), 343–357. Available from: doi:10.1093/oxrep/grh020.

Bilanakos, C., Green, C.P., Heywood, J.S. & Theodoropoulos, N. (2014) Do Dominant Firms Provide More Training? *Journal of Economics & Management Strategy*. [Online] 26 (1), 67–95. Available from: doi:10.1111/jems.12177.

Blundell, R., Dias, M.C., Meghir, C. & Reenen, J.V. (2004) Evaluating the Employment Impact of a Mandatory Job Search Program. *Journal of the European Economic Association*. 2 (4), 569–606.

Card, D. (1990) The Impact of the Mariel Boatlift on the Miami Labor Market. *Industrial and Labor Relations Review*. 43 (2), 245–257.

Cohen, W.M. & Levinthal, D.A. (1990) Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly. Special Issue: Technology,Organizations, and Innovation.* [Online] 35 (1), 128–152. Available from: doi:10.2307/2393553.

Dearden, L., Reed, H. & Reenen, J. V. (2005) *The Impact of Training on Productivity and Wages: Evidence from British Panel Data*. CEP Discussion Paper No 674. London, Centre for Economic Performance.

Dearden, L., Reed, H. & Reenen, J. V. (2000) *Who Gains when Workers Train? Training and Corporate Productivity in a Panel of British Industries*. [Online]. Available from: doi:10.1920/wp.ifs.2000.0004 [Accessed: 7 August 2018].

Freixanet, J. (2012) Export Promotion Programs: Their Impact on Companies' Internationalization Performance and Competitiveness. *International Business Review*. [Online] 21 (6), 1065–1086. Available from: doi:10.1016/j.ibusrev.2011.12.003.

Geldres-Weiss, V.V. & Carrasco-Roa, J.A. (2016) Impact Evaluation of National Export Promotion Programs on Export Firms using Contrast Groups. *International Journal of Export Marketing*. 1 (1), 77–95.

Gençtürk, E.F. & Kotabe, M. (2001) The Effect of Export Assistance Program Usage on Export Performance: A Contingency Explanation. *Journal of International Marketing*. [Online] 9 (2), 51–72. Available from: doi:10.1509/jimk.9.2.51.19886.

Goedhuys, M. & Sleuwaegen, L. (2016) International standards certification, institutional voids and exports from developing country firms. *International Business Review*. [Online] 25 (6), 1344–1355. Available from: doi:10.1016/j.ibusrev.2016.04.006.

Görg, H., Henry, M. & Strobl, E. (2008) Grant Support and Exporting Activity: Evidence from Irish Manufacturing. *Review of Economics and Statistics*. 90 (1), 168–174.

Graffham, A., Karehu, E. & MacGregor, J. (2007) Impact of EurepGAP on small-scale vegetable growers in Kenya. *Fresh Insights*. 6.

Hansen, C.B. (2007a) Asymptotic properties of a robust variance matrix estimator for panel data when is large. *Journal of Econometrics*. [Online] 141 (2), 597–620. Available from: doi:10.1016/j.jeconom.2006.10.009.

Hansen, C.B. (2007b) Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects. *Journal of Econometrics*. [Online] 140 (2), 670–694. Available from: doi:10.1016/j.jeconom.2006.07.011.

Hansen, H. & Trifkovic, N. (2014) Food standards are good - for middle-class farmers. *World Development*. 56, 226–242.

Heckman, J., Ichimura, H., Smith, J. & Petra Todd (1998) Characterizing Selection Bias Using Experimental Data. *Econometrica*. 66 (5), 1017–1098.

Heckman, J. J. & Robb, R. (1986) Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatment on Outcomes. In: *Drawing Inferences from Self-Selected Samples*. Springer-Verlag New Youk Inc. pp. 63–106.

Heckman, J.J. & Hotz, V.J. (1989) Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training. *Journal of the American Statistical Association*. [Online] 84 (408), 862–874. Available from: doi:10.2307/2290059.

Heckman, J.J., Ichimura, H. & Todd, P.E. (1997) Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*. [Online] 64 (4), 605–654. Available from: doi:10.2307/2971733.

Huttunen, K. (2007) The effect of foreign acquisition on employment and wages: Evidence from Finnish establishments. *The Review of Economics and Statistics*. 89 (3), 497–509.

ITC (2016) *SME Competitiveness Outlook: Meeting the Standard for Trade*. SME Competitiveness Outlook 2016. Geneva, Switzerland, ITC.

Kim, D. H. (1998) The Link between Individual and Organizational Learning. *The Strategic Management of Intellectual Capital*. 41–62.

Kotler, P. & Gertner, D. (2002) Country as Brand, Product, and Beyond: A Place Marketing and Brand Management Perspective. *Journal of Brand Management*. 9 (4/5), 249–261.

Lechner, M. (2010) The Estimation of Causal Effects by Difference-in-Difference Methods Estimation of Spatial Panels. *Foundations and Trends in Econometrics*. [Online] 4 (3), 165–224. Available from: doi:10.1561/0800000014.

Lederman, D., Olarreaga, M. & Payton, L. (2009) Export Promotion Agencies Revisited. *World Bank Policy Research Working Paper*. 5125, 31.

Lederman, D., Olarreaga, M. & Zavala, L. (2016) Export promotion and firm entry into and survival in export markets. *Canadian Journal of Development Studies*. [Online] 37 (2), 142–158. Available from: doi:10.1080/02255189.2016.1131671.

Leonidou, L.C. (2011) National export-promotion programs as drivers of organizational resources and capabilities: Effects on strategy, competitive advantage, and performance. Strategic Direction. *Strategic Direction*. 27 (10), 1–29.

Loader, C.R. (1999) Bandwidth selection: classical or plug-in? *The Annals of Statistics*. [Online] 27 (2), 415–438. Available from: doi:10.1214/aos/1018031201.

Martincus, C.V. & Carballo, J. (2010) Entering new country and product markets: does export promotion help? *Review of World Economics*. 146 (3), 437–467.

Martincus, C.V., Carballo, J. & Gallo, A. (2010) The impact of export promotion institutions on trade: Is it the intensive or the extensive margin? *IDB Working Paper Series*. 199, 15.

Mora, R. & Reggio, I. (2012) Treatment Effect Identification Using Alternative Parallel Assumptions. *Working Paper 12-33. Economic Series (48)*. 34.

Moretti, E. (2004) Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*. [Online] 121 (1–2), 175–212. Available from: doi:10.1016/j.jeconom.2003.10.015.

Obenauer, M. L. & Von der Nienburg, B. (1915) Effect of Minimum-Wage Determinations in Oregon. *Bulletin of the U.S. Bureau of Labor Statistics*. 1 (3), 5–8.

Otsuki, T. (2011) Effect of ISO standards on exports of firms in Eastern Europe and Central Asia: An application of the control function approach. *Osaka School of International Public Policy: Discussion paper*. (DP-2011-E-005), 11.

Prakash, A. & Potoski, M. (2007) Collective action through voluntary environmental programs: A club theory perspective. *Policy Studies Journal*. 35 (4), 773–792.

Raines, S.S. (2002) Implementing ISO 14001—an International Survey Assessing the Benefits of Certification. *Corporate Environmental Strategy*. [Online] 9 (4), 418–426. Available from: doi:10.1016/S1066-7938(02)00009-X.

Rosenbaum, P.R. & Rubin, D. B. (1983) The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*. 70 (1), 41–55.

Terlaak, A. & King, A.A. (2006) The effect of certification with the ISO 9000 Quality Management Standard: A signaling approach. *Journal of Economic Behavior & Organization*. [Online] 60 (4), 579–602. Available from: doi:10.1016/j.jebo.2004.09.012.

Viadiu, M.F. & Fransi, E.C. (2005) A Study of the ISO 9000 Certification Process: Consultant Profiles and Company Behaviour. *Managing Service Quality: An International Journal*. [Online] 15 (3), 290–305. Available from: doi:10.1108/09604520510597836.

World Bank (2014) Implementing Enterprise Surveys in Latin America and the Caribbean. *Latin America and the Caribbean Series Note*. 10, 10.

Zwick, T. (2006) The Impact of Training Intensity on Establishment Productivity. *Industrial Relations*. [Online] 45 (1), 26–46. Available from: doi:10.1111/j.1468-232X.2006.00412.x.

IX. Annex

Table 1: Description of variables

Name	Description	Mean	Std. dev
Outcome variables			
Certification status	A binary variable equal to (1) if the firm had an internationally-recognized quality certificate, and (0) if it did not.	0.22	0.42
Export status	A binary variable equal to (1) if the firm is an exporter, and (0) if it is not. A firm is considered an exporter in our sample if 1% or more of its sales come from direct export.	0.22	0.41
Independent variables			
Treat	A treatment dummy which controls for time-invariant differences between the treated and non-treated firms. The question asked was, *"Over the last three years, did this establishment use any services or programs to improve quality control or training to obtain quality certification?"*	0.44	0.50
Time	A time dummy which controls for time-variant differences between 2006 and 2010.	0.50	0.50
Treat*Time	An interaction between the treatment and time dummies, which isolates the effect of being treated over time on an outcome variable. Put differently, the resulting coefficient can be interpreted as the average effect of being treated for the treated group.	-	-
Fixed effects			
Sector	A two-step categorical variable indicating the sector the firm was operating in, in 2006: (1) is equal to manufacturing and (2) to services.	-	-
Country	A variable indicating which country the firm is operating in: (0) Argentina, (1) Bolivia, (3) Chile, (4) Colombia, (5) Ecuador, (6) El Salvador, (7) Guatemala, (8) Honduras, (9) Mexico, (10) Nicaragua, (11) Panama, (12) Paraguay, (13) Peru, (14) Uruguay.	-	-
Firm-level controls			
Firm size (2006)	The natural logarithm of the number of full-time employees in 2006.	3.46	1.39
Sales (2006)	The natural logarithm of annual sales for 2006 in thousands. Sales were converted from local currency to USD using the real annual exchange rate.	7.02	2.10
Firm age (2006)	A continuous variable, defined simply by the age of the firm in 2006.	26.1	20.6
Manager's experience	A continuous variable, defined simply by the number of years of management experience the top manager has.	20.6	10.6
Size of locality	A binary variable indicating if the firm is situated in the capital city or a city of 1 million people or more (1), or a smaller locality (0).	0.76	0.43
Certification status (2006)	A binary variable equal to (1) if the establishment had an internationally-recognized quality certificate in 2006, and (0) if it did not.	0.20	0.40
Exporter status (2006)	A binary variable equal to (1) if the establishment is an exporter in 2006, and (0) if it was not. A firm is considered an exporter in our sample if it 1% or more of its sales come from direct export.	0.22	0.42

Table 2: Descriptive statistics for the sample by treatment status

Indicator	Treated		Non-treated		Difference
	Observations	Mean	Observations	Mean	
2006: Certification status	2716	0.38	3444	0.09	0.29
2010: Certification status	2718	0.51	3452	0.08	0.43
2006: Export status	2732	0.31	3472	0.17	0.15
2010: Export status	2732	0.33	3472	0.15	0.18
2006: Number of employees	2722	183	3462	66	118
2010: Number of employees	2720	257	3458	84	174
2006: Firm age (years)	2700	28.2	3422	24.4	3.72
2010: Firm age (years)	2722	32.4	3458	27.9	4.51
2006: Sales (thousands USD; ln)	2524	7.40	3132	6.22	1.18
2010: Sales (thousands USD; ln)	2426	8.09	2962	6.69	1.40
2006: Manager's experience (years)	2634	20.4	3286	20.8	-0.38
2010: Manager's experience (years)	2524	22.2	3248	22.6	-0.41
2006: Locality	2732	0.75	3472	0.77	-0.02
2010: Locality	2732	0.75	3472	0.77	-0.02

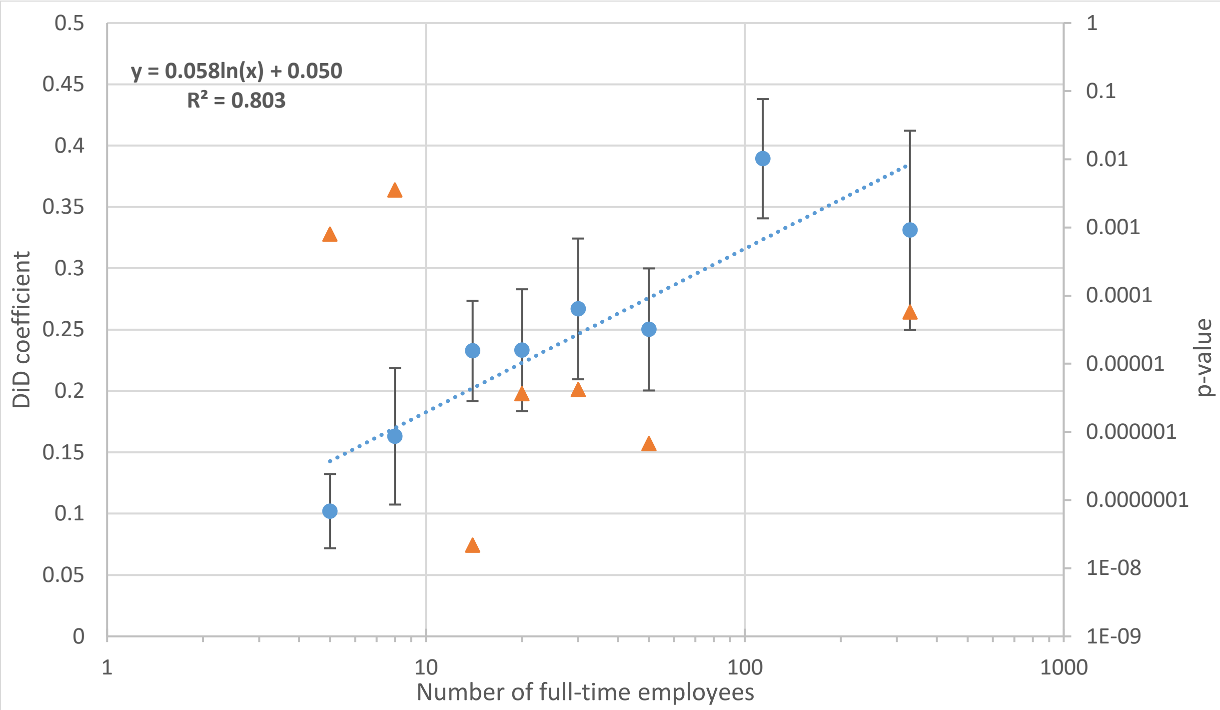
Note: In 2006 there were technically no 'treated' or 'non-treated' firms. Therefore, the label is retroactively applied to identify sample differences.

Table 3: The impact of treatment on gaining or retaining quality certification or exporter status

	Outcome variable (data subset)			
VARIABLES	Certification status (Gain certification)	Certification status (Retain certification)	Export status (Become exporter)	Export status (Remain exporter)
	Model 1	Model 2	Model 3	Model 4
Treat	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Time	0.031***	-0.612***	0.060***	-0.428***
	(0.009)	(0.066)	(0.008)	(0.039)
Treat*Time	0.250***	0.380***	0.057***	0.211***
	(0.019)	(0.060)	(0.018)	(0.034)
Ln(Firm size) (2006)	-0.000	0.023	0.017***	0.004
	(0.004)	(0.021)	(0.006)	(0.010)
Ln(sales) (2006)	0.006	-0.003	-0.001	0.008
	(0.004)	(0.011)	(0.004)	(0.006)
Manager's Experience (2006)	-0.001**	0.000	-0.000	-0.001
	(0.000)	(0.001)	(0.000)	(0.001)
Locality (2006)	0.011	0.009	-0.009	-0.049*
	(0.007)	(0.025)	(0.010)	(0.027)
Firm age (2006)	-0.000	-0.001	-0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Exporter status (2006)	0.037***	0.059***	-	-
	(0.013)	(0.021)	-	-
Certification status (2006)	-	-	0.027**	0.026
	-	-	(0.013)	(0.021)
Sector effects	Yes	Yes	Yes	Yes
Country effects	Yes	Yes	Yes	Yes
Constant	0.000	1.000	0.000	1.000
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	3,789	941	3,728	1,021
RMSE	0.2133	0.3017	0.1927	0.3253

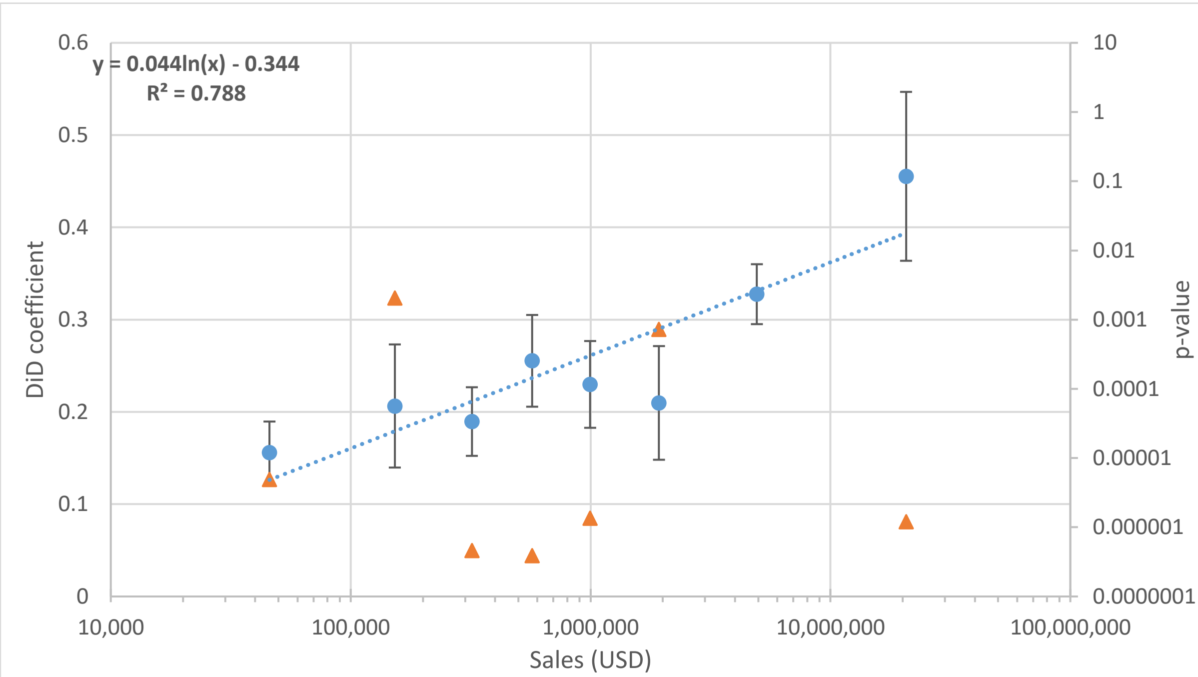
Note: Robust standard errors are in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

Figure 1: Does the effectiveness of treatment depend on the number of employees?



Note: Before executing the regression, firms which held an IRQC in 2006 were dropped. This is done to facilitate the interpretation of the DiD coefficient, which now exclusively represents how effective treatment was for firms transitioning from uncertified to certified status. The DiD coefficient is represented by the blue dots with error bars. The p-value of the coefficients are represented by cyan triangles.

Figure 2: Does the effectiveness of treatment depend on firm sales?



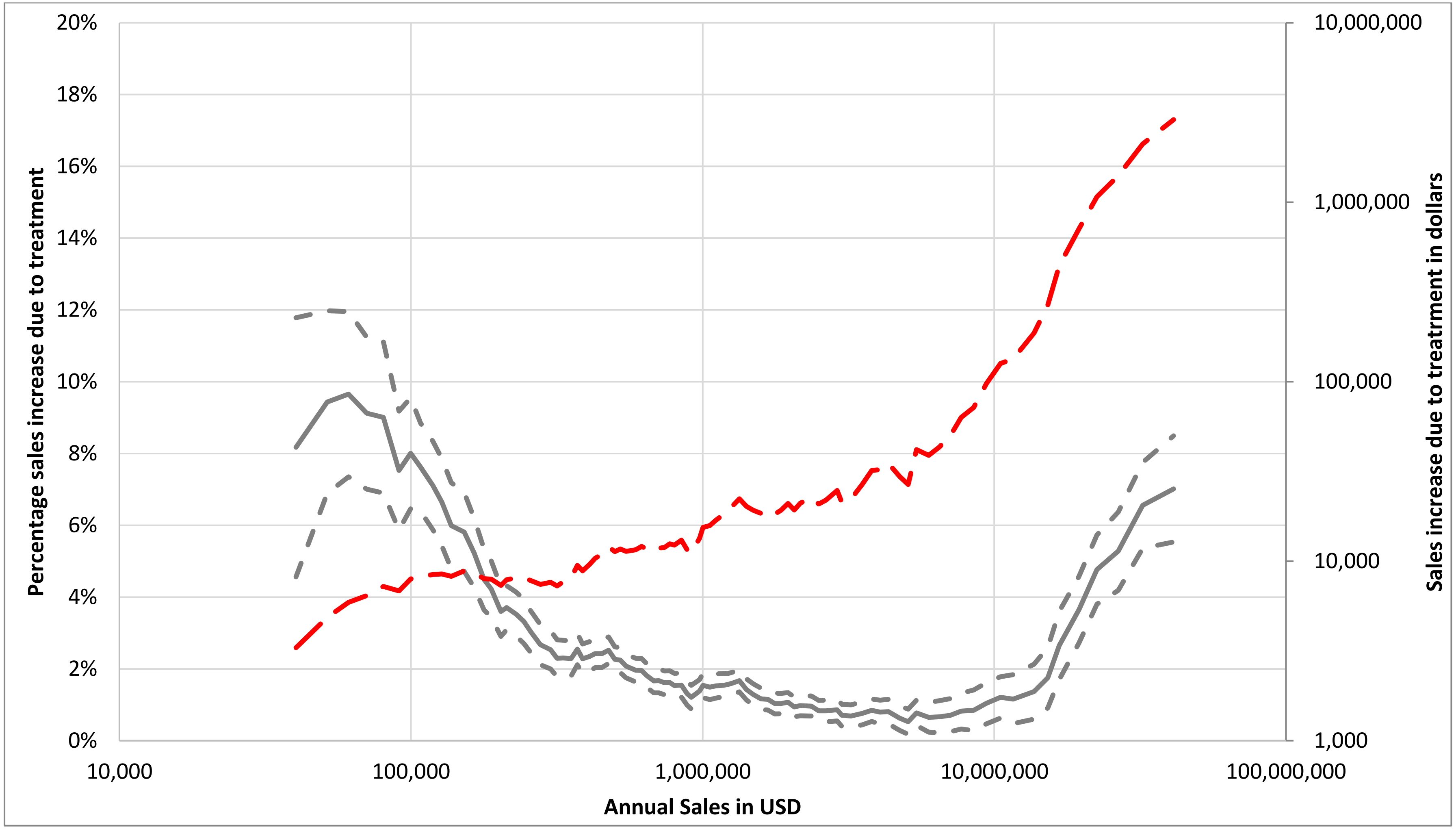
Note: Before executing the regression, firms which held an IRQC in 2006 were dropped. This is done to facilitate the interpretation of the DiD coefficient, which now exclusively represents how effective treatment was for firms transitioning from uncertified to certified status. The DiD coefficient is represented by the blue dots with error bars. The p-value of the coefficients are represented by cyan triangles.

Table 4: Does treatment improve performance outcomes?

VARIABLES	Sales (ln) (No filtering)	Employment (ln) (No filtering)	Sales (ln) (Gained certification)	Sales (ln) (Retain certification)
	Outcome variable (data subset)			
	Model 1	**Model 2**	**Model 3**	**Model 4**
Treat	0.032	0.035**	0.035	0.109
	(0.033)	(0.013)	(0.029)	(0.067)
Time	0.420***	0.063	0.392***	0.712**
	(0.086)	(0.036)	(0.094)	(0.316)
Treat*Time	0.197***	0.138***	0.257**	-0.149
	(0.064)	(0.041)	(0.088)	(0.271)
Ln(Firm size) (2006)	0.229***	0.857***	0.223***	0.229***
	(0.048)	(0.018)	(0.037)	(0.076)
Ln(sales) (2006)	0.783***	0.051***	0.791***	0.758***
	(0.034)	(0.009)	(0.024)	(0.073)
Manager's Experience (2006)	-0.001	-0.000	-0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.002)
Locality (2006)	0.017	0.024	-0.007	0.160
	(0.028)	(0.020)	(0.024)	(0.092)
Firm age (2006)	-0.000	-0.000	-0.001	0.001
	(0.001)	(0.000)	(0.001)	(0.001)
Certification status (2006)	0.116**	0.029	-	-
	(0.046)	(0.028)	-	-
Exporter status (2006)	-0.022	-0.032	-0.024	-0.007
	(0.036)	(0.025)	(0.036)	(0.073)
Sector effects	Yes	Yes	Yes	Yes
Country effects	Yes	Yes	Yes	Yes
Constant	0.705***	0.204***	0.711***	0.867***
	(0.071)	(0.033)	(0.085)	(0.191)
Observations	4,860	5,145	3827	1,033
R-squared	0.864	0.875	0.844	0.841

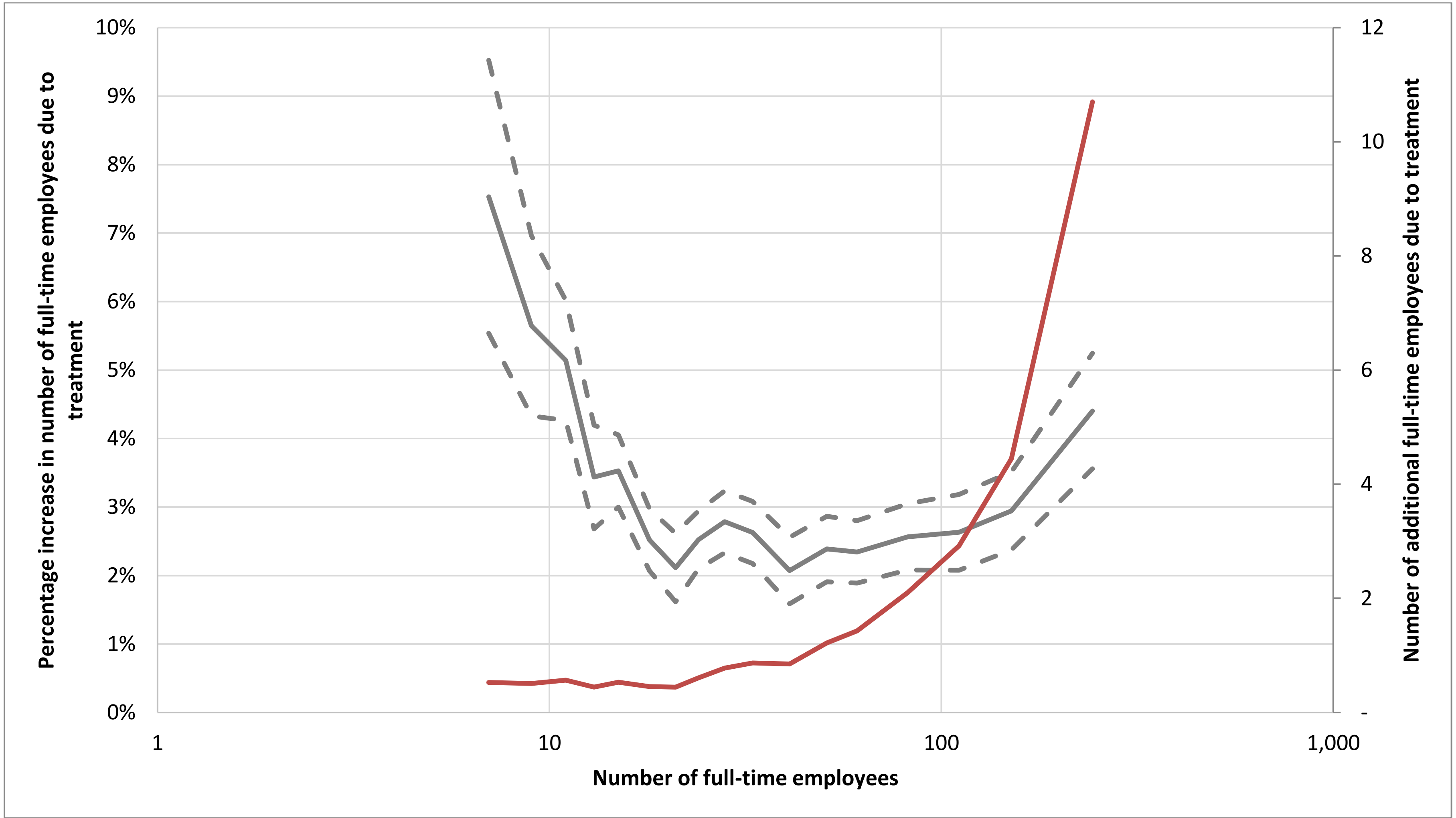
Note: Robust standard errors are in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

Figure 3: The increase in sales due to treatment



Note: The values in this plot were calculated by executing quantile regressions between percentiles of 0.05 and 0.95, in steps of 0.01. The solid grey line is the value of the DiD coefficient, transformed into a percentage sales increase. The dashed grey lines are the one standard deviation errors. The solid red line is the sales increase in dollars.

Figure 4: The increase in employment due to treatment



Note: The values in this plot were calculated by executing quantile regressions between percentiles of 0.1 and 0.90, in steps of 0.05. The solid grey line is the value of the DiD coefficient, transformed into a percentage employment increase. The dashed grey lines are the one standard deviation errors. The solid red line is the value of the number of additional full-time employees due to treatment.

Table 5: Results for logit regression on control variables

Outcome	Participation in treatment		
Independent variables	**Coef.**	**Std. Error**	**P>z**
Firm size in 2006 (ln)	0.232***	(0.056)	0.000
Firm sales in 2006 (ln)	0.087**	(0.040)	0.029
Manager's experience in 2006	-0.002	(0.004)	0.638
Size of locality	-0.074	(0.110)	0.501
Age of enterprise in 2006	0.000	(0.002)	0.954
Exporter status in 2006	0.163	(0.121)	0.178
Holding an IRQC in 2006	1.447***	(0.126)	0.000
Sector effects	Yes		
Country effects	Yes		
Constant	-1.801***	(0.230)	
Observations	2,577		
Pseudo R2	0.1530		
LR chi2(21)	541		

Note: Robust standard errors are in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

Table 6: Propensity score matching robustness checks.

	Average Treatment effect on Treated (ATT)			
Outcome	**Nearest Neighbour (1)**	**Nearest Neighbour (2)**	**Kernel (epanechnikov)**	**Bandwidth (epanechnikov)**
Gaining an IRQC (t-stat)	0.24±0.02 (11.3)	0.23±0.02 (11.4)	0.25±0.02 (13.6)	0.4
Retaining an IRQC (t-stat)	0.40±0.07 (6.1)	0.41±0.06 (6.8)	0.39±0.05 (7.3)	0.06
Becoming an exporter (t-stat)	0.024±0.021 (1.1)	0.026±0.0.018 (1.45)	0.036±0.015 (2.3)	0.025
Remaining an exporter (t-stat)	0.060±0.064 (0.9)	0.071±0.057 (1.25)	0.079±0.054 (1.5)	0.025