

# Economic and social impact of grape growing in Northeastern Brazil

Linda Arata<sup>1</sup>, Sofia Hauschild<sup>2</sup> and Paolo Sckokai<sup>1</sup>

<sup>1</sup> Department of Agricultural and Food Economics, Università Cattolica del Sacro Cuore, Piacenza, Italy

<sup>2</sup> Italian Embassy of Brazil, Brasilia, Brazil

*corresponding author:* <sup>1</sup> Linda Arata, email: [linda.arata@unicatt.it](mailto:linda.arata@unicatt.it)

Submission date: 8<sup>th</sup> June 2017

## Abstract

In recent decades grape production has been developing in some municipalities in Northeastern Brazil as a well-organised, competitive and high value industry. Although so far a few municipalities are involved, the Northeastern viticultural industry has become a model for the whole Brazil and has been identified as a production district. Given the importance of agriculture in the economy of the region our study aims at analysing whether the grape producing activity affects some socio-economic indicators, namely the Theil index, the Human Development Index (HDI) and the unemployment rate over the period 2000-2010. The study is focused on the Northeastern states of Bahia and Pernambuco, two of the poorest and with the highest income inequality among Brazilian States and combines the Difference-in-Differences with the Propensity Score matching method at the municipality level. Results seem to indicate that grape growing plays an important role to guarantee a fairer income distribution. Indeed, the municipalities that grow grape experience a decrease in the level of Theil index by 11.7% compared to the level they would have if they had not participate in grape production. No effect has been found on the HDI and on the unemployment rate. Results are robust to the potential presence of an hidden bias according to the Rosenbaum sensitivity analysis.

Keywords: grape production, socio-economic indicators, Brazil, Propensity Score Matching

JEL Codes: O13, Q13, C21

## **1. Introduction and background**

Although historically concentrated in the Southern states of Brazil, since the '60s grape production has developed also in the Northeastern region of the country, due to public-private investments in agriculture as well as to the development towards a commercial agriculture. One of the most important investment subsidized by the government consisted in irrigation systems which have allowed the setting up of grape growing. Indeed, the Northeast of Brazil is classified as a semi-arid region, with little and unpredictable amount of rain, which could undermine any potentiality for grape production if adequate irrigation systems were not in place. Total area under irrigation in the country expanded more than fivefold between 1960 and 1980 and in two of the Northeastern states, Bahia and Pernambuco, it increased by 100,000 hectares (Selwyn, 2008). The construction of new public infrastructures in Northeastern Brazil, such as roads and airports, has facilitated the development of the grape industry and trade. The investments in the two states and the introduction of rational agricultural practices led to a radical transformation of agriculture and social relationships. Since the '70s-80s there has been a shift from small-scale riverside and flood plain agriculture, where the sharecropping system between landowners and live-in workers was in place, to commercial agriculture, based on a high value horticultural industry.

Grape production represents one of the products of the new regional agricultural system and it is mainly concentrated in the Vale do São Francisco (San Francisco Valley, SF, the region located around the San Francisco River), which includes parts of Bahia and Pernambuco. Viticulture in the semi-arid Northeastern macro-region has specific features due to the climate conditions characterised by a monthly average temperature between 24°C and 30 °C, 500 millimetres/year of precipitation and 50% air humidity. The permanent warm weather is responsible for an acceleration of the physiological process and the propagation is very fast, allowing the first harvest after one year and a half. In addition, irrigated grapes can be produced continuously throughout the year allowing on average 2.5 production cycles per year (Texeira et al., 2007). This pattern leads to an average production of 40 tons per hectare per year, well above the ones obtained in other Brazilian

grape producing regions and in other parts of the world. It also allows harvesting in periods where prices are higher, which turns viticulture into an activity with a lower degree of uncertainty and a high potential profitability (Lima et al, 2009). In 2003 the Integrated Production (IP) method was introduced for grape production in the SF valley. This introduction has improved grape production systems not only for the IP product but for all the grape produced in the region (Camargo et al., 2011).

Together with the subsidisation of irrigation systems and of other infrastructures (roads, airports), the government has supported new grape plantations by subsidised credit and by tax breaks (Tales, 2009). There have also been investments in the training of workers in the grape industry, in research to improve the grape quality, as well as in the promotion of events such as organised paths along the vineyards, festivals and competitions.

Nowadays the SF valley is responsible for 99% of table grape exported by Brazil (Lima et al., 2009) and it is gaining fame as a development model in the Northeast of Brazil, the poorest macro-region of the country (IFAD, 2011) and one of the regions with the highest income inequality (IPEA, 2015). Between 1991 and 2001 grape export from the SF valley increased in volume from 1,000 to 13,000 tons and in value from 4.7 to 20.4 US\$ millions (Selwyn, 2008). Although grape production in this region mainly consists of table grape, recently there has been a growth in the production of grape to be transformed into fine wines such as Cabernet Sauvignon, Syrah, Moscato Canelli, Chardonnay and Chenin Blanc.

In 1992, parallel to an expansion in grape production and trade, the Brazilian Grape Marketing Board (BGMB) was created with the task of exporting grapes mainly to the EU and over the years its importance has increased. The board performs also a quality check and provides its associates (individual farms, cooperatives and producer associations) with training and it is the main exporting organisation of the Northeastern region (Selwyn, 2008). In 2010 the table grape product in the SF valley has been rewarded with the Geographical Indication (GI) and it represents the first GI product in the Northeastern region.

Differently from other Brazilian agricultural industries, dominated by large and extra-large farm size, such as sugar cane, corn and soybeans, small farms organised in cooperatives play an important role in the grape industry, especially in Northeastern Brazil (Tales, 2009). The presence of small farms should reduce the exploitation of workers that characterises other agricultural industries in Brazil and represents a social benefit at the local level. In addition, the grape industry in this area represents an example of 'production district', given the high concentration in the same area of all actors involved in the grape supply chain and the high level of cooperation (Tales, 2009).

The development of a well organised, competitive and high-value agricultural sector such as the grape producing industry in Northeastern Brazil may have an important role in the socio-economic development of the area. This is further supported by the high share (around 20%) of Agricultural Gross Domestic Product (GDP) over total GDP. Although some studies investigate the effect of sugar cane and soybean cultivation on some development indicators in Brazil (Chagas et al., 2012 and Weinhold et al., 2013), we are not aware of studies that consider the effects of grape production. Our paper aims at filling this gap by investigating whether the Brazilian Northeastern municipalities that started to grow grape after 2000 has experienced an improvement in some socio-economic indicators over the period 2000-2010 compared to the municipalities that have never grown grape in that period. Given the specific features of the grape industry in this area, it is likely to have an impact on regional socio-economic development.

The paper is focused on the two Northeastern states of Bahia and Pernambuco. As stated earlier, these two states are among the poorest and also among the ones with the highest income inequality in Brazil. Hence, it is interesting to investigate the impact of grape growing on their

social and economic conditions. In particular, the effect of grape growing on the Theil Index<sup>1</sup>, on the Human Development Index (HDI)<sup>2</sup>, and on the unemployment rate is explored.

## 2. Methodology

### 2.1 Propensity Score matching and Difference-in-Difference

We investigate the effect of grape growing on the socio-economic conditions of the municipalities of Bahia and Pernambuco by means of the Propensity Score Matching (PSM) methodology. PSM is a semi-parametric method that allows to assess the effect of a treatment in a non-experimental setting. Indeed, in non-experimental conditions the treatment is randomly assigned and individuals self-select to the treatment according to their characteristics. If those characteristics are related to the outcome to be evaluated, the simple comparison between treated and non treated individuals leads to a bias evaluation. PSM aims at overcoming the selection bias problem by matching each treated unit with one or more non treated units with similar observed characteristics, such that the difference in the outcome between the units can be interpreted as the effect of treatment (Smith and Todd, 2005).

Rosenbaum and Rubin (1983) propose to combine the observed characteristics ( $X$ ) that potentially affect both the treatment and the outcomes in one summary measure, the propensity score  $P(X)$ , that is the probability of being treated, such that the conditional distribution of ( $X$ ) given  $P(X)$  is independent of the treatment assignment. PSM provides a consistent evaluation of the treatment when two assumptions are satisfied. The first assumption is the mean independence

---

<sup>1</sup> The Theil Index is a measure of income inequality and was developed by Theil in 1967. A value of 0 reflects total equality, a value of 1 represents maximum inequality.

<sup>2</sup> This indicator encompasses three dimensions of social conditions: education (measured by rates of literacy and enrollment), longevity (life expectancy at birth), and income (per capita gross domestic product - GDP) (Chagas et al, 2012). According to the United Nations Development Programme (UNDP), the HDI was created to emphasize that not only economic growth but also people and their capabilities should be the ultimate criteria for assessing the development of a determined region.

assumption which states that after conditioning on the propensity score, the mean outcome is independent of the treatment assignment. The second assumption is the common support condition which guarantees that each treated unit potentially finds a matched untreated unit by restricting the probability of the treated to be lower than 1. The presence of unobservables that simultaneously affect the outcomes and the decision to participate into treatment lead to biased results by violating the first assumption. To partially overcome the problem of selection bias on unobservable, Heckamn, Ichimura and Todd (1997) propose to combine the PSM estimator with the Difference-in-Differences (DID) estimator such that the effect of the treatment is evaluated by comparing the before-after outcome of the treated units with that of the matched non treated units and the matching is based on the propensity score:

$$DID = \frac{1}{N} \sum_{i=1}^{I \cap S} \left[ (Y_{it}^1 - Y_{it}^0 | D = 1) - \sum_{j=1}^{J \cap S} W_{ij} (Y_{jt}^0 - Y_{jt}^0 | D = 0) \right]$$

where  $t'$  is the pre-treatment period,  $t$  is the post-treatment period,  $i$  identifies the treated individuals,  $j$  identifies the non treated individuals,  $N$  is the number of units of the treated group falling in the region of common support,  $W_{ij}$  indicates the weights ( $0 \leq W_{ij} \leq 1$ ), which depend on the distance between  $P_i$  and  $P_j$ , and  $S$  indicates the region of common support.

The DID estimator controls for unobservables that are constant over time and are responsible for outcome level differences. Although DID allows for time-invariant differences in outcome levels between the treated and the control group, it requires that conditional on the propensity score the outcome in the two group follows parallel path in the absence of the treatment (i.e. the DID mean independence).

The aim of the PSM is to estimate the average treatment effect on the treated (ATT) which, in the case of the DID-PSM, can be expressed as

$$ATT = \{E(Y_t^1 - Y_t^0 | D = 1, P(X)) - E(Y_t^0 - Y_t^0 | D = 0, P(X))\}$$

and represents the difference in the average outcome growth between the treated and the matched control group.

The combination of PSM and DID has been used to investigate the impact of some agricultural practices on farm production choices and economic performances in developed countries (Arata and Sckokai, 2016; Udagawa et al., 2014; Pufahl and Weiss, 2009;). In developing countries, where agriculture represents a large share of total GDP, PSM-DID has been applied to analyse the impact of agricultural activities on social and economic development (Chagas et al., 2012 and Weinhold et al., 2013). Our analysis belongs to this second stream of literature, since the combination of PSM and DID is suitable for our aim of analysing the effect of growing grape on some socio-economic indicators at the municipality level. Indeed, it is likely that municipalities that started to grow grape differs from the municipalities that never grew grape over the period considered and this difference may be related to the values of the outcomes. In addition, the use of DID allows to use the pre-treatment outcome as a control variable in the propensity score and to remove the bias for the time-invariant unobserved characteristics. This later feature allows to reduce the set of control variables to be used in the propensity score.

## *2.2 Sensitivity analysis*

The combination of PSM and DID preserves from the bias due to time-invariant unobserved characteristics. However, the presence of unobservables that vary over time and that are related to the decision to grow grape and to the outcome indicators undermines the results of the treatment effect. Matching estimators are not robust to this kind of hidden bias. As it is not possible to check the presence and the magnitude of the hidden bias, a sensitivity analysis is required in order to measure how strongly an unobserved variable must affect the odds ratio of treatment assignment to undermine the conclusions about the treatment effect. Rosenbaum (2002) propose to get a bound on the significance level of the treatment effect according to the extent of the hidden bias. We can express the log of the odds as:

$$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = f(X_i) + \gamma u_i$$

where,  $\pi_i$  is the probability of unit  $i$  to participate into the treatment,  $f(X_i)$  is a general function that relates the observed covariates to the odds of participation,  $u_i$  is an unobservable component and  $\gamma$  is its effect on the odds of participation. The odds ratio between two observationally identical units is:

$$\Gamma = \frac{\frac{\pi_i}{1-\pi_i}}{\frac{\pi_j}{1-\pi_j}} = \frac{e^{f(X_i)+\gamma u_i}}{e^{f(X_j)+\gamma u_j}}$$

The last term of the equality reduces to  $e^{\gamma(u_i-u_j)}$  as unit  $i$  and unit  $j$  are identical with respect to the observed variables. Therefore, if there are no differences in the unobserved variables, or if the unobservables do not affect the treatment assignment ( $\gamma=0$ ), the odds ratio is equal to 1. Conversely if there is hidden bias due to unobservables the odds ratio conditional to the observable characteristics departs from 1. Thus,  $\Gamma$  measures the magnitude of the hidden bias. The larger the value of  $\Gamma$ , the stronger is the influence of an unobserved variable on the decision to participate and the wider is the confidence interval around the treatment effect. For each level of  $\Gamma$  the bounds for the significance level of each outcome is computed.

The Rosenbaum sensitivity analysis represents one of the most widely applied method to check the robustness of the results in the PSM (Caliendo and Kopening, 2008; Chagas et al., 2008, Liu and Lynch, 2011). We implement this analysis in our study in order to check the robustness of our results to the potential presence of an unobserved variable that simultaneously affect the decision to grow grape and the indicator outcomes.

### 3. Data and empirical model



The state of Bahia is 564.733 km<sup>2</sup>, has 15 million inhabitants and ranks the fourth most populous Brazilian state and the fifth-largest in size, with a total of 417 municipalities (Brazilian Institute of Geography and Statistics - IBGE). In 2010 (the year of our evaluation), Bahia's HDI ranked 22<sup>nd</sup> among the 27 Brazilian states, and the 3<sup>rd</sup> most unequal Brazilian state according to the Theil Index (IBGE). Pernambuco has a total area of 98.149 km<sup>2</sup>, 185 municipalities and 9.2 million people; it is the seventh most populous state of Brazil and the sixth most densely populated. Pernambuco ranks 19<sup>th</sup> among the 27 Brazilian states for the HDI and is the 12<sup>th</sup> most unequal considering the Theil Index (IBGE).

As the number of total municipalities and their boundaries changes periodically, municipalities have been consolidated into Minimum Comparable Areas (MCAs) to make them consistently comparable among years. In our analysis the treated group is represented by the 16 MCAs of Bahia and Pernambuco that did not grow grape in 2000 but started grape production between 2000 and 2010. Conversely, the non-treated group is represented by the 363 MCAs of the same region that never grew grape over the period 2000-2010. We exclude from the non-treated group those MCAs that are located inside the so called "sertão nordestino" since they did not have irrigation systems in the period under analysis. The rationale behind that decision was that, if the region does not have neither a minimal amount of rain nor a good irrigation system, it gets really difficult to succeed in producing grape, and this may have a more general effect on the socio-economic indicators we analyse. Also, the metropolitan areas of Recife and Salvador are excluded from the analysis.

The first step of the matching procedure consists in running a probit model of the probability (propensity score) of a MCA to produce grape on a set of control variables. The control variables used to compute the propensity score are selected in order to control for characteristics that may affect both the decision of starting grape production and the development outcomes. These control variables are: the share of agricultural employees over total population, the share of agricultural GDP over total GDP, the average number of tractors per farm, the value of agricultural production

per hectare, the average temperature in each season and the average rainfall in each season. These variables come from the agricultural census of 2006 carried out by the Brazilian Institute of Geography and Statistics (IBGE). Unfortunately, the time when the Brazilian agricultural census is carried out does not overlap with the general census, thus we do not have data of our control variables for the year 2000. As we apply the DID we also use the 2000 values of the outcomes (pre-treatment period) as control variables. Since we are interested in analysing the effect of grape growing on the socio-economic conditions of MCAs, the outcome variables are the Theil Index, the HDI, and the unemployment rate. These outcomes are measured at the MCAs level both in 2000 and in 2010, the last two years of the Brazilian census carried out by IBGE.

Based on the propensity score from the first step, we implement the 10 nearest neighbour (10 NN) matching estimator with replacement which assigns to each MCAs starting grape production later than 2000 (treated group) ten MCAs that had never grown grape in 2000-2010 (non-treated group). For each treatment unit, the ten closest matched non-treated municipalities in terms of propensity score are selected. In order to reduce the bias that may derive from an estimator that assigns multiple non-treated units to each treated unit, we introduce a caliper of 0.1; thus, among the ten nearest neighbours matched units, only the ones whose propensity score differs from the propensity score of the treated group no more than 0.1 are selected. The 10 NN matching estimator is a good compromise between bias and variance. Indeed, assigning to each treated unit multiple non treated units reduces the variance of the estimator compared to the single nearest neighbour at a small cost in terms of bias (Lawley and Towe, 2014). In addition, the bias is controlled by the imposition of a caliper and by allowing for replacement.

Before matching, the share of agricultural employees on total population is 30.6% in the treated group and 24.3% in the non treated group and the share of agricultural GDP on total GDP is 20.4% and 18.5% in the two groups respectively (Table 1). While these two variables do not show differences that are statistically significant, the average number of tractors per farm significantly

**Table 1. Control variables mean and standardized bias before and after the matching**

control variables	Mean		t-test on the mean difference		% bias	% bias reduction
	treated	control	t	p value		
share of workers in the agricultural sector over total population ^	30.65	24.32	1.33	0.186	29.10	
	31.41	29.90	0.19	0.853	6.90	76.10
share of agricultural GDP over total GDP	20.38	18.51	0.60	0.546	15.10	
	20.79	23.67	-0.50	0.621	-23.40	-54.60
average number of tractors per farm	0.21	0.05	3.51	0***	40.20	
	0.08	0.09	-0.10	0.923	-1.10	97.20
agricultural GDP per hectare (R\$/ha)		1906.3				
	2103.30	0	0.25	0.807	7.60	
	2332.0	2278.5				
Theil index	0	0	0.04	0.966	2.10	72.80
	0.56	0.50	1.84	0.067*	48.40	
HDI	0.54	0.53	0.23	0.817	9.30	80.80
	0.45	0.43	1.52	0.129	33.10	
unemployment rate	0.45	0.44	0.33	0.741	13.70	58.60
	13.47	14.68	-0.67	0.503	-17.60	
summer temperature (°C)	14.44	14.26	0.08	0.938	2.60	85.00
	25.55	25.23	1.12	0.263	28.40	
autumn temperature (°C)	25.60	25.62	-0.06	0.953	-2.40	91.40
	24.78	24.39	1.45	0.148	39.90	
winter temperature (°C)	24.83	24.90	-0.17	0.869	-6.40	83.80
	22.98	21.90	3.67	0***	92.80	
spring temperature (°C)	22.89	23.06	-0.34	0.737	-14.70	84.20
	25.41	23.98	4.77	0***	117.30	
summer precipitation (mm/month)	25.37	25.48	0.84	-0.210	-9.00	92.30
	110.14	84.02	2.61	0.009***	55.30	
autumn precipitation (mm/month)	101.42	104.63	-0.16	0.877	-6.80	87.70
	79.74	111.77	-2.80	0.005***	-82.00	
winter precipitation (mm/month)	73.10	78.41	-0.61	0.547	-13.60	83.40
	17.28	92.91	-5.10	0***	-157.20	
spring precipitation (mm/month)	18.66	21.44	-0.21	0.833	-5.80	96.30
	49.97	62.08	-1.29	0.198	-33.20	
	46.34	47.20	-0.07	0.945	-2.40	92.90

^ The first row of each control variable refers to unmatched groups while the second row refers to matched group

\*, \*\*, \*\*\* indicate 10%, 5% and 1% significance level respectively

differs between the two groups: it is 0.21 in the treated MCAs and 0.05 in the non treated MCAs. If we look at the value of the outcome variable before the treatment (thus to the value employed as control variable in the matching) we notice that the Theil index in the treated MCAs is significantly larger than the Theil index in the non treated group (0.56 vs. 0.50). The HDI is 0.45 and 0.43 and

the unemployment rate is 13.5 and 14.7 in the two groups respectively, but these differences are not statistically significant.

In order to check the goodness of our matching technique (i.e. the ability to make the distribution of the control variables independent of the decision to participate into the treatment) we perform three tests as suggested by Caliendo and Kopeinig (2008). The first test is the covariates balancing property which consists in a t-test on the mean difference of each control variable between the treated and the non-treated group. The second criteria consists in measuring the standardized bias before and after the matching; the standardized bias measures the distance of the marginal distribution of the covariates between the two groups. It is calculated for each covariate as the difference of the sample mean between treated and matched control group over the square root of the average of the samples variances. The third criteria is the pseudo R square, which consists in re-estimating the probit model after the matching when the pseudo R-square should appear very small (Sianesi, 2004).

Once the matching has been performed and the matching quality is verified, the ATT is computed to get the difference in the average growth of the outcomes between the MCAs which has started grape production after 2000 and MCAs which had never produced grape in the period 2000-2010.

## **4. Results**

### *4.1 Quality of matching*

As discussed in Section 3, we compute three indicators to check the quality of our matching analysis: the balancing test, the standardized bias and the pseudo R square. Results on the balancing test (Table 1) show that after the matching there are not statistically significant differences in the level of control variables between the treated and the control group. That is not the case before the matching, since some of the variables differ between the two groups at the 1% (average number of

tractors per farm, winter and spring temperature, winter precipitation), 5% (summer precipitation) and 10% (Theil index in the pre-treatment period) significance level. The percentage reduction in the standardized bias between the two groups ranges from 54.6 to 97.2% according to the variable. Finally, the F-test of all the control variables together against the probability of participation into the treatment is significantly different from zero before the matching (pseudo R<sup>2</sup> equal to 0.405 and p-value of the likelihood ratio lower than 0.01) while it is not after the matching (pseudo R<sup>2</sup> equal to 0.039 and p-value equal to 1). Thus, all the three indicators conclude that our matching analysis successfully reaches the goal of removing the differences in observed variables between the two groups such that, conditional on  $P(X_i)$ , the distribution of each covariate is independent of the treatment status.

#### *4.2 Impact of Grape Production*

The probit model of the probability of growing grape against the set of observed covariates indicates that the average number of tractors per farm, the temperature in autumn as well as the summer precipitation increase the probability of producing grape (significance level of 10%). On the other hand, the summer and winter temperature and the average precipitation level in spring decrease this probability. The other variables included in the binary model do not affect statistically significantly the decision of producing grape (Table 2).

MCAs which started to grow grape after 2000 experience a decrease in the value of the Theil index over the period 2000-2010, while the matched MCAs which had never grown grape in the same period showed an increase in the same indicator (Table 3). The treated group has an average decrease in the Theil index of 3.5% compared to the level in 2000 while the control group records a rise of 9.4%. The difference in the average growth between the two groups is significant at the 5% level and seems to indicate that grape production in the two Brazilian states of Bahia and Pernambuco contributes to a better income distribution and to reduce inequality. One of the reasons of this result may be the high quality level of grape production as compared to other agricultural

industries, which may generate a better remuneration of workers in the grape industry. In addition, parallel to the development of large farms in the grape industry in Northeastern Brazil, also a large number of small farms started grape production (Selwyn, 2008). The presence of a large number of small farms may guarantee a fairer income distribution. This result is opposite to what Weinhold et al. (2012) found for the effect of soybean production on inequality in the Brazilian Amazon. This may be explained by the different characteristics of the grape industry as compared to soybean, where the high share of large farms does not guarantee a fair income distribution.

**Table 2. Coefficient estimates of the probit model on the probability of growing grape**

	coefficient estimates	standard error	p-value
share of workers in the agricultural sector over total population	1.16	0.97	0.233
share of agricultural GDP over total GDP	-2.16	1.71	0.205
average number of tractors per farm	1.39	0.75	0.063*
agricultural GDP per hectare (R\$/ha)	0.00	0.00	0.136
Theil index	0.86	1.61	0.595
HDI	0.03	0.03	0.399
unemployment rate	-1.53	3.62	0.673
summer temperature (°C)	-1.42	0.74	0.056*
autumn temperature (°C)	2.22	1.11	0.045*
winter temperature (°C)	-1.00	0.60	0.094*
spring temperature (°C)	0.30	0.45	0.502
summer precipitation (mm/month)	0.04	0.03	0.08*
autumn precipitation (mm/month)	0.00	0.02	0.806
winter precipitation (mm/month)	-0.01	0.02	0.664
spring precipitation (mm/month)	-0.07	0.04	0.039*
constant	-6.22	5.54	0.262
pseudo R <sup>2</sup>	0.40		
total number of MCAs	378		
number of treated MCAs	16		
number of non treated MCAs	362		

We did not find any effect of grape production neither on the HDI nor on the unemployment rate in this area. Over the period 2000-2010, the HDI in the MCAs participating in grape production and the matched non grape producing MCAs increases with a parallel path and the unemployment rate decreases by 5 points in both groups. Given the absence of a reduction in the unemployment rate due to grape production, the improving of the Theil index may be the consequence of a shift of the labour force from a sector where the employee/landowner wage ratio was very low to a sector where the income is more equally distributed and where small farms have better conditions to survive.

**Table 3. Average Treatment Effect on the Treated (ATT) of growing grape, 2000-2010**

	Average growth in the treated group	Average growth in the control group	ATT	t
Theil index	-0.019	0.050	-0.070** (-0.034)	-2.070
HDI	0.162	0.164	-0.002 (0.013)	-0.160
Unemployment rate	-5.132	-4.990	-0.142 (1.585)	-0.090

#### *4.3 Impact of Grape Sensitivity Analysis*

The results of the Rosenbaum sensitivity analysis indicate that the positive effect of grape production on the Theil index is robust when an unobserved variable affects the odds ratio of the treatment assignment by no more than 25%-30% (Table 4). The absence of an effect on the HDI because of the starting of grape production is questioned by a critical level of  $\Gamma$  between 2.2 and 2.3, while the critical level of  $\Gamma$  in the case of the effect on the unemployment rate is between 2.1 and 2.2. In the last two cases, it means that a hidden bias, that causes the odds ratio of the probability to participate to change by more than 2, may undermine the validity of the conclusion on the absence of an effect of grape production on the HDI and on the unemployment rate.

According to the sensitivity analysis, our results seem to be robust against the potential presence of an unobserved factor that affects simultaneously the probability to grow grape and the

outcomes. It is worth to remind that we also control for unobserved factors that are constant over time by means of the DID. In addition, as mentioned by DiPrete and Gangl (2004), the results of the sensitivity analysis are the worst case scenarios. For example, if the hidden bias affects the odds ratio more than 30% it does not mean that there is no effect of grape production on the Theil index, but it means that the confidence interval of the Theil index would become wider and include the value of zero.

**Table 4. Results of the Rosenbaum sensitivity analysis**

Gamma ( $\Gamma$ )	p-critical		Unemployment rate
	Theil Index	HDI	
1	0.021	0.413	0.365
1.05	0.025	0.382	0.335
1.1	0.030	0.353	0.308
1.15	0.035	0.326	0.282
1.2	0.040	0.301	0.259
1.25	0.046	0.278	0.237
1.3	0.052	0.256	0.217
1.35	0.059	0.236	0.199
1.4	0.065	0.218	0.182
1.45	0.072	0.201	0.167
1.5	0.079	0.185	0.153
1.55	0.087	0.170	0.140
1.6	0.094	0.157	0.128
1.65	0.102	0.144	0.117
1.7	0.110	0.133	0.107
1.75	0.118	0.122	0.098
1.8	0.126	0.113	0.090
1.85	0.134	0.104	0.082
1.9	0.142	0.096	0.075
1.95	0.151	0.088	0.069
2	0.159	0.081	0.063
2.05	0.167	0.075	0.057
2.1	0.176	0.069	0.053
2.15	0.184	0.063	0.048
2.2	0.193	0.058	0.044
2.25	0.201	0.054	0.040



## 5. Discussion and Conclusions

In recent decades, grape production has become a well-organised, competitive and high quality agricultural industry in Northeastern Brazil. Given the importance of agriculture in the overall economy of the two Northeastern states of Bahia and Pernambuco (around 20% of GDP), it is likely that the development of a modern agricultural industry may have an impact on some socio-economic indicators. In addition, differently from other agricultural industries in Brazil, small family farms play a key role in the grape industry and the high concentration and cooperation among the actors of the industry in Northeastern Brazil identifies a production district strictly related with regional development paths. Our study investigate whether grape production affects income distribution measured by the Theil index, the HDI and the unemployment rate in the two states at the MCA level. By combining the PSM with the DID we compare the development of the value of each outcome between a treated group (MCAs that started to grow grape after 2000) and a control group (matched MCAs that never grew grape in the 2000-2010 period). Results seem to indicate that grape production contributes to a fairer income distribution within the treated MCAs. Indeed, MCAs that started grape production experience a decrease in the Theil index of 11.7% compared to the level they would have experienced without developing grape production. One of the reasons for the positive effect on the Theil index may be the high value added of this agricultural industry which may contribute to reduce the worker exploitation and guarantee a better remuneration. Another reason may be represented by the large share of small family farms that work in the grape industry, which may contribute to a fairer remuneration. No effect has been shown for the HDI and the unemployment rate. In order to promote the HDI and reduce the unemployment rate public policies specifically targeted to them are required.

As stated in this paper, the setting up of grape production in Northeastern Brazil has been supported by private-public investments in infrastructures as well as by subsidised credit and tax breaks. Thus, the grape industry in this region is an example of how public support to agriculture

leads to general socio-economic benefits to the society as a whole, since agriculture represents an important share of the economy. Given this result, it would be interesting to analyse whether the positive effects of grape production on some general socio-economic indicators, such as the Theil index, is confirmed also for the Southern states, the historical grape producing area in Brazil.

## References

- Arata, L., and Sckokai, P. (2016). The Impact of Agri-environmental Schemes on Farm Performance in Five EU Member States: A DID-Matching Approach. *Land Economics* 92(1), 167-186.
- Caliendo, M., Kopeinig, S., 2008. Some Practical Guidance for the Implementation of the Propensity Score Matching. *Journal of Economic Surveys* 22 (1), 31-72.
- Camargo, U. A., Tonietto, J., Hoffmann, A. (2011). Progressos na viticultura brasileira. *Revista Brasileira de Fruticultura* 33: 144-149
- Chagas, A. L. S., Toneto, R., & Azzoni, C. R. (2012). A spatial propensity score matching evaluation of the social impacts of sugarcane growing on municipalities in Brazil. *International Regional Science Review*, 35(1), 48-69.
- DiPrete, T., and M. Gangl. 2004. Assessing Bias in the Estimation of Causal Effects: Rosenbaum Bounds on Matching Estimators and Instrumental Variables Estimation With Imperfect Instruments. *Sociological Methodology* 34: 271-310.
- Heckman, J.J., Ichimura, H., Todd, P. E., 1997. Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *Review of Economic Studies* 64, 605-654.
- IBGE - Brazilian Institute of Geography and Statistics. <http://www.ibge.gov.br/english/>. Assessed on July 2016.
- IFAD (2011): International Fund for Agricultural Development. *Habilitando os pobres rurais a superar a pobreza no Brasil*. Report available at [http://www.ifad.org/operations/projects/regions/PL/factsheet/brazil\\_p.pdf](http://www.ifad.org/operations/projects/regions/PL/factsheet/brazil_p.pdf). Assessed on July 2016.
- IPEADATA- Institute of Applied Economic Research. <http://www.ipeadata.gov.br>. Assessed on July 2015.
- Lawley, C., Towe, C., 2014. Capitalized Costs of Habitat Conservation Easements. *American Journal of Agricultural Economics*, 96(3), 657-672.
- Lima, M. A. C, Sa, I. B., Kiill, L. H. P., Araujo, J. L. P., Borges, R. M. E., Lima Neto, F. P., Soares, J. M., Leao, P. C. de S., Silva, P. C. G. da, Correia, R. C., Silva, A. de S., Sá, I. I. S., Silva, D. F. da (2009). *Subsídios técnicos para a indicação geográfica de procedência do Vale do Submédio São Francisco: uva de mesa e manga*. Embrapa Semiárido. Documentos. ISSN 1808-9992.
- Liu, X., Lynch, L., 2011. Do Agricultural Land Preservation Programs Reduce Farmland Loss? Evidence from Propensity Score Matching Estimator. *Land Economics* 87 (2), 183-201.
- Pufahl, A., Weiss., C. (2009).. “Evaluating the Effects of Farm Programmes: Results from Propensity Score Matching.” *European Review of Agricultural Economics* 36 (1): 79–101.
- Rosenbaum, P. 2002. *Observational Studies*. New York, NY: Springer.
- 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.

Selwyn, B. (2008). Institutions, upgrading and development: evidence from north east Brazilian export horticulture. *Competition & change*, 12(4), 377-396.

Smith, J., Todd, P. E., 2005. Does Matching Overcome LaLonde's Critique of Non-Experimental Estimators? *Journal of Econometrics* 125, 305-353.

Tales, V. (2009). Vitivinicultura no Nordeste do Brasil: Situação Recente e Perspectivas. *Revista Economica do Nordeste* 40 (3): 499-524.

Teixeira, A. D. C., Bastiaanssen, W. G. M., & Bassoi, L. H. (2007). Crop water parameters of irrigated wine and table grapes to support water productivity analysis in the Sao Francisco river basin, Brazil. *Agricultural Water Management*, 94(1), 31-42.

Udagawa, Chihiro, Ian Hodge, and Mark Reader (2014). "Farm Level Costs of Agri-environment Measures: The Impact of Entry Level Stewardship on Cereal Farm Incomes." *Journal of Agricultural Economics* 65 (1): 212-33.

Weinhold, D., Killick, E., & Reis, E. J. (2013). Soybeans, poverty and inequality in the Brazilian Amazon. *World Development*, 52, 132-143.