Regional Science and Urban Economics 59 (2016) 24–36

Contents lists available at ScienceDirect

Regional Science and Urban Economics

journal homepage: www.elsevier.com/locate/regec

A spatial difference-in-differences analysis of the impact of sugarcane production on respiratory diseases

André L.S. Chagas[a,*,1], Carlos R. Azzoni[a], Alexandre N. Almeida[b]

[a]Department of Economics, University of São Paulo, Av. Luciano Gualberto, 908, São Paulo, SP, Brazil
[b]Department of Applied Economics, Esalq-University of Sao Paulo, Avenida Pádua Dias, 11, Piracicaba, SP, Brazil

ARTICLE INFO

Article history:
Received 11 May 2015
Received in revised form 6 April 2016
Accepted 9 April 2016
Available online 16 April 2016

JEL classification:
C21
C23
I18
R11
Q18
Q42

Keywords:
Spatial diff-in-diff
Sugarcane producing
Health conditions
Spatial Econometrics

ABSTRACT

Sugarcane production represents around 10% of the agricultural area and 1% of GDP in Brazil, and has grown substantially in recent years. The traditional harvest method involves burning the field to facilitate access to the canes, resulting in well-documented negative effects on health. The existing studies do not consider the effects on health in the surrounding areas. This article presents a new variety of a spatial diff-in-diff model to control for the effects of sugarcane production in neighboring non-producing regions. This method is an addition to the Spatial Econometrics literature, as it includes spatial effects on treated and untreated regions, so that the effects on both producing and surrounding non-producing regions can be properly estimated. The results indicate that the effects on the producing regions are 78% larger than if the effects on the surrounding areas were ignored. Moreover, the effects on the surrounding areas, typically ignored in other studies, are relevant, and almost as large as the effects on the producing areas.

1. Introduction

Brazil is a traditional producer of sugar and has been an important player in the international market for centuries. In 2013, the country was the largest producer in the world, producing almost 27% more than the second largest producer, India. Although this market has somewhat stagnated in recent years, its growth was substantive in recent decades. Sugar is produced from sugarcane, an input that is also used to produce ethanol as fuel for automobiles. A governmental incentive program to substitute ethanol for fossil fuels was established in the late 1970s and reached full steam in the first decade of this century, as the automobile producers developed techniques to allow cars to run on both gasoline and/or ethanol. High oil prices powered the fuel substitution and the demand for ethanol increased dramatically, and production followed. As a result of these two influences, the production of sugarcane has increased sharply in the last 20 years, with the ethanol industry representing approximately 3.5% of Brazilian industrial GDP. The sector as a whole employs more than 6 million people and the planted area doubled in the last 20 years, occupying 10% of the agricultural area of the country.

The ethanol program has been considered a success in terms of emissions reduction by replacing pollutant fossil fuels (Goldemberg et al., 2008), but there are many issues related to the possible negative by-products of sugarcane production. There are doubts about the quality of the employment in the sugarcane fields, because the activity is hazardous and physically demanding. There are also questions on environmental aspects, such as soil contamination, atmospheric pollution generated by the burning of the fields, water consumption, and dislocation of other crops towards native forests (Noronha et al., 2006). Some studies have shown that the balance of costs and benefits is positive from the standpoint of the entire country (BNDES and CGEE, 2008), but not so evidently in the growing regions that disproportionately bear the negative impacts.

* Corresponding author.
E-mail addresses: achagas@usp.br (A. Chagas), cazzoni@usp.br (C. Azzoni), alex.almeida@usp.br (A. Almeida).

[1] The author was partially supported by CNPq, Proc.481027/2011-4.

http://dx.doi.org/10.1016/j.regsciurbeco.2016.04.002
0166-0462/

The most studied aspect is related to the labor market, and the negative impacts of manual harvesting are highlighted (Alves, 2006, 2007; Baccarin et al., 2008). Toneto-Jr and Liboni (2008) indicated that sugarcane generates more jobs than soybean, and only slightly less than corn. As it generates more value per hectare and more jobs as well, cane growing generates more income per area planted than other staple crops. Because transportation costs on the raw material are high, processing plants (sugar mills and/or ethanol distilleries) must be located close to the fields, increasing the sector's indirect effects on the producing region. Chagas et al. (2011) evaluated the impact of sugarcane on the local Human Development Index using spatial propensity score matching, controlling for the fact that sugarcane production in one specific region is not random. The results suggest that sugarcane growing is not relevant to determine local social conditions.

This paper deals with the impacts of sugarcane production on health conditions in the planting areas and their neighborhood. Because harvesting involves burning the fields, it releases fine and coarse particulate matter that affects the population in the vicinity. We explore a new ground in presenting a spatial difference-in-differences model (SDID) to control for the effect of sugarcane production on both producing (treated) and nonproducing (untreated) neighboring regions. This procedure for measuring the effects is more complete than the ones used in previous studies, such as Heckert and Mennis (2012) and Dubé et al. (2014). It brings a new way to look at both the true effects of sugarcane production on health and the measurement of spatial effects in general.

The article is organized in six sections, including this introduction. The next section deals with a review of the literature of the effects of sugarcane production on human health. Section 3 presents a review of the methodological questions present in the literature, the methodology proposed to identify the possible impacts of sugar-cane production on the respiratory health conditions in the producing regions, and the data used. Section 4 presents the results followed by robustness checks of the estimates, as presented in Section 5. The last section contains the final remarks of the analysis.

2. Sugarcane production, air pollution, and human health

Sugarcane is harvested by unskilled workers mostly manually. This traditional harvest method involves burning the planting area to facilitate access to the canes. There are concerns about the possible negative direct and indirect effects on health in the planting regions. The burning of the fields is intended to increase workers? productivity, as it eases access to the plants, saves on time otherwise spent in the separation of leaves, and reduces work hazards (dry leaves are harmful and there might be poisonous insects and snakes). It takes place at the beginning of harvest, which coincides with the dry season in the production areas. Many studies highlight the increase in both fine and coarse particulate matter, black carbon concentration, especially during burning hours (Lara et al., 2005), and the increase of the air concentration of substances as nitrite, sulfite, oxide of carbon, and others in the air (Allen et al., 2004). Considering smoke dispersion, the literature relates that short and long-term exposition to classical pollutants (matter, sulfite, nitrite, oxide carbon, etc.) can negatively affect the economy of a country by damaging the health status of the workers, specifically among the young and the elderly (Braga et al., 1999; Fischer et al., 2003; Gangadharan and Valenzuela, 2001; Goncalves et al., 2005; Roseiro, 2002; Sicard et al., 2010; Sun and Gu, 2008; Wen and Gu, 2012).

Sugarcane burning generates a massive quantity of particles and toxic gases that spread all over the region, reaching cities and becoming a potential threat to human health. According to Mazzoli-Rocha et al. (2008), pollution from sugarcane burning may be as harmful as pollution from traffic and manufacturing activities. There are many studies in this topic on the Brazilian case, mostly coming from the public health literature (Arbex et al., 2000, 2004, 2007, 2014; Cançado et al., 2006; Carneseca et al., 2012; Goto et al., 2011; Ribeiro, 2008; Santejo Silveira et al., 2013; Uriarte et al., 2009). The study of Nicolella and Belluzzo (2015) is an exception. They use a classical difference-in-differences approach to evaluate the impact of the reduction in the pre-harvest burning sugarcane on respiratory health. The results indicate that reducing the area where sugarcane is harvested after burning reduces the number of hospitalization cases. These are mostly case studies focusing on the effects of burning on respiratory health problems at the local level. They concentrate on the short-distance effects, failing to capture the consequences of burning events on other places (spillover effect), which is the focus of this work.

The literature on spillover effects of environmental events is increasing rapidly, but it is still limited. There are many papers testing the well-known Environmental Kuznets Curves (EKC), associating low levels of environmental problems both at low or high per capita income levels, and at high levels of environmental problems at intermediate income levels (Dinda, 2004; Grossman and Krueger, 1991, 1995). Spatial econometrics techniques were used to measure if per capita emissions in a country (county) were spatially dependent on the environmental characteristics of the neighboring countries (counties), as in Ciriaci and Palma (2010), Hao and Liu (2016), Maddison (2006, 2007), Rupasingha et al. (2004), Stern (2000), Su et al. (2009).

3. Methodology and data

Spatial econometrics techniques are becoming more popular in the study of environmental interactions, such as Hosseini and Kaneko (2013), at the institutional level, Renard and Xiong (2012) and Li et al. (2014), on industrial structure similarity, Pandit and Laband (2007), on imperiled species, Won Kim et al. (2003) and Chen and Ye (2015), on housing and gasoline prices, Li et al. (2014), on local economic development, and air quality and urbanization, Fang et al. (2015) on automobile and population density, and Chen and Ye (2015) on the levels of precipitation and the direction and speed of the wind. However, to the best of our knowledge, there are still only few studies measuring the effects of pollution of any source on health indicators considering the spatial correlation (Lagravinese et al., 2014; Wang et al., 2014, 2015, are exceptions).

3.1. The difference-in-differences model

The literature on impact evaluation sets to measure the impact, or the marginal effect, of a single binary regressor that equals one if the treatment occurs and zero otherwise (Ashenfelter and Card, 1985). The simplest case is one where outcomes are observed for two groups in two time periods. One of the groups receives a treatment in the second period, and the other group is not exposed to the treatment during either period. In the case where the same units within a group are observed in each time period, the average gain in the second (control) group is subtracted from the average gain in the first (treatment) group. This should remove any biases in second-period comparisons between the treatment and control groups that could be the result of permanent differences between those groups, as well as biases from comparisons over time in the treatment group that could be the result of common trends.

In the equations that follow, y_{it} is the variable of interest (hospitalizations due to respiratory diseases) and $\mathbf{x}_{it}$ is a vector of observable characteristics specific to region i in period t. We consider two situations for each region: before (b) and after (a) treatment. Additionally, we introduce a fixed effect φ_i and a drift term θ_t.

Then, the pre-treatment and post-treatment outcomes are given by, respectively,

$$y^b_{it,0} = \varphi_i + \theta_t + \mu(\mathbf{x}_i) + \varepsilon_{it}$$
$$y^b_{it,1} = y^b_{it,0} \tag{1}$$

where, $y^b_{it,0}$ is the dependent variable in the untreated region, before treatment, and $y^b_{it,1}$ is the dependent variable in the treated region. After treatment, we have another situation to the treated region, that is

$$y^a_{it,0} = y^b_{it,0}$$
$$y^a_{it,1} = y^a_{it,0} + \alpha. \tag{2}$$

The parameter α captures the direct effect of the treatment on the treated region. Defining D_{it} as a region-i specific indicator of treatment in time t, we can write

$$\begin{aligned} y_{it} &= (1 - D_{it})y_{it,0} + D_{it}y_{it,1}, \\ &= \varphi_i + \theta_t + \alpha D_{it} + \mu(\mathbf{x}_i) + \varepsilon_{it}. \end{aligned} \tag{3}$$

Using the "before" and "after" formulations, we obtain the Average Treatment Effect (ATE)

$$\begin{aligned} \text{ATE} &= E[y^a_{it,1} - y^b_{it,1}] - E[y^a_{it,0} - y^b_{it,0}] \\ &= \alpha. \end{aligned} \tag{4}$$

An important assumption for identification is that one, and only one, of the potential outcomes is indeed observable for every member of the population. This assumption, sometimes called the observation rule, follows from the so-called Stable Unit Treatment Value Assumption (SUTVA). In other words, it is required that the potential outcome in one unit should not be affected by the particular assignment of treatments to the other units (Cox, 1958; Rosenbaum, 2010). Importantly, it implies that the treatments are completely represented and, in particular, that there are no relevant interactions between the members of the population.

3.2. The spatial difference-in-differences model (SDID)

As usual in spatial studies, we need to take into account that regions are interrelated. This generates the possibility of propagation of the effects on both the regions where production takes place (treated region) and on the surrounding areas (untreated region), violating the SUTVA assumption. This makes causal inference more difficult.

In recent years, some studies added spatial effects to the difference-in-differences approach. Heckert and Mennis (2012) measured the impact of Philadelphia's innovative vacant land greening program on residential property values. The authors concluded that the program's impact decreases with the distance to the treated lots and then compared the observed changes in property values surrounding treated vacant lots with changes in lots that could have been treated but were not. The author was worried with spatial nonstationarity and so used geographically weighted regressions to compare spatial variations of the program's impact across districts. The author found a positive impact of the program on property values surrounding greened vacant lots in comparison to properties surrounding nongreened ones. Comparing different versions of the model, the study also explored spatial variations in the impacts of the program, offering insights into which kind of neighborhood might produce the greatest economic benefit from vacant land greening programs.

Dubé et al. (2014) evaluated the impact of public mass transit systems expansion on real-estate values in Montreal, Canada, taking into account possible spatial spillover effects. The authors claimed that the SDID method generated better estimates of the impact of the establishment of new train stations and allowed for richer interpretations of the marginal effects. They indicated that the consideration of spillover effects "is a major methodological gain as compared to the DID version" (Dubé et al., 2014, p.38). The authors put particular emphasis on the development of a suitable weight matrix to account for the spatial links between observations. Their experimental design takes the estimated reduction in car travel time due to the inauguration of new stations in Montreal as the treatment. The SAR model specification made it possible to account for possible spatial spillover effects in the price of houses determination process. Interestingly enough, the authors concluded that there was little gain in using the SDID estimator, in comparison to the usual diff-in-diff method in that specific case, given some particularities of the situation. In spite of such disappointing results, they argued in favor of the SDID estimator, as it allows for a simple t-test of the presence of spatial autocorrelation in the dependent variable, which is better than assuming that the problem does not exist.

Both Heckert and Mennis (2012) and Dubé et al. (2014) ignored the problems involved in the possible violation of the SUTVA. Delgado and Florax (2015) introduced this concern, considering the treatment effect in a difference-in-differences approach for spatial data with local spatial interaction. In their case, the potential outcome of observed units depends on their own treatment as well as on the treatment status of proximate neighbors. The authors emphasize the indirect effect of the treatment over the treated region due to neighborhood structure, but do not pay attention to the effect on the untreated region closed to treated one, as we discuss ahead in this paper. Unfortunately, the authors also do not offer an empirical applications, as we do in this paper.

In our work, this problem is explicitly taken into account. In the problem dealt with in this paper, the nature of the treatment is such that it affects both treated and untreated regions, thus violating the SUTVA assumption. Thus, we have to model the effect of the treatment also on untreated regions neighboring the treated ones. This is what the following model intends to do.

3.3. The model

In the model constituted by Eqs. (1) to (3), we incorporate in Eq. (2) the possibility of a propagation effect of the treatment in both regions, treated and untreated. We can thus identify two different impacts in the after-treatment situation: in the treated region and in the untreated region. The latter depends on the proximity of the regions. In the after-treatment situation, we have

$$y^a_{it,0} = y^b_{it,0} + \mathbf{w}'_i d_{it}\beta$$
$$y^a_{it,1} = y^a_{it,0} + \alpha \tag{5}$$

where $\mathbf{w}_i$ is an $n \times 1$ vector associating each region to all the other regions, and d_{it} is an $n \times 1$ vector of values $d_{it} = 1$ if the region is treated, and $d_{it} = 0$ otherwise. The parameter α captures the direct effect of the treatment on the treated region; β captures the indirect effect of the treatment on all regions, treated and untreated, conditioned on the neighbor treated, which is captured by $\mathbf{w_i}'d_{it}$. Based on the definition of D_{it}, as before we have $y_{it} = (1 - D_{it})y_{it,0} + D_{it}y_{it,1}$.

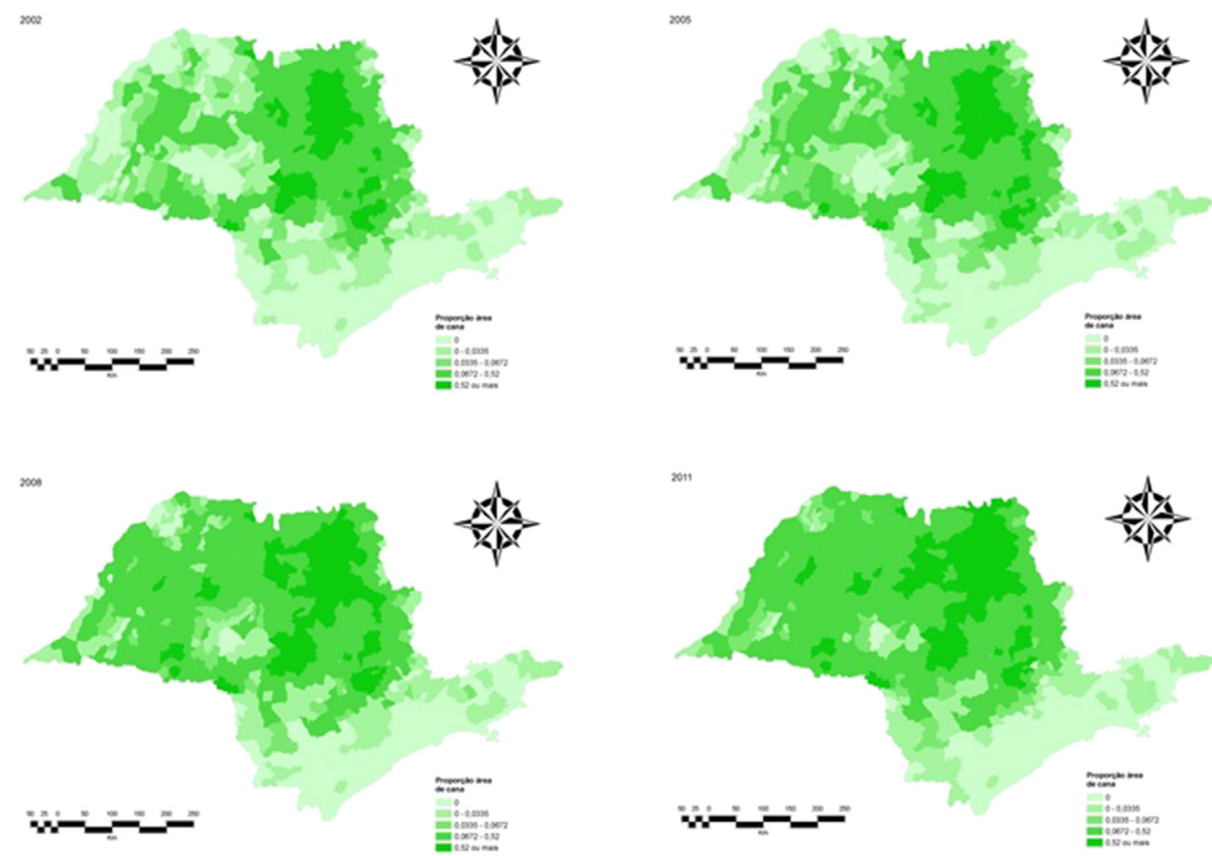

Fig. 1. Sugarcane production in São Paulo state by municipality, 2002, 2005, 2008, and 2011.
Source: IBGE, Municipal Agricultural Research.

Using the "before" and "after" definitions, now three effects can be computed: ATE (Average Treatment Effect), ATET (Average Treatment Effect on the Treated), and ATENT (Average Treatment Effect on the Non-Treated), as follows

$$\begin{aligned}
\text{ATE} &= E\left[y^a_{it,1} - y^b_{it,1}\right] - E[y^a_{it,0} - y^b_{it,0}] \\
&= \alpha \\
\text{ATET} &= E\left[y^a_{it,1} - y^b_{it,1}\right] \\
&= \alpha + \mathbf{w}_i' d_{it}\beta \\
\text{ATENT} &= E\left[y^a_{it,0} - y^b_{it,0}\right] \\
&= \mathbf{w}_i' d_{it}\beta.
\end{aligned}$$

Table 1
Number of treated regions.

Year	Number of municipalities	% of total
2002	230	0.357
2003	236	0.366
2004	242	0.376
2005	260	0.404
2006	291	0.452
2007	324	0.503
2008	348	0.540
2009	366	0.568
2010	387	0.601
2011	387	0.601
2012	385	0.598
2013	408	0.634

Source: IBGE, authors' calculations.

These expressions make clear the bias in the usual diff-in-diff approach if the spatial spillover in the treatment effect, represented by $\mathbf{w}_i' d_{it}\beta$, is not considered. In matrix notation, with a database structured as a panel data, we have

$$Y_t = \phi + \theta_t + \mu(\mathbf{X}_t) + (\alpha + \mathbf{W}\beta)D_t + \Xi_t \tag{6}$$

where $Y_t = (Y_{1t}, \cdots, Y_{nt})'$ is an $nt \times 1$ vector of observations, $\phi = (\phi_1, \cdots, \phi_n)'$ is an $nt \times 1$ vector of regional fixed-effects, $\theta_t = (\theta_1, \cdots, \theta_t)'$ is an $nt \times 1$ vector of time fixed-effects, $\mathbf{X}_t = (\mathbf{X}_{1t}, \cdots, \mathbf{X}_{nt})'$ is an $nt \times k$ matrix of covariates, $D_t = (D_1, \cdots, D_t)'$ is a *dummy* variable indicating treated regions, $\mathbf{W}$ is a $n \times n$ neighborhood weight matrix, and $\Xi_t = (U_{1t}, \cdots, U_{nt})'$ is a vector of errors of $nt \times 1$ dimension. α and β are parameters to be estimated, and μ is a function relating $\mathbf{X}_t$ to Y_t.

The term $\beta\mathbf{W}D_t$ indicates the indirect effect of the treatment on both regions, treated and untreated. This effect is usually ignored in estimations of this type.[2] However, this is an average effect, affecting both types of regions.

It is possible, however, that the incidence of the indirect effect could be different among treated and untreated regions. Consider a situation in which the indirect effect in the treated region is small, because the direct effect is more important. At the same time, the indirect effect on the untreated region is large, because it is the only effect impacting the region. In this situation, estimating β as an average to all regions might underestimate the real effect of the

[2] Angelucci and Giorgi (2009), Berniell et al. (2013), Kaboski and Townsend (2012) are some exceptions. However, these studies do not control for different structures of neighborhood.

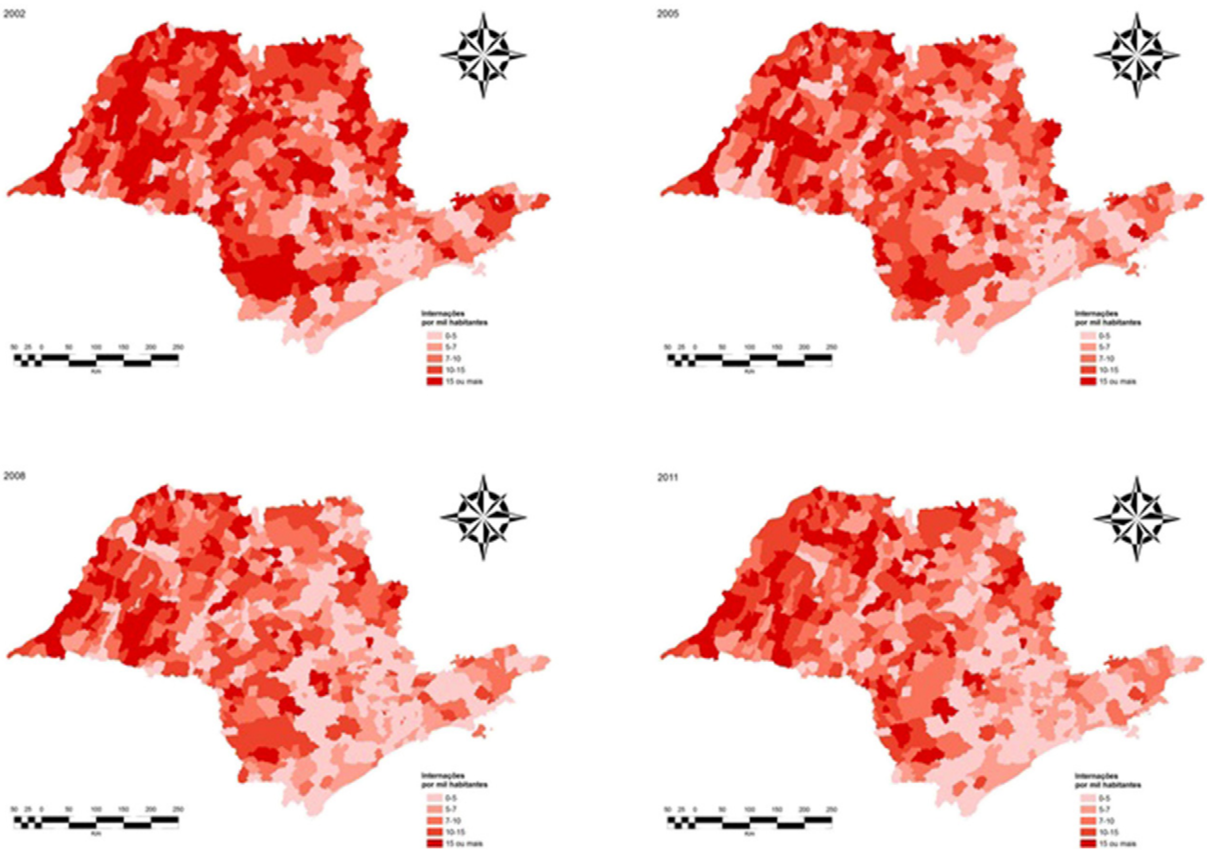

Fig. 2. Hospitalization due to respiratory problems in São Paulo state by municipality, 2002, 2005, 2008, and 2011.
Source: Datasus, Health Ministry.

treatment, because β will be estimated as an average of the indirect effects on the treated and untreated regions.

For clarity, consider, at each point in time, the following decomposition of the **W** matrix,

$$\mathbf{W} = \mathbf{W}_{T,T} + \mathbf{W}_{T,NT} + \mathbf{W}_{NT,T} + \mathbf{W}_{NT,NT}$$

where

$$\begin{aligned} \mathbf{W}_{T,T} &= \mathcal{D}_t \times \mathbf{W} \times \mathcal{D}_t \\ \mathbf{W}_{T,NT} &= \mathcal{D}_t \times \mathbf{W} \times \mathcal{D}_t^C \\ \mathbf{W}_{NT,T} &= \mathcal{D}_t^C \times \mathbf{W} \times \mathcal{D}_t, \text{ and} \\ \mathbf{W}_{NT,NT} &= \mathcal{D}_t^C \times \mathbf{W} \times \mathcal{D}_t^C \end{aligned}$$

where $\mathcal{D}_t = \mathrm{diag}(D_t)$ is an $n \times n$ matrix with D_t in the main diagonal and zeros elsewhere, and $\mathcal{D}_t^C = \mathrm{diag}(\boldsymbol{\iota}_n - D_t)$, with $\boldsymbol{\iota}_n$ a vector of 1's. In this way $\mathbf{W}_{ij}$ represents the neighborhood effects of the j-region on i-region, $i, j = T$ (treated) or NT (untreated). Substituting in Eq. (6), results in

$$Y_t = \phi + \theta_t + \mu(\mathbf{X}_t) + [\alpha + (\mathbf{W}_{T,T} + \mathbf{W}_{T,NT} + \mathbf{W}_{NT,T} + \mathbf{W}_{NT,NT})\beta]D_t + \Xi_t.$$

Then, it is clear that β represents an average effect, as we have mentioned above. A more realistic model considers different effects

Table 2
Hospitalization due respiratory health problem, by municipality.

Year	Hospitalization due to respiratory problems			
	Mean	Std. dev.	Maximum	Minimum
2002	11.678	7.216	45.366	0.810
2003	11.337	7.332	41.895	0.272
2004	10.742	6.831	45.326	0.000
2005	9.996	6.291	42.415	0.000
2006	10.632	6.860	54.111	0.262
2007	9.780	6.271	54.756	1.175
2008	8.745	5.690	37.122	0.949
2009	9.798	6.367	39.896	1.291
2010	9.355	6.204	46.392	1.166
2011	9.264	5.967	45.065	0.458
2012	4.956	3.227	22.942	0.681
2013	5.077	3.200	21.889	0.000

Source: IBGE, authors' calculations.

Table 3
Summary statistics for the variables.

Variable	Mean	Std. dev.	Max	Min	N
Treatment	0.500	0.500	1.000	0.000	7728
WD	0.507	0.371	1.000	0.000	7728
W_{11}D	0.399	0.427	1.000	0.000	7728
W_{21}D	0.108	0.203	1.000	0.000	7728
Workers	0.202	0.143	2.202	0.036	7728
Urbanization	0.831	0.146	1.000	0.221	7728
Olders	0.125	0.030	0.251	0.043	7728
Children	0.229	0.034	0.366	0.071	7728
Doctors	0.675	0.847	7.000	0.000	7728
Wworkers	0.202	0.056	0.456	0.089	7728
Wurbanizaton	0.832	0.076	0.997	0.472	7728
Wolders	0.125	0.021	0.198	0.057	7728
Wchildren	0.229	0.027	0.332	0.164	7728
Wmedicos	0.682	0.232	1.766	0.170	7728

Source: Authors' calculations.

Table 4
Linear correlations between the variables of the model.

Variables	Treatment	$\mathbf{W}D$	$\mathbf{W}_{11}D$	$\mathbf{W}_{21}D$	Workers	Urbanization	Olders
Treatment	1.0000						
$\mathbf{W}D$	0.7856	1.0000					
$\mathbf{W}_{11}D$	0.9342	0.8805	1.0000				
$\mathbf{W}_{21}D$	−0.5327	−0.0271	−0.4977	1.0000			
Workers	0.1475	0.1486	0.1649	−0.0759	1.0000		
Urbanization	0.2667	0.2703	0.2998	−0.1375	0.1930	1.0000	
Olders	0.0715	0.1970	0.0764	0.1994	−0.0807	−0.1718	1.0000
Children	−0.2502	−0.3684	−0.2730	−0.0983	−0.1386	−0.1266	−0.7296
Doctors	0.0117	0.0069	0.0136	−0.0160	0.2104	0.3675	0.0727
Wworkers	0.2644	0.3242	0.3338	−0.1106	0.2840	0.3242	−0.1230
Wurbanizaton	0.3973	0.4567	0.4786	−0.1737	0.2363	0.5389	−0.1499
Wolders	0.2094	0.3071	0.1910	0.1591	−0.0609	−0.1209	0.7088
Wchildren	−0.3491	−0.4665	−0.3729	−0.0673	−0.1160	−0.1445	−0.5804
Wmedicos	0.0356	0.0337	0.0948	−0.1383	0.1939	0.2446	−0.2536
Variables	**Children**	**Doctors**	**Wworkers**	**Wurbanizaton**	**Wolders**	**Wchildren**	**Wmedicos**
Children	1.0000						
Doctors	−0.2501	1.0000					
Wworkers	−0.2222	0.1412	1.0000				
Wurbanizaton	−0.2225	0.1241	0.5695	1.0000			
Wolders	−0.6651	−0.1044	−0.1608	−0.1875	1.0000		
Wchildren	0.8050	0.0014	−0.2773	−0.2622	−0.8269	1.0000	
Wmedicos	0.0019	0.0661	0.5626	0.5101	−0.3202	−0.0520	1.0000

Source: Authors' calculations.

for dissimilar $\mathbf{W}$ matrices. As, by construction, $\mathbf{W}_{T,NT}D_t$ and $\mathbf{W}_{NT,NT}D_t$ are $\mathbf{0}$-vectors[3], the unrestricted model is

$$Y_t = \phi + \theta_t + \mu(\mathbf{X}_t) + [\alpha + (\mathbf{W}_{T,T}\beta_1 + \mathbf{W}_{NT,T}\beta_2)]D_t + \Xi_t. \tag{7}$$

The models in Eqs. (6) and (7) are special forms of the SDID.[4] It is important to register that they do not contain a traditional spatial interaction effect, such as in the Spatial Autoregressive Model (SAR) and the Spatial Error Model (SEM) (Anselin, 1988; LeSage and Pace, 2009). However, we can model the control effects, $\mu(\mathbf{X})$, including an auto-regressive spatial term or the error as a Spatial Error Model or both

$$\mu(\mathbf{X}_t) = \rho\mathbf{W}Y_t + \mathbf{X}_t\gamma'$$

and/or

$$\Xi_t = \lambda\mathbf{W}\Xi_t + \Upsilon_t.$$

In the first equation, γ is a $1 \times k$ parameter vector to be estimated, and ρ is the spatial auto-regressive parameter. In the second equation, Υ_t is an error vector, not spatially associated, and λ is the spatial error parameter to be estimated. Thus, a complete version of models 6 and 7 is

$$\begin{aligned} Y_t =& (\mathbf{I}_n - \rho\mathbf{W})^{-1}\{\phi + \theta_t + \mathbf{X}_t\gamma' + [\alpha + (\mathbf{W}_{T,T}\beta_1 + \mathbf{W}_{NT,T}\beta_2)]D_t \\ &+ (\mathbf{I}_n - \lambda\mathbf{W})^{-1}\Upsilon_t\}. \end{aligned} \tag{8}$$

To the best of our knowledge, this formulation is new to the Spatial Econometrics literature.

3.4. Data

A balanced panel of 644 municipalities belonging to the state of Sao Paulo, the largest producer of sugarcane in the country was chosen as study area. Annual data covered the period 2002–2013. Information on sugarcane production, planted area, and harvested area is based on the annual survey on agricultural production developed by IBGE, the Brazilian statistics office.

As mentioned before, the expansion in the sugarcane growing area has prompted a series of questions on the possible conflicts between land used to produce food versus energy. This does not seem to be a problem at the national level: Brazil has over 800 million hectares of landmass, of which over 300 million are suitable for farming and ranching activities. Of these, about 60 million are used to grow permanent and temporary crops and some 200 million are used for animal husbandry. Thus, there is plenty of suitable land to increase production, and this can be even larger if degraded land is recovered and if productivity in animal production, which is very low in the country, would increase (Chagas et al., 2008). In São Paulo state, however, the crop represents nearly 50% of the area suitable for farming. Fig. 1 shows a map of the evolution of sugarcane production in São Paulo state, by municipality, during the period analyzed.[5] It is clear that there was a steady increase in the production in the northwest region of the state, and a sprawl to the west of the state, an area previously used for cattle ranching.

The production of sugarcane is important to define our treatment variable. In terms of physical conditions, any municipality in the state is capable of producing sugarcane, a plant with minimum soil and weather requirements. Even when it is not produced commercially, farmers typically plant a small quantity to feed animals or to make a highly appreciated drinkable juice. Although some restrictions on production in some areas were recently introduced by the state and federal governments[6] , they do not affect the capability of the restricted areas to produce the cane. In any case, they were implemented at the end of our period of analysis. From all the municipalities in the state, we consider as treated those in which the share

[3] The elements of both matrices $\mathbf{W}_{T,NT}$ and $\mathbf{W}_{NT,NT}$ are given, respectively, by $\mathbf{W}_{T,NT\,ij} = \begin{cases} \mathbf{W}_{ij} & \text{if } D_{it} = D^C_{jt} = 1 \\ 0 & \text{otherwise} \end{cases}$; and $\mathbf{W}_{T,NT\,ij} = \begin{cases} \mathbf{W}_{ij} & \text{if } D^C_{it} = D^C_{jt} = 1 \\ 0 & \text{otherwise} \end{cases}$. Then, $\mathbf{W}_{T,NT}D_t = \sum_j \mathbf{W}_{ij}|_{D_{it}=D^C_{jt}=1} \times D_{it}$, but this is a 0-vector because $\mathbf{W}_{T,NT\,ij}$ is different from zero only when D_{it} is equal to zero. The same holds for $\mathbf{W}_{NT,NT}D_t$.

[4] Delgado and Florax (2015) consider only the restricted case.

[5] We select some years in this period, but the evolution is evident.

[6] These restrictions aimed at forest and water sources preservation. At the state level, they were introduce only in 2008.

Table 5
Results.

	Classical	SLX		SDEM	
		Restricted	Unrestricted	Restricted	Unrestricted
Treatment	0.8145***	0.7184**	1.4914***	0.6828**	1.441***
	(0.2099)	(0.2835)	(0.3514)	(0.2867)	(0.379)
WD		0.2678		0.3561	
		(0.5189)		(0.6654)	
$W_{T,T}D$			−0.6671		−0.5902
			(0.5762)		(0.7357)
$W_{NT,T}D$			1.3445**		1.3503*
			(0.5939)		(0.7413)
R-square	0.7815	0.7830	0.7834	0.7828	0.7832
AIC	5.0426	5.0374	5.0359	5.0126	5.0116
Moran's I	0.0683	0.0608	0.0600		
p-Value	0.0000	0.0000	0.0000		
LM_lag	268.5993	228.7994	221.7469		
p-Value	0.0000	0.0000	0.0000		
LM_error	337.6452	268.1517	261.0415		
p-Value	0.0000	0.0000	0.0000		
rob LM_lag	0.0377	1.3812	1.5001		
p-Value	0.8460	0.2399	0.2207		
rob LM_error	69.0836	40.7336	40.7947		
p-Value	0.0000	0.0000	0.0000		
λ				0.3707***	0.3746***
				(0.0270)	(0.0269)
N	644	644	644	644	644
T	12	12	12	12	12

Note: All models include control variables. SLX and SDEM models also include spatial lags of controls variables. Additionally, the models include a constant and spatial and temporal fixed effects.
Statistical significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

of the area planted with sugarcane is above 6,7%, the median of the distribution of the production area. In Table 1 we report the number of treated areas in each year, showing an increase from 38.4% to 62.4% in the period.

Table 6
Complete results of the SDEM model (unrestricted case).

Dependent variable		Intern per th	
R-squared		0.7832	
corr-squared		0.0501	
Within R2		0.1867	
Between R2		0.0186	
Overall R2		0.0099	
σ^2		9.4838	
Log-likelihood		19,349.634	
AIC		5.0116	
Wald, p-value		271.2591, 0.0000	
Nobs,Nvar,#FE		7728, 15, 656	
# iterations		17	
Min and max λ		−0.9900, 0.9900	
Variable	Coefficient	Standard-error	p-Value
Constant	10.0166	8.0812	0.1851
Treatment	1.4410	0.3790	0.0003
$\mathbf{W}_{T,T}D_t$	−0.5902	0.7357	0.2892
$\mathbf{W}_{NT,T}D_t$	1.3503	0.7413	0.0759
trend_d	0.2891	0.4778	0.3322
trend_{wd}	−1.9698	0.8008	0.0194
Workers	−1.5337	0.5634	0.0098
Urbanization	6.8558	1.8923	0.0006
Olders	15.4285	9.0286	0.0926
Children	−8.6372	8.9771	0.2511
Doctors	0.6034	0.2364	0.0154
W-workers	−4.7291	3.0582	0.1207
W-urbanization	−9.2632	9.2298	0.2411
W-olders	−52.0746	35.3292	0.1346
W-children	40.4848	31.7538	0.1770
W-doctors	−1.0709	1.2023	0.2683
λ	0.3746	0.0269	0.0000

Our variable of interest is the number of persons hospitalized due to respiratory problems (per 1000 inhabitants). The data was provided by DATASUS[7] , the statistical agency of the Ministry of Health, and includes hospitalizations in public and private hospitals. The information is highly disaggregated in spatial terms, and we used data at the municipality level. As Fig. 2 indicates, the number of cases of hospitalizations due to respiratory diseases is decreasing over time. This could be associated to changes in the federal legislation, which introduced limitations for burning in certain areas and times. This is especially true in the state of Sao Paulo, in which a state law broadens the limitations imposed by the federal law. The practice of burning the canes to facilitate harvesting is expected to end in a few years' time in the state, both by restrictions coming from the legislation (both environment and labor market related) and by economic stimuli for the economical use of the leaves and the straws. Table 2 illustrates the situation.

Given that, we have introduced a trend variable interacting with the treated regions to adjust for this empirical evidence in all models estimated, and the coefficients are negative and significant in all cases.[8] We have included variables to control for socioeconomic

[7] http://www2.datasus.gov.br/DATASUS/index.php.
[8] As suggested by one of the anonymous reviewers, we have replaced the general trend with temporal fixed effects interacting with the treatment effects. The results are available upon request. The interaction dummies are not significant, in general, and they are not robust to different specifications. Although the presence of a large number of non-significant variables does not generate biases, it tends to compromise inference. However, the absence of an important explanatory variable, as the trend seems to be, can bias the results, if the missing variable is correlated with another explanatory variable. In fact, the effect of the untreated region becomes non-significant when the trend term is omitted. This might be due to a possible positive relationship between the trend and the way the region interacts with its neighbors over time. This is an interesting aspect to explore in a future work.

Table 7
Monte Carlo simulation results.

Models	β								
	0.1			0.4			0.8		
	Coef	95% conf. interval		Coef	95% conf. interval		Coef	95% conf. interval	
	N = 200, T = 10, λ = 0.1								
Contiguity matrix									
Classical	0.0852	−0.2210	0.3914	0.2020	−0.1042	0.5082	0.3578	0.0516	0.6640
SLX	−0.0173	−0.4069	0.3722	0.1022	−0.2874	0.4917	0.2615	−0.1281	0.6510
SDEM	−0.0168	−0.4063	0.3728	0.1024	−0.2871	0.4919	0.2613	−0.1283	0.6508
k-Nearest matrix									
Classical	0.1665	−0.3210	0.6540	0.4665	−0.0210	0.9540	0.8665	0.3790	1.3540
SLX	0.0982	−0.5184	0.7147	0.3982	−0.2184	1.0147	0.7982	0.1816	1.4147
SDEM	0.0987	−0.5171	0.7145	0.3987	−0.2171	1.0145	0.7987	0.1830	1.4145
k-Nearest Euclidean matrix									
Classical	0.1391	−0.3203	0.5985	0.3457	−0.1137	0.8051	0.6213	0.1619	1.0807
SLX	0.0009	−0.5636	0.5655	0.2106	−0.3540	0.7751	0.4901	−0.0744	1.0547
SDEM	0.0038	−0.5602	0.5678	0.2121	−0.3519	0.7761	0.4897	−0.0742	1.0536
	N = 200, T = 100, λ = 0.1								
Contiguity matrix									
Classical	−0.1812	−0.2746	−0.0878	−0.0728	−0.1662	0.0206	0.0718	−0.0216	0.1652
SLX	0.0143	−0.1044	0.1329	0.1229	0.0043	0.2416	0.2678	0.1492	0.3865
SDEM	0.0154	−0.1032	0.1340	0.1231	0.0045	0.2417	0.2667	0.1481	0.3853
k-Nearest matrix									
Classical	−0.4752	−0.6282	−0.3221	−0.1752	−0.3282	−0.0221	0.2248	0.0718	0.3779
SLX	0.0959	−0.0943	0.2862	0.3959	0.2057	0.5862	0.7959	0.6057	0.9862
SDEM	0.0957	−0.0941	0.2855	0.3957	0.2059	0.5856	0.7957	0.6059	0.9855
k-Nearest Euclidean matrix									
Classical	−0.3304	−0.4695	−0.1914	−0.1297	−0.2688	0.0093	0.1379	−0.0012	0.2769
SLX	0.0432	−0.1284	0.2148	0.2442	0.0726	0.4158	0.5122	0.3406	0.6838
SDEM	0.0429	−0.1284	0.2142	0.2430	0.0717	0.4143	0.5098	0.3384	0.6811
	N = 200, T = 10, λ = 0.5								
Contiguity matrix									
Classical	−0.1230	−0.4824	0.2364	−0.0200	−0.3794	0.3394	0.1174	−0.2420	0.4768
SLX	0.0810	−0.3584	0.5205	0.1858	−0.2536	0.6252	0.3254	−0.1140	0.7649
SDEM	0.0651	−0.3601	0.4904	0.1631	−0.2622	0.5884	0.2937	−0.1317	0.7191
k-Nearest matrix									
Classical	−0.3316	−0.9947	0.3314	−0.0316	−0.6947	0.6314	0.3684	−0.2947	1.0314
SLX	0.1061	−0.6523	0.8645	0.4061	−0.3523	1.1645	0.8061	0.0477	1.5645
SDEM	0.1040	−0.5832	0.7912	0.4040	−0.2833	1.0913	0.8039	0.1167	1.4911
k-Nearest Euclidean matrix									
Classical	−0.2050	−0.7674	0.3573	−0.0120	−0.5744	0.5503	0.2453	−0.3171	0.8076
SLX	−0.0227	−0.6925	0.6471	0.1714	−0.4984	0.8413	0.4303	−0.2395	1.1001
SDEM	0.0047	−0.6245	0.6338	0.1950	−0.4341	0.8241	0.4490	−0.1802	1.0782
	N = 200, T = 100, λ = 0.5								
Contiguity matrix									
Classical	0.0219	−0.0910	0.1349	0.1282	0.0152	0.2411	0.2698	0.1568	0.3827
SLX	0.0362	−0.0955	0.1678	0.1425	0.0108	0.2741	0.2842	0.1526	0.4158
SDEM	0.0357	−0.0892	0.1606	0.1385	0.0136	0.2634	0.2755	0.1506	0.4004
k-Nearest matrix									
Classical	0.0554	−0.1515	0.2623	0.3554	0.1485	0.5623	0.7554	0.5485	0.9623
SLX	0.0955	−0.1280	0.3190	0.3955	0.1720	0.6190	0.7955	0.5720	1.0190
SDEM	0.0976	−0.0964	0.2917	0.3977	0.2037	0.5916	0.7976	0.6036	0.9917
k-Nearest Euclidean matrix									
Classical	0.0375	−0.1364	0.2114	0.2411	0.0672	0.4151	0.5127	0.3388	0.6866
SLX	0.0272	−0.1710	0.2254	0.2316	0.0334	0.4297	0.5040	0.3059	0.7022
SDEM	0.0349	−0.1425	0.2123	0.2330	0.0556	0.4105	0.4972	0.3198	0.6747
	N = 200, T = 10, λ = 0.9								
Contiguity matrix									
Classical	−0.0707	−0.8090	0.6676	0.0461	−0.6922	0.7845	0.2020	−0.5363	0.9403
SLX	0.0288	−0.8134	0.8711	0.1453	−0.6970	0.9875	0.3005	−0.5417	1.1427
SDEM	0.0345	−0.4344	0.5033	0.1449	−0.3239	0.6136	0.2920	−0.1767	0.7608

(*continued on next page*)

Table 7 (*continued*)

Models	β								
	0.1			0.4			0.8		
	Coef	95% conf. interval		Coef	95% conf. interval		Coef	95% conf. interval	
	N = 200, T = 10, λ = 0.9								
k-Nearest matrix									
Classical	−0.1165	−1.7578	1.5247	0.1835	−1.4578	1.8247	0.5835	−1.0578	2.2247
SLX	0.0728	−1.4046	1.5502	0.3728	−1.1046	1.8502	0.7728	−0.7046	2.2502
SDEM	0.0862	−0.5891	0.7614	0.3862	−0.2890	1.0613	0.7861	0.1108	1.4614
k-Nearest Euclidean matrix									
Classical	0.0245	−1.3887	1.4377	0.2344	−1.1788	1.6476	0.5143	−0.8988	1.9275
SLX	0.1159	−1.1619	1.3936	0.3273	−0.9505	1.6051	0.6092	−0.6686	1.8869
SDEM	0.0452	−0.5761	0.6664	0.2494	−0.3721	0.8709	0.5214	−0.1001	1.1429
	N = 200, T = 100, λ = 0.9								
Contiguity matrix									
Classical	−0.1135	−0.3648	0.1378	−0.0067	−0.2579	0.2446	0.1358	−0.1155	0.3871
SLX	0.0372	−0.2387	0.3131	0.1455	−0.1304	0.4214	0.2899	0.0140	0.5658
SDEM	0.0333	−0.1123	0.1788	0.1327	−0.0128	0.2782	0.2654	0.1198	0.4109
k-Nearest matrix									
Classical	−0.3299	−0.9323	0.2725	−0.0299	−0.6323	0.5725	0.3701	−0.2323	0.9725
SLX	0.1038	−0.4462	0.6538	0.4038	−0.1462	0.9538	0.8038	0.2538	1.3538
SDEM	0.1017	−0.1042	0.3075	0.4017	0.1958	0.6075	0.8016	0.5958	1.0075
k-Nearest Euclidean matrix									
Classical	−0.1904	−0.6398	0.2590	−0.0123	−0.4617	0.4371	0.2253	−0.2241	0.6747
SLX	0.0812	−0.3397	0.5020	0.2618	−0.1590	0.6826	0.5026	0.0818	0.9235
SDEM	0.0583	−0.1345	0.2512	0.2290	0.0361	0.4219	0.4567	0.2637	0.6496

Source: Authors' calculations.

conditions that influence people's behavior towards health prevention, such as the proportion of workers in the population, and urbanization. We have also included the proportion of elderly and young people, to control for the presence of groups more susceptible to respiratory health problems, as indicated in the literature (Braga et al., 1999; Goncalves et al., 2005; Roseiro, 2002). Finally, we have introduced the proportion of doctors in the population, to control for the presence of regular assistance. Table 3 reports the descriptive statistics and Table 4 reports the correlation matrix.

4. Results

This section presents the results.[9] We computed and compared three different types of models. The first is a classical panel data regression with fixed effects, to set a baseline for comparing our results. The second includes spatial controls on the x variables. This is similar to the spatial lag of X (SLX) model case suggested by Vega and Elhorst (2015) and is our baseline case to perform the model search in Spatial Econometrics. The third involves traditional models in Spatial Econometrics (SDM/SDEM). For this last case, the selection was based on the Lagrange Multiplier (LM) and LM robust tests (Anselin et al., 1996), following the suggestion by Florax et al. (2003).[10] In all situations, the tests indicated the use of the SDEM model. Additionally, we consider both restricted and unrestricted cases, as models 6 and 7, respectively. We consider a k-nearest neighbors distance matrix, with k varying between 20 and 100, and use the Akaike Information Criterion (AIC) for pooled models, without spatial effects, to choose the size of k. Thus, we chose the matrix that best fits the data.

Classical panel

$$Y_t = \phi + \theta_t + (\alpha + \tau t)D_t + \mathbf{X}_t\gamma\prime + \Xi_t$$

SLX model

Restricted model

$$Y_t = \phi + \theta_t + (\alpha + \beta\mathbf{W} + \tau t + \psi\mathbf{W}t)D_t + \mathbf{X}_t\gamma\prime + \mathbf{W}\mathbf{X}_t\delta\prime + \Xi_t$$

Unrestricted model

$$Y_t = \phi + \theta_t + (\alpha + \beta_1\mathbf{W}_{T,T} + \beta_2\mathbf{W}_{NT,T} + \tau t + \psi\mathbf{W}t)D_t + \mathbf{X}_t\gamma\prime + \mathbf{W}\mathbf{X}_t\delta\prime + \Xi_t$$

SDEM model

Restricted model

$$Y_t = \phi + \theta_t + (\alpha + \beta\mathbf{W} + \tau t + \psi\mathbf{W}t)D_t + \mathbf{X}_t\gamma\prime + \mathbf{W}\mathbf{X}_t\delta\prime + (\mathbf{I}_n - \lambda\mathbf{W})^{-1}\Upsilon_t$$

Unrestricted model

$$Y_t = \phi + \theta_t + (\alpha + \beta_1\mathbf{W}_{T,T} + \beta_2\mathbf{W}_{NT,T} + \tau t + \psi\mathbf{W}t)D_t + \mathbf{X}_t\gamma\prime + \mathbf{W}\mathbf{X}_t\delta\prime + (\mathbf{I}_n - \lambda\mathbf{W})^{-1}\Upsilon_t$$

The results are reported in Table 5. The Classical Panel Model indicates that sugarcane production increases hospitalizations by only 0.81 cases per thousand, and this conclusion is significant at 1%, which is in line with the results of Nicolella and Belluzzo (2015). However, as we have mentioned above, this is an incomplete account of the effects. The introduction of spatial controls increases the influence of sugarcane production on hospitalizations, and suggests the relevance of introducing spatial controls on untreated regions. The impact of sugarcane production in the treated region is of 1.49 cases

[9] We based our estimates on Elhorst's routine for spatial panel data models (Elhorst, 2010a,b). Elhorst uses Maximum Likelihood (ML) because the number of studies considering Instrumental Variables/Generalized Method of Moments (IV/GMM) estimators of spatial panel data models is still relatively sparse. One exception is Kelejian et al. (2006), who used IV to estimate a spatial lag model with time-period fixed effects. They point out that the model cannot be combined with a spatial weight matrix with non-diagonal elements equal to 1/(N − 1). In this situation, the spatially lagged dependent is asymptotically proportional and thus collinear with the time-period fixed effects, as N goes to infinity. Elhorst (2010a) provides Matlab routines to estimate spatial panel data models, including the bias correction procedure proposed by Lee and Yu (2010) if the spatial panel data model contains spatial and/or time-period fixed effects, the direct and indirect effects estimates of the explanatory variables proposed by LeSage and Pace (2009), and a selection framework to determine which spatial panel data model best describes the data.

[10] These tests have become popular in empirical research. However, Elhorst (2014) calls attention to the fact that the power of these tests to spatial panel data models must still be investigated.

Table 8
Robustness check — neoplasm-related hospitalization.

Dependent variable		neopl per th	
R-squared		0.4166	
corr-squared		0.0284	
Within R2		0.0111	
Between R2		0.1042	
Overall R2		0.0081	
sigma2^		0.4844	
Log-likelihood		−7836.0529	
AIC		2.0318	
Wald test, p-value		174.3972, 0.0000	
Nobs,Nvar,#FE		7728, 15, 656	
# iterations		14	
Min and max λ		−0.9900, 0.9900	
Variable	Coefficient	Standard-error	p-Value
Constant	2.1565	1.6701	0.1733
Treatment	0.0780	0.0826	0.2553
$\mathbf{W}_{T,T}D_t$	0.1376	0.1420	0.2495
$\mathbf{W}_{NT,T}D_t$	0.1880	0.1454	0.1729
trend_d	−0.0517	0.1071	0.3550
trend_{wd}	−0.3723	0.1592	0.0259
Workers	0.0800	0.1254	0.3256
Urbanization	0.8412	0.4248	0.0562
Olders	11.8091	2.0527	0.0000
Children	−8.8971	2.0480	0.0000
Doctors	−0.0492	0.0529	0.2585
W-workers	−1.8169	0.5803	0.0030
W-urbanization	−2.9752	1.7688	0.0969
W-olders	−6.6432	6.4381	0.2343
W-children	15.2469	6.1739	0.0189
W-doctors	−0.6271	0.2303	0.0098
spat.aut.	0.2321	0.0282	0.0000

per thousand, 84% larger than in the previous model. Moreover, the influence on the neighboring untreated regions is approximately 90% of the effect on producing areas (1.34/1.49).

The spatial effects may take place through other channels than spatial lags in the independent variable. Therefore, we introduced spatial controls as suggested by the LM and LM robust tests, which have indicated the SDEM specification in all cases. In this situation, the treatment effect diminished to 1.44, in the unrestricted model. However, the relative importance of the effect on untreated regions (1.35) has increased to 94% of the effect on treated regions (1.35/1.44).

Table 6 shows the complete results of the SDEM model (unrestricted case). The signs of the control variables are as expected. Better social conditions reduce the number of admissions and so do a larger proportion of workers in the population. Urbanization increases the number of hospitalizations, probably reflecting easier access to hospitals and/or that cane burning worsens the pollution in the cities. More elderly people in the region leads to larger numbers of hospitalizations, but the same does not show for the number of children (not significant). The number of doctors increases hospitalizations, also reflecting access to hospitals. The spatial parameter (λ) controls for common shocks to the dependent variable, and is positive and significant.

5. Robustness checks

To verify the robustness of our results, we have evaluated four situations: the identification power of the model, the application to diseases not related, in principle, to cane production, and the use of different forms of measuring the neighborhood effects.

We ran Monte Carlo simulations to investigate the identification power of the unrestricted model in Eq. (8). We have considered the identification of the effect of the treatment on the non-treated as a weighted average, with the weights being the proximity of the neighboring unities. The ML method employed may not identify correctly the effect due to possible multicollinearities, a problem that is more serious in small samples. Therefore, we have set $\beta_1 = 0$, that is, we have simulated only the effects of sugarcane production in non-producing regions in the neighborhood of producing regions. To all simulations, we considered the following model

$$\begin{aligned} Y_t =& C + \phi + \theta_t + (\alpha + \beta_1 \mathbf{W}_{T,T} + \beta_2 \mathbf{W}_{NT,T} + \tau t + \psi \mathbf{W} t) D_t \\ &+ \mathbf{X}_t \gamma\prime + \mathbf{W}\mathbf{X}_t \delta\prime + (\mathbf{I}_n - \lambda \mathbf{W})^{-1} \Upsilon_t. \end{aligned} \tag{9}$$

The dimension of the panel was set in $n = 200$[11] and $t = 10$ (a short panel) and $t = 100$ a long one). The simulation consists of varying β_2, the effect of the treatment on the untreated regions. We simulated $\beta_2 = 0.1, 0.4$, and 0.8, to identify the magnitude at which the effects begin to appear. We chose $C = 10$, similar with the value estimated in the restricted SDEM model (Table 6); and $\tau = \psi = -1$, compatible with the estimated negative trend (Table 6); $\lambda = 0.5$, an average spatial effect. A k-nearest neighbor matrix of spatial weights, with $k = \sqrt{N}$, was constructed to form the y vectors used in the simulations. Pseudo-geographical coordinates were generated from random normal variables. The x-variable is formed by a random vector, with $\gamma = \delta = 1$. The treatment indicator was generated from a random uniform variable with values 0 or 1, and the treatment effect over the treated, α, was set to 1. We ran 1000 draws to form different Υ-vectors of errors. The tests were implemented with the similar models estimated with the real data (a Classical Panel, SLX and SDEM models). We used three different types of W-matrices: Queen-contiguity, k-nearest, and k-nearest weighted by the inverse Euclidean distance.

The simulation results are displayed in Table 7. The coefficients correspond to the average of the 1000 coefficients estimated in each case. Observing the results, it is clear that the SLX and SDEM models, with the true matrix (k-nearest matrix), always hit the parameter (these methods are consistent), but even so inefficient when the parameter indicating that these methods are consistent. However, they are inefficient when the parameter magnitude is tiny (0.1), even when T is high ($T = 100$). In panel data with small dimension ($T = 10$), the SDEM models provided more efficient estimation than the SLX ones (the confidence interval is always smaller in SDEM models than in SLX). In small panels, with larger parameters ($\beta = 0.8$) the classical method always underestimated the true value. Thus, the SDEM model with fixed effects, which is our preferred model, is capable of identifying the true effect.

As another form of robustness check, we ran the same models using the incidence of hospitalizations related to neoplasm pathologies, which are not, in principle, related to sugarcane production, at least in the short run. Given the possibility that some respiratory or skin-related neoplasm cases could be associated to cane burning in the long term, we have excluded these cases from the neoplasm hospitalization set. As the results presented in Table 8 indicate, we found no relationship whatsoever between sugarcane production and the incidence of hospitalizations related to this sort of pathologies. This result suggests that there is no concentration of hospitalizations in the cane-producing areas other than the ones related to the negative externalities generated by cane production.

As another robustness check, we considered different forms for the W-matrix. In the first case, we changed the number of neighbors located within a 100-km radius between 0 and 50; in the second case, we fixed a maximum of $k = 22$ neighbors, and changed the radius between 0 and 100 km. Fig. 3 shows the effects over untreated regions. As the figure shows, the mean effects are close to the estimated SDEM model.

[11] This choice corresponds to a medium sample data, smaller than our empirical one.

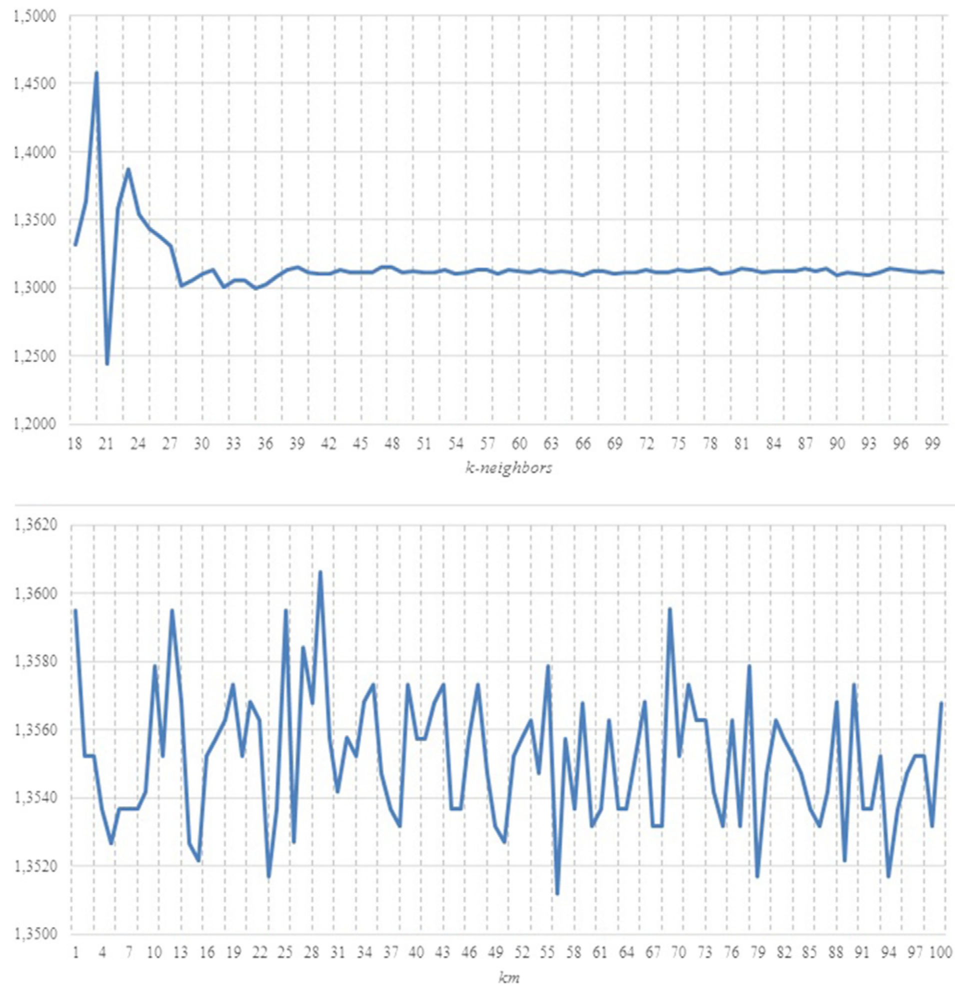

Fig. 3. Robustness checks —different *W* matrices.
Source: Authors' calculation.

Finally, we have considered the direction of the wind in the construction of the neighborhood matrix. Because the transmission mechanism is wind-related, as the particles are transported with the smoke coming from the burning, municipalities located within the same distance could receive different influences depending on their relative position concerning the wind. We have used data from the National Institute of Meteorology, covering 12 measuring stations within São Paulo state and 29 stations located at a maximum distance of 200 km from the state borders. We have used yearly information for the harvest season (from April to September) for all years between 2002 and 2013. Fig. 4 shows the predominant wind direction in the state, according to the parameters adopted. We have considered a 120° window from the municipality centroid for wind dispersion (60° on each side) and a maximum distance of 75 km. Table 9 shows the results of the simulation. The estimated coefficients are similar to the ones obtained in the previous regressions. A Wald-type test of restriction does not allow the rejection of the null hypothesis of equal coefficients for the effect on neighboring municipalities in both regressions (Tables 6 and 9).[12]

[12] This result is in line with the conclusion of (LeSage and Pace, 2014 ,p. 247): “due to the number of common elements in these weight matrices and selection of parameters that give the best fit for each *W*, good fitting models using these different forms of *W* are not likely to produce estimates and inferences that materially differ.”.

Considering these robustness checks, it seems that our results are firm, and the proposed method of measuring the effects of cane production on neighboring regions is adequate.

6. Conclusions

The increasing importance of ethanol as fuel for cars in Brazil has attracted attention for many reasons. Being a biofuel that is environment-friendly, it appeared as a potential solution for the world's dependence on fossil. On the other hand, many negative aspects have been pointed out, such as poor working conditions, soil contamination, dislocation of land used to produce other products and into forested areas. In this article we have investigated one negative externality widely recognized in the literature, that is, the impacts of the burning of the canes on respiratory diseases. Although harvest methods are changing in recent years, both by law enforcement and by new economic incentives for the use of the leaves and the bagasse, the practice of burning the fields to facilitate access to the sugarcane is still prevalent, and will remain so for a while in areas with spiky topography.

The existing studies on the impact of sugarcane production on health conditions do not consider the effects on areas surrounding the plantations. We have developed a new variety of spatial diff-in-diff model to control for the effects of sugarcane production on neighboring non-producing regions, introducing spatial effects also in the independent variables, through a SDEM model. This method

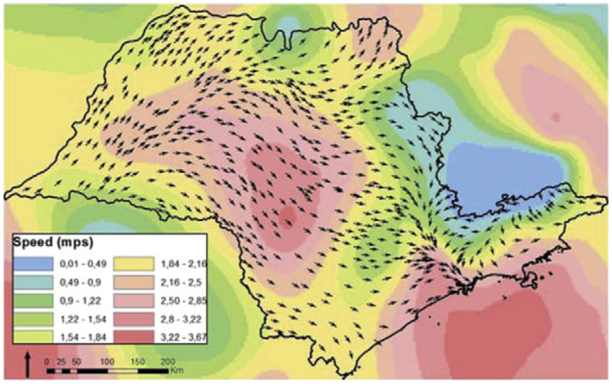

Fig. 4. Wind speed and directions.
Source: INMET. Authors calculation.

is an addition to the Spatial Econometrics literature, as it includes spatial effects on treated and untreated regions in a comprehensive way, so that the effects on both producing and surrounding non-producing regions can be properly estimated.

We have introduced control variables related to socioeconomic conditions in the regions, such as the presence of children and elders, the share of population employed, the number of doctors, and the urbanization rate. The estimated coefficients for these variables came out as expected. As for the spatial effects, the results indicate that the impacts on the producing regions are 78% larger than if the effects on the surrounding areas were ignored. This indicates that ignoring the effects on surrounding areas in the calculations underestimates the effects on the producing areas in the calculations underestimates the effects on the producing areas themselves. Moreover, the effects on the surrounding areas, typically ignored in other studies, are relevant, and almost as large as the effects on the producing areas. Again, ignoring the neighborhood effects underestimates the impacts on hospitalizations in the area at large. We have implemented robustness checks that gave us more confidence on the results, as they have indicated that they are not related to specificities of the regions considered.

These findings are important for the planning of the distribution of health facilities across regions. It is clear that sugarcane production tends to increase hospitalizations due to respiratory causes not only in the producing municipalities, but also in the vicinity. In addition, the quantitative effects are much larger than if the spatial effects were ignored. Therefore, planning the organization of the health services to cope with this kind of negative externality must consider larger numbers of hospitalization requests, and should consider broader areas, involving both producing and non-producing municipalities.

Table 9
Robustness check —wind neighbors matrix.

Dependent variable		intern per th	
R-squared		0.7831	
corr-squared		0.0502	
Within R2		0.1842	
Between R2		0.025	
Overall R2		0.0121	
sigma2^		9.4824	
Log-likelihood		−19,349.773	
AIC		5.0116	
Wald test, p-value		276.6881, 0.0000	
Nobs,Nvar,#FE		7728, 15, 670	
# iterations		16	
Min and max λ		−0.9900, 0.9900	
Variable	Coefficient	Standard-error	p-Value
Constant	10.8825	8.0906	0.1615
Treatment	1.1777	0.3352	0.0008
$\mathbf{W}_{T,T}D_t$	−0.1326	0.5312	0.3867
$\mathbf{W}_{NT,T}D_t$	1.3679	0.5652	0.0213
$trend_d$	0.2447	0.4609	0.3465
$trend_{wd}$	−2.0516	0.6786	0.0041
Workers	−1.5459	0.5618	0.0091
Urbanization	6.8309	1.8919	0.0006
Olders	15.6748	9.0280	0.0884
Children	−8.7973	8.9873	0.2471
Doctors	0.6006	0.2365	0.0159
W-workers	−5.2374	3.0255	0.0892
W-urbanization	−10.0147	9.2938	0.2232
W-olders	−50.2464	33.5926	0.1303
W-children	39.1678	31.7675	0.1866
W-doctors	−1.1589	1.2055	0.2513
spat.aut.	0.3793	0.0268	0.0000
Wald restriction test, p-value		0.0010, 0.9752	

Acknowledgments

The authors acknowledge the relevant and interesting comments by Paul Elhorst and Raymond Florax on a preliminary version of this paper, presented at the European Regional Science Association meeting held in Lisbon, August 2015, and also the comments and suggestions of the two anonymous referees. Of course, the content of this article is the sole responsibility of the authors.

References

Allen, A.G., Cardoso, A.A., Rocha, G.O., 2004. Influence of sugar cane burning on aerosol soluble ion composition in Southeastern Brazil. Atmospheric Environment 38 (30), 5025–5038.
Alves, F.J.C., 2006. Por que morrem os cortadores de cana? Saúde e Sociedade 15 (3). 90–98.
Alves, F.J.C., 2007. Migração de trabalhadores rurais do Maranhão e Piauípara o corte de cana em são paulo - será este um fenômeno casual ou recorrente da estratégia empresarial do complexo agroindustrial canavieiro? In: Novaes, R., Alves, F.J.C. (Eds.), Migrantes: trabalho e trabalhadores no complexo agroindustrial canavieiro - os heróis do agronegócio brasileiro. EDUFSCar, São Carlos, pp. 21–54.
Angelucci, M., Giorgi, G.D., 2009. Indirect effects of an aid program: how do cash transfers affect ineligibles' consumption? Am. Econ. Rev. 99 (1). 486–508.
Anselin, L., 1988. Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Dordecht.
Anselin, L., Bera, A.K., Florax, R.J.G.M., Yoon, M.J., 1996. Simple diagnostic tests for spatial dependence. Reg. Sci. Urban Econ. 26 (1), 77–104.
Arbex, M.A., Bohm, G.M., Saldiva, P.H., Conceição, G., 2000. Assessment of the effects of sugar cane plantation burning on daily counts of inhalation therapy. J. Air Waste Manag. Assoc. 50 (10), 1745–1749.
Arbex, M.A., Cançado, J.E.D., Pereira, L.A.M., Braga, A.L.F., Saldiva, P.H.N., 2004. Queima de biomassa e efeitos sobre a saúde. J. Bras. Pneumol. 30 (2), 158–175.
Arbex, M.A., Martins, L.C., Oliveira, R.C., Pereira, L.A.A., Arbex, F.F., Cançado, J.E.D., Saldiva, P.H.N., Braga, A.L.F., 2007. Air pollution from biomass burning and asthma hospital admissions in a sugar cane plantation area in Brazil. J. Epidemiol. Community Health 61 (5), 395–400.
Arbex, M.A., Pereira, L.A.A., Carvalho-Oliveira, R., Saldiva, P.H.N., Braga, A.L.F., 2014. The effect of air pollution on pneumonia-related emergency department visits in a region of extensive sugar cane plantations: a 30-month time-series study. J. Epidemiol. Community Health 68 (7), 669–674.
Ashenfelter, O., Card, D., 1985. Using the longitudinal structure of earnings to estimate the effect of training programs. Rev. Econ. Stat. 67 (4), 648–660.
Baccarin, J.G., Alves, F.J.C., Gomes, L.F.C., 2008. Emprego e condições de trabalho dos canavieiros no centro-sul do Brasil, entre 1995 e 2007. Anais do XLVI Congresso da Sober. Sociedade Brasileira de Economia e Sociologia Rural, Rio Branco.
Berniell, L., de la Mata, D., Valdés, Nieves, 2013. Spillovers of health education at school on parents' physical activity. Health Econ. n/a, n/a–n/a.
BNDES, CGEE, 2008. Bioetanol de cana-de-açúcar: energia para o desenvolvimento. BNDES, Rio de Janeiro.
Braga, A.L.F., Conceição, G.M.S., Pereira, L.A.A., Kishi, H.S., Pereira, J.C.R., Andrade, M.F., Gonçalves, F.L.T., Saldiva, P.H.N., Latorre, M.R.D.O., 1999. Air pollution and pediatric respiratory hospital admissions in São Paulo, Brazil. J. Environ. Med. 1 (2), 95–102.
Cançado, J.E.D., Saldiva, P.H.N., Pereira, L.A.A., Lara, L.B.L.S., Artaxo, P., Martinelli, L.A., Arbex, M.A., Zanobetti, A., Braga, A.L.F., 2006. The impact of sugar cane burning emissions on the respiratory system of children and the elderly. Environ. Health Perspect. 114 (5), 725–729.
Carneseca, E.C., Achcar, J.A., Martinez, E.Z., 2012. Association between particulate matter air pollution and monthly inhalation and nebulization procedures in Ribeirão Preto, São Paulo State, Brazil. Cad. Saude Publica 28 (8), 1591–1598.

Chagas, A.L.S., Toneto-Jr, R., Azzoni, C.R., 2008. Teremos que trocar energia por comida? análise do impacto da expansão da produção de cana-de-açácar sobre o preço da terra e dos alimentos. Economia (Brasília) 39-61, 39–61.
Chagas, A.L.S., Toneto-Jr, R., Azzoni, C.R., 2011. A spatial propensity score matching evaluation of the social impacts of sugarcane growing on municipalities in Brazil. Int. Reg. Sci. Rev. 35, 48–69.
Chen, Xiaoguang, Ye, Jingjing, 2015. When the wind blows: spatial spillover effects of urban air pollution, environment for development. Working Paper.
Ciriaci, D., Palma, D., 2010. Geography, environmental efficiency and Italian economic growth: a spatially-adapted environmental Kuznets curve. Working Paper. European Commission, JRC Institute for Prospective Technological Studies, Enea,
Cox, D.R., 1958. Planning of Experiments. Wiley, New York.
Delgado, M.S., Florax, R.J., 2015. Difference-in-differences techniques for spatial data: local autocorrelation and spatial interaction. Econ. Lett. 137, 123–126.
Dinda, S., 2004. Environmental Kuznets curve hypothesis: a survey. Ecol. Econ. 49 (4), 431–455.
Dubé, J., Legros, D., Thériault, M., Rosiers, F.D., 2014. A spatial difference-in-differences estimator to evaluate the effect of change in public mass transit systems on house prices. Transp. Res. B 64, 24–40.
Elhorst, J.P., 2010. Matlab software for spatial panels. Paper Presented at 4th World Conference of the Spatial Econometric Association, Chicago.
Elhorst, J.P., 2010. Spatial panel data models. In: Fischer, M.M., Getis, A. (Eds.), Handbook of Applied Spatial Analysis. Springer, pp. 377–407.
Elhorst, J.P., 2014. Spatial Econometrics: From Cross-sectional Data to Spatial Panels. Springer
Fang, C., Liu, H., Li, G., Sun, D., Miao, Z., 2015. Estimating the impact of urbanization on air quality in China using spatial regression models. Sustainability 7 (11),15570.
Fischer, P., Hoek, G., Brunekreef, B., Verhoeff, A., van Wijnen, J., 2003. Air pollution and mortality in the Netherlands: are the elderly more at risk?. Eur. Respir. J. 21 (Suppl. 40), 34s–38s.
Florax, R.J., Folmer, H., Rey, S.J., 2003. Specification searches in spatial econometrics: the relevance of Hendry's methodology. Reg. Sci. Urban Econ. 33, 557–579.
Gangadharan, L., Valenzuela, M., 2001. Interrelationships between income, health and the environment: extending the environmental Kuznets curve hypothesis. Ecol. Econ. 36 (3), 513–531.
Goldemberg, J., Coelho, S.T., Guardabassi, P., 2008. The sustainability of ethanol production from sugarcane. Energy Policy 36 (6), 2086–2097.
Goncalves, F.L.T., Carvalho, L.M.V., Conde, F.C., Latorre, M.R.D.O., Saldiva, P.H.N., Braga, A.L.F., 2005. The effects of air pollution and meteorological parameters on respiratory morbidity during the summer in So Paulo City. Environ. Int. 31 (3), 343–349.
a, M.L., Goto, D.M., Obuti, C.A., ao Barbosa, C.M.G., Saldiva, P.H.N., Zanetta, D.M.T., Lorenzi-Filho, G., de Paula Santos, U., Nakagawa, N.K., 2011. Effects of biomass burning on nasal mucociliary clearance and mucus properties after sugarcane harvesting. Environ. Res. 111 (5), 664–669.
Grossman, G.M., Krueger, A.B., 1991. Environmental impacts of a north American free trade agreement, National Bureau of Economic Research. Working Paper 3914.
Grossman, G.M., Krueger, A.B., 1995. Economic growth and the environment. Q. J. Econ. 110 (2), 353–377.
Hao, Y., Liu, Y.M., 2016. The influential factors of urban pm2.5 concentrations in China: a spatial econometric analysis. J. Clean. Prod. 112 (Part 2), 1443–1453. Preventing Smog Crises.
Heckert, M., Mennis, J., 2012. The economic impact of greening urban vacant land: a spatial difference-in-differences analysis. Environ. Plan. A 44, 3010–3027.
Hosseini, H.M., Kaneko, S., 2013. Can environmental quality spread through institutions?. Energy Policy 56, 312–321.
Kaboski, J.P., Townsend, R.M., 2012. The impact of credit on village economies. Am. Econ. J. Appl. Econ. 4 (2), 98–133.
Kelejian, H.H., Prucha, I.R., Yuzefovich, Y., 2006. Estimation problems in models with spatial weighting matrices which have blocks of equal elements. J. Reg. Sci. 46 (3), 507–515.
Lagravinese, R., Moscone, F., Tosetti, E., Lee, H., 2014. The impact of air pollution on hospital admissions: evidence from Italy. Reg. Sci. Urban Econ. 49, 278–285.
Lara, L.L., Artaxo, P., Martinelli, L.A., Camargo, P.B., Victoria, R.L., Ferraz, E.S.B., 2005. Properties of aerosols from sugar-cane burning emissions in Southeastern Brazil. Atmos. Environ. 39 (26), 4627–4637.
Lee, L.F., Yu, J., 2010. Estimation of spatial autoregressive panel data models with fixed effects. J. Econ. 154, 165–185.
LeSage, J., Pace, R.K., 2009. Introduction to Spatial Econometrics. CRC Press, Boca Raton.
LeSage, J.P., Pace, R.K., 2014. The biggest myth in spatial econometrics. Econometrics 2, 217–249.
Li, Q., Song, J., Wang, E., Hu, H., Zhang, J., Wang, Y., 2014. Economic growth and pollutant emissions in China: a spatial econometric analysis. Stochastic Environmental Research and Risk Assessment 28 (2), 429–442.
Maddison, D., 2006. Environmental Kuznets curves: a spatial econometric approach. J. Environ. Econ. Manag. 51 (2), 218–230.
Maddison, D., 2007. Modelling sulphur emissions in Europe: a spatial econometric approach. Oxf. Econ. Pap. 59 (4), 726–743.
Mazzoli-Rocha, F., Magalhães, C.B., Malm, O., Saldiva, P.H., Zin, W.A., Faffe, D.S., 2008. Comparative respiratory toxicity of particles produced by traffic and sugar cane burning. Environ Res. 108 (1), 35–41.
Nicolella, A.C., Belluzzo, W., 2015, 2. The effect of reducing the pre-harvest burning of sugar cane on respiratory health in Brazil. Environ. Dev. Econ. 20, 127–140.
Noronha, S., Ortiz, L., Schlesinger, S., 2006. Agronegócio e biocombustveis: uma mistura explosiva - Impactos da expansão das monoculturas para a produção de Bioenergia. Núcleo Amigos da Terra, Rio de Janeiro.
Pandit, R., Laband, D.N., 2007. General and specific spatial autocorrelation: insights from country-level analysis of species imperilment. Ecol. Econ. 61 (1), 75–80.
Renard, M.-F., Xiong, H., 2012. Strategic interactions in environmental regulation enforcement: evidence from Chinese provinces. Etudes et Documents E 2012.07, CERDI
Ribeiro, H., 2008. Queimadas de cana-de-açúcar no Brasil: efeitos á sade respiratória. Revista de Saúde Pública 42, 370–376.
Roseiro, M.N.V., 2002. Morbidade por problemas respiratórios em Ribeirão Preto-SP, de 1995 a 2001, segundo indicadores ambientais, sociais e econâmicos. Tese de Doutorado, Universidade de São Paulo, escola de Enfermagem de Ribeirão Preto.
Rosenbaum, P.R., 2010. Design of Observational Studies. Springer, New York.
Rupasingha, A., Goetz, S.J., Debertin, D.L., Pagoulatos, A., 2004. The environmental Kuznets curve for US counties: a spatial econometric analysis with extensions. Pap. Reg. Sci. 83 (2), 407–424.
Santejo Silveira, H.C., Schmidt-Carrijo, M., Henrique Seidel, E., Scapulatempo-Neto, C., Longatto-Filho, A., Lopes Carvalho, A., Vieira Reis, R.M., Nascimento Saldiva, P.H., 2013. Emissions generated by sugarcane burning promote genotoxicity in rural workers: a case study in Barretos, Brazil. Environ. Health 12 (1),
Sicard, P., Mangin, A., Hebel, P., Malléa, P., 2010, Mar. Detection and estimation trends linked to air quality and mortality on French Riviera over the 1990–2005 period. Sci. Total Environ. 408 (8), 1943–1950.
Stern, D.I., 2000. Applying recent developments in time series econometrics to the spatial domain. Prof. Geogr. 52 (1), 37–49.
Su, Z.-f., HU, R.-d., LIN, S.-q., 2009. Spatial econometric analysis of Kuznets' relationship between environmental quality and economic growth. Geogr. Res. 28 (2), 303
Sun, R., Gu, D., 2008. Air pollution, economic development of communities, and health status among the elderly in urban China. Am. J. Epidemiol. 168 (11), 1311–1318.
Toneto-Jr, R., Liboni, L.B., 2008. Mercado de trabalho da cana-de-açúcar. Anais do I Workshop do Observatório do Setor Sucroalcooleiro, Ribeirão Preto.
Uriarte, M., Yackulic, C.B., Cooper, T., Flynn, D., Cortes, M., Crk, T., Cullman, G., McGinty, M., Sircely, J., 2009. Expansion of sugarcane production in So Paulo, Brazil: implications for fire occurrence and respiratory health. Agric. Ecosyst. Environ. 132 (1–2), 48–56.
Vega, S.H., Elhorst, J.P., 2015. The SLX model. J. Reg. Sci. http://dx.doi.org/10.1111/jors.12188.
Wang, L., Wei, B., Li, Y., Li, H., Zhang, F., Rosenberg, M., Yang, L., Huang, J., Krafft, T., Wang, W., 2014. A study of air pollutants influencing life expectancy and longevity from spatial perspective in China. Sci. Total Environ. 487, 57–64.
Wang, W., Ying, Y., Wu, Q., Zhang, H., Ma, D., Xiao, W., 2015. A GIS-based spatial correlation analysis for ambient air pollution and AECOPD hospitalizations in Jinan, China. Respir. Med. 109 (3), 372–378.
Wen, M., Gu, D., 2012. Air pollution shortens life expectancy and health expectancy for older adults: the case of China. J. Gerontol. A Biol. Sci. Med. Sci. 67 (11), 1219–1229.
Won Kim, C., Phipps, T.T., Anselin, L., 2003. Measuring the benefits of air quality improvement: a spatial hedonic approach. J. Environ. Econ. Manag. 45 (1), 24–39.