

Chagas, Andre; Azzoni, Carlos; Almeida, Alexandre

Conference Paper

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

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Chagas, Andre; Azzoni, Carlos; Almeida, Alexandre (2015) : A spatial Diff-in-Diff analysis of the impact of sugarcane production on respiratory diseases, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124670>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

A Spatial Diff-in-Diff Analysis of the Impact of Sugarcane Production on Respiratory Diseases

André L. S. Chagas

Department of Economics, University of São Paulo

achagas@usp.br

Carlos R. Azzoni

Department of Economics, University of São Paulo

cazzoni@usp.br

Alexandre N. Almeida

Department of Applied Economics, Esalq-University of Sao Paulo

alex.almeida@usp.br

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 spatial Diff-in-Diff model to control for the effects of sugarcane production in neighboring non-producing regions. This method is a contribution 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 about three times as large as the situation in that the effects on the surrounding areas were ignored. Moreover, the effects on the surrounding areas, typically ignored in other studies, are relevant, amounting to 80% of the effects on the producing areas.

Key-Words: Spatial Diff-in-Diff; Sugarcane Producing, Health Conditions, Spatial Econometrics

Jel Classification: C21, C23, I18, R11, Q18, Q42

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 both on 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 around 3.5% of Brazilian industrial GDP. The sector as a whole employs over 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, 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.

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 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.

Sugarcane is harvested by unskilled workers, mostly manually. The traditional harvest method involves the burning of the cane to facilitate access to the plant, and one issue relates to the possible negative health effects of this process. Burning the field 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 as 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 relate sugarcane burning to the increase in fine particulate matter, coarse particulate matter, and black carbon concentration, especially during burning hours (Lara et al., 2005), and increases the air concentration of substances as nitrite, sulfite, oxide of carbon, and others (Allen et al., 2004). The literature also relates that short and long-term exposition to classical pollutants (matter, sulfite, nitrite, oxide carbon, etc.) can negatively affect the workers (Sicard et al., 2010), especially the young, and the elderly (Braga et al., 1999; Roseiro, 2002; Gonçalves et al., 2005).

Sugarcane burning generates a massive quantity of smoke that spreads all over the region, reaching cities and becoming a potential threat to human health. Pollution from sugarcane burning may be as harmful as pollution from traffic and manufacturing activities (Mazzoli-Rocha et al., 2008). Some studies related its impacts on health for specific municipalities or regions (Arbex et al., 2000, 2004; Cançado et al., 2006; Arbex et al., 2007; Ribeiro, 2008; Uriarte et al., 2009; Carneseca et al., 2012). These studies focused on the short-distance effects of burning, considering only the association of burning and the incidence of respiratory health problems at the local level, but they fail to capture the consequences of burning events on other places.

This paper presents a spatial difference-in-difference model that is developed and applied to control for the effect of sugarcane production on both producing (treated) and non-producing (untreated) regions located in the vicinity. This method of measuring the effects is more complete than the methods used in previous studies, and provides contributions both to the discussion of the true effects of sugarcane production on health and to the measurement of spatial effects in general. The article is organized in four sections, including this introduction. The next section presents the methodology used to identify the possible impacts of growing sugarcane on the respiratory health conditions in the producing regions, and the data utilized. The third section presents the results; the fourth section checks the robustness; the fifth section contains the final remarks.

2 Methodology and data

2.1 The model

As usual in spatial studies, we take into account that regions are interrelated. This generates the possibility of propagation of the effects both on the region where production takes place (treated region) and on the surrounding areas (untreated region). In the equations that follow, y_{it} is the variable of interest, hospitalizations due to respiratory diseases; x is a vector of observable characteristics; w_i is a $n \times 1$ vector associating each region to all the other regions; d_{it} is a $n \times 1$ vector of values $d_{it} = 1$ if the region is treated, and $d_{it} = 0$ otherwise. The subscript i refers to region, and the subscript t refers to time (year). We consider two situations for each region: before (b) and after (a) the treatment. In the before-treatment situation, we have

$$\begin{aligned} y_{it,0}^b &= \mu(\mathbf{x}) + u_{it} \\ y_{it,1}^b &= y_{it,0}^b \end{aligned}$$

where, $y_{it,0}^b$ is the dependent variable in the untreated region, before the treatment, and $y_{it,1}^b$ is the dependent variable in the treated region. In the after-treatment situation, we can identify two impacts: 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

$$\begin{aligned} y_{it,0}^a &= \mu(\mathbf{x}) + \mathbf{w}_i d_{it} \beta + u_{it} \\ y_{it,1}^a &= y_{it,0}^a + \alpha \end{aligned}$$

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 $\mathbf{w}_i d_{it}$. Defining D_{it} as a region- i specific indicator of treatment in time t , we can write

$$y_{it} = (1 - D_{it})y_{it,0} + D_{it}y_{it,1} \tag{1}$$

Using the “before” and “after” definitions, three effects can be computed: ATE (Av-

verage Treatment Effect), ATET (Average Treatment Effect on the treated), and ATENT (Average Treatment Effect on the untreated), as follows

$$\begin{aligned}
\text{ATE} &= E[y_{it,1}^a - y_{it,1}^b] - E[y_{it,0}^a - y_{it,0}^b] \\
&= \alpha \\
\text{ATET} &= E[y_{it,1}^a - y_{it,1}^b] \\
&= \alpha + \mathbf{w}_i d_{it} \beta \\
\text{ATENT} &= E[y_{it,0}^a - y_{it,0}^b] \\
&= \mathbf{w}_i d_{it} \beta
\end{aligned}$$

In matrix notation, we have

$$Y = \mu(\mathbf{X}) + (\alpha + \mathbf{I}_t \otimes \mathbf{W}\beta)D + U \quad (2)$$

where Y is a $nt \times 1$ matrix of observations, \mathbf{X} is a $nt \times k$ matrix of covariates, D is a *dummy* variable indicating treated regions, \mathbf{I}_t is a square identity matrix of $t \times t$ dimension, \mathbf{W} is a $n \times n$ neighborhood weight matrix and U is a vector of errors of $nt \times 1$ dimension. μ , α and β are parameters to be estimated.

The term $\mathbf{I}_t \otimes \mathbf{W}\beta$ indicates the indirect effect of the treatment on both regions, treated and untreated. This effect is usually ignored in estimations of this type¹. 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 of the treatment 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 treatment, because β will be estimated as an average of the indirect effects on the treated and on the untreated regions.

Consider, for clarity, the following decomposition of the \mathbf{W} matrix,

¹Angelucci and Giorgi (2009); Kaboski and Townsend. (2012); Berniell et al. (2013) are some exceptions. However, these studies do not control for different structures of neighborhood.

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

where

$$\begin{aligned}\mathbf{W}_{T,T} &= \text{diag}(D) \times (\mathbf{I}_t \otimes \mathbf{W}) \times \text{diag}(D) \\ \mathbf{W}_{T,NT} &= \text{diag}(D) \times (\mathbf{I}_t \otimes \mathbf{W}) \times \text{diag}(\iota - D) \\ \mathbf{W}_{NT,T} &= \text{diag}(\iota - D) \times (\mathbf{I}_t \otimes \mathbf{W}) \times \text{diag}(D), \text{ and} \\ \mathbf{W}_{NT,NT} &= \text{diag}(\iota - D) \times (\mathbf{I}_t \otimes \mathbf{W}) \times \text{diag}(\iota - D)\end{aligned}$$

As D is a *dummy* variable associated to the information on the treatment, and $\text{diag}(D)$ is a $nt \times 1$ matrix with D in the main diagonal and zeros elsewhere, ι represents a vector of 1's, and \mathbf{W}_{ij} represents the neighborhood effects of the j -region on i -region, $i, j = T$ (treated) or NT (untreated). Substituting in (2), results in

$$Y = \mu(\mathbf{X}) + [\alpha + (\mathbf{W}_{T,T} + \mathbf{W}_{T,NT} + \mathbf{W}_{NT,T} + \mathbf{W}_{NT,NT})\beta]D + U$$

Then, it is clear that β represents an average effect, as we mentioned above. A more realistic model considers different effects for dissimilar \mathbf{W} matrices. As, by construction, $\mathbf{W}_{T,NT}D$ and $\mathbf{W}_{NT,NT}D$ are $\mathbf{0}$ -vectors, the unrestricted model is

$$Y = \mu(\mathbf{X}) + [\alpha + (\mathbf{W}_{T,T}\beta_1 + \mathbf{W}_{NT,T}\beta_2)]D + U \quad (3)$$

The models in (2) and (3) are a special form of the Spatial Diff-in-Diff models (SDID). It is important to register that they do not contain a traditional spatial interaction effect, such as in the Spatial Autoregressive Model (SAR) and 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

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

and/or

$$E = \lambda(\mathbf{I}_t \otimes \mathbf{W})U$$

In the first equation, \mathbf{X} is a $n \times k$ vector of observable characteristics, \mathbf{W} is a spatial weight matrix of $n \times n$ dimension, γ is a $1 \times k$ parameter vector to be estimated, and ρ is the spatial auto-regressive parameter. In the second equation, E is an error vector, not spatially associated, λ is the spatial error parameter to be estimated. Thus, a complete version of models 2 and 3 is

$$Y = [\mathbf{I}_{nt} - \rho(\mathbf{I}_t \otimes \mathbf{W})]^{-1} \{ \mathbf{X}\gamma' + [\alpha + (\mathbf{W}_{T,T}\beta_1 + \mathbf{W}_{NT,T}\beta_2)]D + [\mathbf{I}_{nt} - \lambda(\mathbf{I}_t \otimes \mathbf{W})]^{-1}U \} \quad (4)$$

This is an addition to the Spatial Econometrics literature. Previously, Heckert and Mennis (2012) have estimated a Diff-in-Diff model to study observed changes in property values in the city of Philadelphia, after the removal of debris coupled with planting grass and trees in vacant areas, as a means of improving blighted communities. They use a geographically weighted regression model to estimate the effects on surrounding treated vacant lots with observed changes around lots that might have been treated but were not. They concluded that properties surrounding greened vacant lots had a greater increase in value than properties surrounding non-greened vacant lots. By developing both global and local versions of the model, they explored the spatial variations in the impacts of the program. Dubé et al. (2014) proposed a spatial Diff-in-Diff estimator to study the impacts of a new commuter rail transit system on real-estate values in Montreal, Canada. The proposed estimator accounts for possible spatial spillovers and they compare the results with the ones obtained from the usual DID model.

These were steps forward in considering the expected influence on untreated regions as well. However, the proposed models do not estimate the influence in the way the problem at hand requires. Applying their technique in our case would be equivalent to considering that hospitalizations in one region influence hospitalizations in the neighborhood. Instead, what we need to know is the influence of the production of sugarcane in the treated region on hospitalizations in both regions, treated and untreated. Therefore, we use an extension of the SEM model that includes spatial lags of the independent regressors, the

Spatial Durbin Error Model (SDEM), and consider both restricted (2) and unrestricted (3) situations.

2.2 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 lands 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. In Fig. 1 we map the evolution of sugarcane production in São Paulo state, by municipality, during the period 2002-2011 ². It is clear that there was a steady increase in the production in northwest region of the state, and a sprawl to the west of the state, a previously pasture area.

[FIGURE 1 HERE]

The production of sugarcane is important to define our treatment variable. We consider as treated any municipality in which the share 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.

Our variable of interest is the number of persons hospitalized to be treated for respiratory problems (per 1,000 inhabitants). The data is provided by DATASUS³, the statistical

²We select some years in this period, but the evolution is evident.

³<http://www2.datasus.gov.br/DATASUS/index.php>.

Table 1: Number of treated region and proportion on total, 2002-2011

Year	Number of treated region	Prop. 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.

agency of the Ministry of Health, and includes hospitalization records for public and private hospitals. The information is highly disaggregated in spatial terms, and we use data at the municipality level. Fig. 2 exhibits the evolution of the number of hospitalizations of interest in São Paulo state. Over time, the number of cases of hospitalizations due to respiratory diseases is decreasing. 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 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.

[FIGURE 2 HERE]

Given that, we have introduced a trend variable to adjust for this empirical evidence in all models estimated, and the coefficients are negative and significant in all cases. We have included variables to control for socioeconomic 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 literature (Braga et al., 1999; Roseiro, 2002; Gonçalves et al., 2005). Finally, we consider 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

Table 2: Hospitalization due respiratory health problem, by region, 2002-2011

Year	Hospitalization due to respiratory problems			
	Mean	Std. Dev	Maximum	Minimun
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.

correlation matrix.

3 Results

This section presents the results⁴. We compute and compare five models. The first is a classical panel data regression with fixed effects, to set a baseline for comparing the results with spatial controls included. The second includes only spatial lag on treatment variable, without including spatial lag in other independent variables. In this way, this model is similar to the literature on peer effects. In the third case we include 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 fourth case involves the traditional SAR and SEM methods in spatial econometrics. The selection was based on the Lagrange Multiplier (LM) tests - LM and

⁴We based our estimates on Elhorst's routine for spatial panel data models (Elhorst, 2010b,a). 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 weights 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 contain 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.

Table 3: Summary statistics for the variables

Variable	Mean	Sdt. Dev.	Max	Min	N
Treatment	0.500	0.500	1.000	0.000	7,728
WD	0.507	0.371	1.000	0.000	7,728
W_1D	0.399	0.427	1.000	0.000	7,728
W_2D	0.108	0.203	1.000	0.000	7,728
Workers	0.202	0.143	2.202	0.036	7,728
Urbanization	0.831	0.146	1.000	0.221	7,728
Olders	0.125	0.030	0.251	0.043	7,728
Children	0.229	0.034	0.366	0.071	7,728
Doctors	0.675	0.847	7.000	0.000	7,728
Wworkers	0.202	0.056	0.456	0.089	7,728
Wurbanization	0.832	0.076	0.997	0.472	7,728
Wolders	0.125	0.021	0.198	0.057	7,728
Wchildren	0.229	0.027	0.332	0.164	7,728
Wmedicos	0.682	0.232	1.766	0.170	7,728

Source: Authors calculations.

Table 4: Linear correlation between variable of the model

Variabels	Treatment	WD	W_1D	W_2D	Workers	Urbanization	Olders
Treatment	1.0000						
WD	0.7856	1.0000					
W_1D	0.9342	0.8805	1.0000				
W_2D	-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
Wurbanization	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

Variabels	Children	Doctors	Wworkers	Wurbanization	Wolders	Wchildren	Wmedicos
Treatment							
WD							
W_1D							
W_2D							
Workers							
Urbanization							
Olders							
Children	1.0000						
Doctors	-0.2501	1.0000					
Wworkers	-0.2222	0.1412	1.0000				
Wurbanization	-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.

LM robust tests (Anselin et al., 1996), following the suggestion by Florax et al. (2003)⁵. For all situations the tests indicated the use of SEM model. Finally, we considered an extension of the SEM model, including spatial lags of the independent regressors, a SDEM - Spatial Durbin Error Model (LeSage, 2014). Additionally, we consider both restricted and unrestricted cases, as models (2) and (3), respectively. We consider a k -nearest neighbors distance matrix, with k varying between 20 and 40, and used the Akaike Information Criterion(AIC) for pooled models, without spatial effects, to choose the order that minimizes the AIC criterion. Thus, we chose the matrix that best fits the data.

The models estimated are shown in Table 5. The dependent variable (y_{it}) is the number of hospitalizations (per 1,000 inhabitants), in region i in time t ; \mathbf{w}_i is a vector of neighborhood weights; $\mathbf{w}_{iNT,T}$ is a vector of weights associating treated to untreated neighbors; and $\mathbf{w}_{iT,T}$ is a vector of weights associating treated to treated neighbors; \mathbf{x}_{it} is a vector of control variables, including the constant; u_{it} is an error term, normally distributed, centered in zero and with constant variance. As usual in spatial econometrics, $\omega_i(\lambda)$ represents the i -row of matrix $(I - \lambda\mathbf{W})^{-1}$, a global matrix that associates shocks in a specific region to all the other regions. D_{it} is a indicator of treatment, and \mathbf{d}_t is a vector of all indicators in t . The parameters α , β 's, γ , δ , and λ are estimated.

Table 5: Models estimated

Case		Expression
Classical Panel		$y_{it} = \alpha D_{it} + \mathbf{x}_{it}\gamma' + u_{it}$
Restricted	Spatial control	$y_{it} = \alpha D_{it} + \beta \mathbf{w}_i' \mathbf{d}_t + \mathbf{x}_{it}\gamma' + u_{it}$
	SLX model	$y_{it} = \alpha D_{it} + \beta \mathbf{w}_i' \mathbf{d}_t + \mathbf{x}_{it}\gamma' + \mathbf{w}\mathbf{x}_{it}\delta' + u_{it}$
	SEM model	$y_{it} = \alpha D_{it} + \beta \mathbf{w}_i' \mathbf{d}_t + \mathbf{x}_{it}\gamma' + \omega_i(\lambda)u_{it}$
	SDEM model	$y_{it} = \alpha D_{it} + \beta \mathbf{w}_i' \mathbf{d}_t + \mathbf{x}_{it}\gamma' + \mathbf{w}\mathbf{x}_{it}\delta' + \omega_i(\lambda)u_{it}$
Unrestricted	Spatial control	$y_{it} = \alpha D_{it} + \beta_1 \mathbf{w}_{iT,T}' \mathbf{d}_t + \beta_2 \mathbf{w}_{iNT,T}' \mathbf{d}_t + \mathbf{x}_{it}\gamma' + u_{it}$
	SLX model	$y_{it} = \alpha D_{it} + \beta_1 \mathbf{w}_{iT,T}' \mathbf{d}_t + \beta_2 \mathbf{w}_{iNT,T}' \mathbf{d}_t + \mathbf{x}_{it}\gamma' + \mathbf{w}\mathbf{x}_{it}\delta' + u_{it}$
	SEM model	$y_{it} = \alpha D_{it} + \beta_1 \mathbf{w}_{iT,T}' \mathbf{d}_t + \beta_2 \mathbf{w}_{iNT,T}' \mathbf{d}_t + \mathbf{x}_{it}\gamma' + \omega_i(\lambda)u_{it}$
	SDEM model	$y_{it} = \alpha D_{it} + \beta_1 \mathbf{w}_{iT,T}' \mathbf{d}_t + \beta_2 \mathbf{w}_{iNT,T}' \mathbf{d}_t + \mathbf{x}_{it}\gamma' + \mathbf{w}\mathbf{x}_{it}\delta' + \omega_i(\lambda)u_{it}$

The results are reported in Table 6. A trend term is introduced in all models, for the reasons mentioned above, with negative and significant coefficients, as expected. The

⁵Elhorst provided the routines for the spatial panel data case. These tests have become very 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.

Classical Panel Model indicates that sugarcane production increases hospitalizations by only 0.48 cases per thousand, and this conclusion is significant at 1%. 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 1.92 cases per thousand, and the influence on neighboring untreated regions is about 45% of the effect on producing areas (0.86/1.92).

These models do not consider the social and demographic conditions in the neighbours, which might influence hospitalizations, such as urbanization and the share of children and elderly in the region's population. In the presence of these controls (SLX model), the treatment effect increases marginally to 1.97, revealing the importance of the controls, as compared to the previous models. The unrestricted model shows an impact of 1.11 on treated regions, and impacts on neighboring untreated regions around 56% of the impact on treated regions (1.11/1.97).

The spatial effects may take place through other channels than spatial lags in the independent variable. Therefore, we introduce spatial controls as suggested by the LM and LM robust test, which have indicated the SEM specification in all cases. In this situation, the treatment effect diminished, and is now between 0.57 and 1.69. However, the effect on untreated regions (1.18) increased substantively, to almost 70% of the effect on treated regions (1.18/1.69). However, the results also suggest the need to include spatial controls to the independent variables. In the SDEM model, the impact of sugarcane production ranges from 0.58 (in the restricted model) to 1.66 (in the unrestricted model). When we introduce controls over neighboring untreated regions, we find the largest estimated impact, and so is the estimated impact on untreated regions (1.31), which represents 81% of the effect on treated regions.

Table 7 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 does 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 or for the number of doctors (not significant). The spatial control for the number of doctors in the neighborhood is significant and negative, indicating a reduction in the number of hospitalizations in the neighboring areas. The spatial parameter λ controls for common shocks to the dependent variable, and is positive and significant.

4 Robustness checks

To verify the robustness of our results, we have produced three situations to check if our results stand: 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. 4. We set $\beta_1 = 0$, that is, we are simulating only the effects of sugarcane production in non-producing regions in the neighborhood of producing regions. To all simulations, we consider the following model

$$\begin{aligned} y_{it} &= \delta + \tau_t \mu_i + x_{it} \gamma + \alpha d_{it} + \beta_1 \mathbf{w}_{iTT}' d_{it} + \beta_2 \mathbf{w}_{iNT}' d_{it} + u_{it} \\ u_{it} &= \lambda \mathbf{w}_i' u_{it} + e_{it} \end{aligned} \quad (5)$$

where $i = 1, 2, \dots, N = 200$ indicates regions⁶ and $t = 1, 2, \dots, T = 100$ indicates time. The simulation consists in varying β_2 , the effect of the treatment on the untreated regions, which we have estimated to be around 80% of the effect on treated ones (Table 6). We simulate $\beta_2 = 0.1, 0.4$, and 0.8 , to identify the magnitude at which the effects begin to appear. We chose $\delta = 6$, similar to the value estimated in the restricted SDEM model (Table 6); $\tau = -0.1$, similar to the estimated negative trend (Table 7); $\lambda = 0.5$, an average spatial effect. The x -variable is formed by a random vector, with $\gamma = 1$. 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. 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. We ran 1,000 draws to form different e -vectors of errors. Therefore, we came up with 3,000 y -vectors (one for each draw and one for each β). The tests were implemented with the SDM and SDEM models, with and without fixed effects, with 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 8. The first column shows the average of the 1,000 coefficients estimated in each case, for the Pooled and Fixed-Effects models. Observing the results, it is clear that the SDEM model with fixed effects is the only model

⁶This choice corresponds to a medium sample data, smaller than our empirical one.

that produces average values of β similar to the true (simulated) β values, but only with the contiguity and k -nearest matrices. The significance increases as β increases, such that, with $\beta = 0.8$, the average of the estimated coefficients with the SDEM with fixed effect is 0.81 (0.17 \vdash 1.45 confidence interval) for the contiguity matrix, and 0.80 (0.20 \vdash 1.41) for the k -nearest matrix. Thus, the SDEM model with fixed effects, which is our preferred model, is capable of identifying the true effect, at least in two forms of neighborhood.

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 term. 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 9 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.

Finally, we considered different forms for the W -matrix. In the first case, we change 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. 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.

5 Conclusion

The increasing importance of ethanol as fuel for cars in Brazil has created attention for many reasons. Being a biofuel that is very environment-friendly, it appeared as a potential solution for the world's dependence on fossil fuels in the future. 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, and so on. 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 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 is a contribution 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 about three times as large as the situation in that 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 themselves. Moreover, the effects on the surrounding areas, typically ignored in other studies, are relevant, representing between 48% and 80% of 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 main results, as they have indicated that the results 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.

References

Allen, A., Cardoso, A., and Rocha, G. (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 Piauí para o corte de cana em São Paulo - será este um fenômeno casual ou recorrente da estratégia empresarial do complexo agroindustrial canavieiro? In Novaes, R. and Alves, F. J. C., editors, *Migrantes: trabalho e trabalhadores no complexo agroindustrial canavieiro - os heróis do agronegócio brasileiro.*, pages 21–54. EDUFSCar, São Carlos.
- Angelucci, M. and Giorgi, G. D. (2009). Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption? *American Economic Review*, 99(1):486–508.
- Anselin, L. (1988). *Spatial Econometrics: Methods and models*. Kluwer Academic Publishers, Dordrecht.
- Anselin, L., Bera, A. K., Florax, R. J., and Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, 26(1):77–104.
- Arbex, M. A., Bohm, G. M., Saldiva, P. H., and 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):1745–9.
- Arbex, M. A., Cançado, J. E. D., Pereira, L. A. M., Braga, A. L. F., and Saldiva, P. H. N. (2004). Queima de biomassa e efeitos sobre a saúde. *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., and Braga, A. L. F. (2007). Air pollution from biomass burning and asthma hospital admissions in a sugar cane plantation area in Brazil. *Journal of Epidemiology and Community Health*, 61(5):395–400.
- Baccarin, J. G., Alves, F. J. C., and Gomes, L. F. C. (2008). Emprego e condições de trabalho dos canavieiros no centro-sul do Brasil, entre 1995 e 2007. In *Anais do XLVI Congresso da Sober*, Rio Branco. Sociedade Brasileira de Economia e Sociologia Rural.
- Berniell, L., de la Mata, D., and Valdés, N. (2013). Spillovers of health education at school on parents' physical activity. *Health Economics*, n/a:n/a–n/a.
- BNDES and CGEE (2008). *Bioetanol de cana-de-açúcar: energia para o desenvolvimento*. BNDES, Rio de Janeiro.

- Braga, A. L., Conceição, G. M., Pereira, L. A., Kishi, H. S., Pereira, J. C., Andrade, M. F., Gonçalves, F. L., Saldiva, P. H., and Latorre, M. R. (1999). Air pollution and pediatric respiratory hospital admissions in São paulo, Brazil. *Journal of Environmental Medicine*, 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., and 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., and Martinez, E. Z. (2012). Association between particulate matter air pollution and monthly inhalation and nebulization procedures in Ribeirão Preto, São Paulo State, Brazil. *Cadernos de Saúde Pública*, 28(8):1591–1598.
- Chagas, A. L. S., Toneto-Jr, R., and Azzoni, C. R. (2008). Teremos que trocar energia por comida? análise do impacto da expansão da produção de cana-de-açúcar sobre o preço da terra e dos alimentos. *Economia (Brasília)*, 39-61:39–61.
- Chagas, A. L. S., Toneto-Jr, R., and Azzoni, C. R. (2011). A spatial propensity score matching evaluation of the social impacts of sugarcane growing on municipalities in Brazil. *International Regional Science Review*, 35:48 – 69.
- Dubé, J., Legros, D., Thériault, M., and Rosiers, F. D. (2014). A spatial difference-in-differences estimator to evaluate the effect of change in public mass transit systems on house prices. *Transportation Research Part B*, 64:24–40.
- Elhorst, J. P. (2010a). Matlab software for spatial panels. In *Paper presented at 4th World Conference of the Spatial Econometric Association, Chicago*.
- Elhorst, J. P. (2010b). Spatial panel data models. In Fischer, M. and Getis, A., editors, *Handbook of applied spatial analysis*, pages 377–407. Springer.
- Elhorst, J. P. (2014). *Spatial Econometrics: from cross-sectional data to spatial panels*. Sp.
- Florax, R. J., Folmer, H., and Rey, S. J. (2003). Specification searches in spatial econometrics: the relevance of hendry’s methodology. *Regional Science and Urban Economics*, 33:557–579.
- Gonçalves, F., Carvalho, L., Conde, F., Latorre, M., Saldiva, P., and Braga, A. (2005). The effects of air pollution and meteorological parameters on respiratory morbidity during the summer in São Paulo City. *Environment International*, 31(3):343 – 349.

- Heckert, M. and Mennis, J. (2012). The economic impact of greening urban vacant land: a spatial difference-in-differences analysis. *Environment and Planning A*, 44:3010–3027.
- Kaboski, J. P. and Townsend, R. M. (2012). The impact of credit on village economies. *American Economic Journal: Applied Economics*, 4(2):98–133.
- Kelejian, H. H., Prucha, I. R., and Yuzefovich, Y. (2006). Estimation problems in models with spatial weighting matrices which have blocks of equal elements. *Journal of Regional Science*, 46(3):507–515.
- Lara, L., Artaxo, P., Martinelli, L., Camargo, P., Victoria, R., and Ferraz, E. (2005). Properties of aerosols from sugar-cane burning emissions in Southeastern Brazil. *Atmospheric Environment*, 39(26):4627 – 4637.
- Lee, L. F. and Yu, J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154:165–185.
- LeSage, J. and Pace, R. K. (2009). *Introduction to spatial econometrics*. CRC Press, Boca Raton.
- LeSage, J. P. (2014). What regional scientists need to know about spatial econometrics.
- Mazzoli-Rocha, F., Magalhães, C. B., Malm, O., Saldiva, P. H., Zin, W. A., and Faffe, D. S. (2008). Comparative respiratory toxicity of particles produced by traffic and sugar cane burning. *Environ Res.*, 108(1):35–41.
- Noronha, S., Ortiz, L., and Schlesinger, S. (2006). *Agronegócio e biocombustíveis: uma mistura explosiva - Impactos da expansão das monoculturas para a produção de Bioenergia*. Núcleo Amigos da Terra, Rio de Janeiro.
- Ribeiro, H. (2008). Queimadas de cana-de-açúcar no Brasil: efeitos à saúde respiratória. *Revista de Saúde Pública*, 42:370–6.
- Roseiro, M. N. V. (2002). *Morbidade por problemas respiratórios em Ribeirão Preto-SP, de 1995 a 2001, segundo indicadores ambientais, sociais e econômicos*. Tese de doutorado, Universidade de São Paulo, escola de Enfermagem de Ribeirão Preto.
- Sicard, P., Mangin, A., Hebel, P., and Malléa, P. (2010). Detection and estimation trends linked to air quality and mortality on French Riviera over the 1990-2005 period. *Sci Total Environ*, 408(8):1943–50.
- Toneto-Jr, R. and Liboni, L. (2008). Mercado de trabalho da cana-de-açúcar. In *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., and Sircely, J. (2009). Expansion of sugarcane production in São Paulo, Brazil: Implications for fire occurrence and respiratory health. *Agriculture, Ecosystems & Environment*, 132(1-2):48 – 56.
- Vega, S. H. and Elhorst, J. P. (2015). The SLX model. *Journal of Regional Science*, doi: 10.1111/jors.12188.

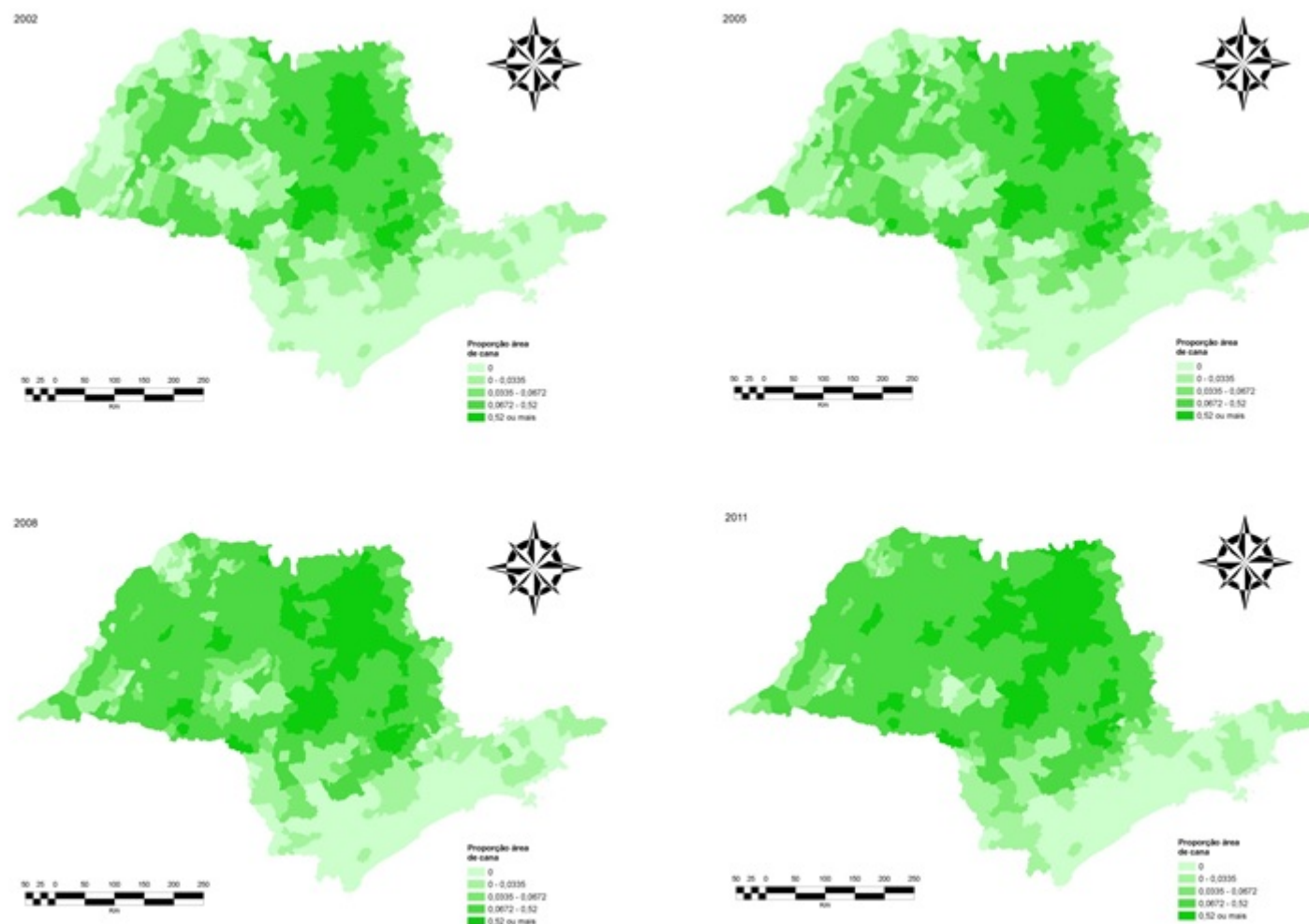


Figure 1: Sugarcane production in São Paulo state by municipality, 2002, 2005, 2008, and 2011
Source: IBGE, Municipal Agricultural Research.

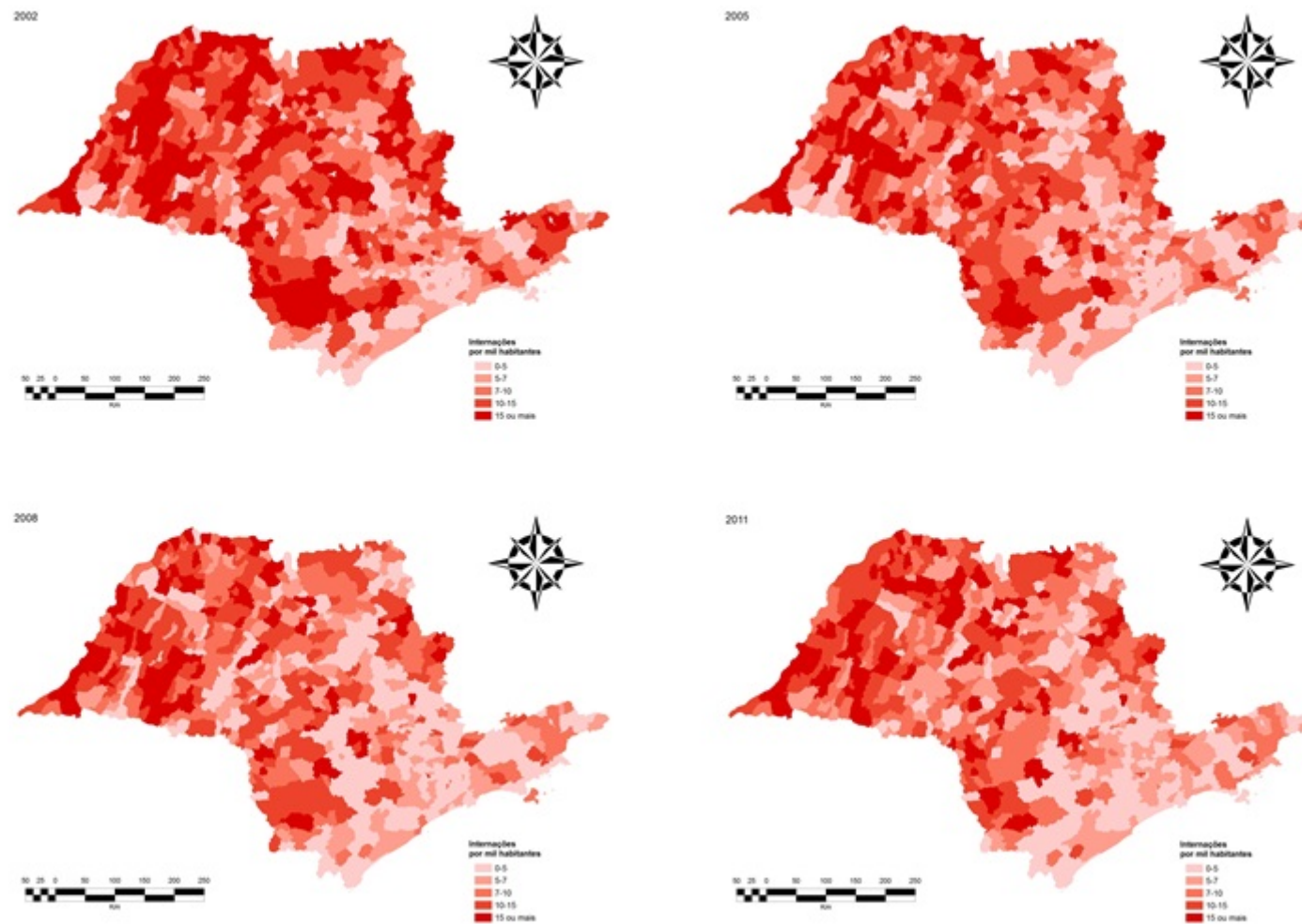


Figure 2: Hospitalization due to respiratory problems in São Paulo state by municipality, 2002, 2005, 2008, and 2011
Source: Datasus, Health Ministry.

Table 6: Results

	Constant	Treatment	$\mathbf{W}D$	$\mathbf{W}_{11}D$	$\mathbf{W}_{12}D$	λ	Spatial lag controls
Classical Panel	12.4253*** (2.6622)	0.4807*** (0.1645)					Yes
Spatial control	12.233*** (2.6685)	0.5723*** (0.1869)	-0.3815 (0.3692)				Yes
	10.9963*** (2.6771)	1.9253*** (0.3419)		-1.6587*** (0.4572)	0.8559** (0.4522)		Yes
SLX Model	42.5687*** (8.8151)	0.6018*** (0.1858)	-0.0353 (0.4122)				Yes
	37.1583*** (8.8789)	1.9698*** (0.3488)		-1.4532*** (0.5129)	1.114** (0.4806)		Yes
SEM Model	7.8946*** (2.4768)	0.5672*** (0.1707)	0.156 (0.5998)			0.617*** (0.0201)	Yes
	7.2519*** (2.4856)	1.6865*** (0.3722)		-0.9035 (0.6843)	1.1831** (0.6756)	0.625*** (0.0198)	Yes
SDM Model	6.1705 (8.2346)	0.5833*** (0.1713)	0.4081 (0.6305)			0.612*** (0.0203)	Yes
	3.1693 (8.2997)	1.6644*** (0.3741)		-0.6908 (0.7169)	1.3376**	0.62***	Yes

Source: Authors calculations.

Notes: *** significant to less than 1%; ** significant to 5 % ; * significant to 10 %; standard errors in parenthesis.

Table 7: Results to SDEM case

Dependent Variable =		intern	
R-squared =		0.7560	
corr-squared =		0.2469	
Within R2 =		0.2469	
Between R2 =		0.1290	
Overall R2 =		0.0694	
σ^2 =		9.5646	
log-likelihood =		-19,475.01	
Nobs,Nvar,NFE =		7,728, 14, 658	
Num iterations =		16	
min and max rho =		-0.9900, 0.9900	
Variable	Coefficient	Standard error	z-prob
constant	3.1693	8.2997	0.7026
treatment	1.6644	0.3741	0.0000
$\mathbf{W}_{11}D$	-0.6908	0.7169	0.3353
$\mathbf{W}_{21}D$	1.3376	0.6929	0.0535
trend	-5.8183	2.8223	0.0393
workers	-1.5641	0.5822	0.0072
urbanization	7.9258	1.9335	0.0000
olders	15.8238	9.0706	0.0811
children	-9.3626	9.0410	0.3004
doctors	0.3982	0.2442	0.1029
Wworkers	-5.5533	3.4533	0.1078
Wurbanization	10.0615	10.4611	0.3362
Wolders	-30.5229	39.9199	0.4445
Wchildren	6.0643	36.7513	0.8689
Wdoctors	-3.8623	1.3826	0.0052
λ	0.6200	0.0200	0.0000
LM test no spatial lag, probability =			1968.3649, 0.0000
robust LM test no spatial lag, probability =			10.2073, 0.0000
LM test no spatial error, probability =			2053.3284, 0.0000
robust LM test no spatial error, probability =			95.1708, 0.0000

Table 8: Monte Carlo results

Models	β								
	0.1			0.4			0.8		
	Coef	95% Conf. Interval		Coef	95% Conf. Interval		Coef	95% Conf. Interval	
Contiguity Matrix									
SDM	0.266237	-0.37487	0.907341	0.35049	-0.27308	0.974058	0.758873	0.13599	1.381755
SDEM	0.266237	-0.37487	0.907341	0.510086	-0.11881	1.138984	0.909998	0.280908	1.539089
SDM_FE	-0.12649	-0.76745	0.514459	0.202011	-0.42691	0.830935	0.612547	-0.01533	1.240427
SDEM_FE	0.09147	-0.55567	0.738611	0.410221	-0.22873	1.049174	0.810146	0.171235	1.449057
k-nearest Matrix									
SDM	0.214329	-0.38596	0.814613	0.208982	-0.36535	0.783315	0.629112	0.056352	1.201873
SDEM	0.214329	-0.38596	0.814613	0.49984	-0.08419	1.083866	0.899836	0.3158	1.483871
SDM_FE	-0.33542	-0.91977	0.248944	0.029321	-0.53724	0.595876	0.459115	-0.10613	1.024361
SDEM_FE	0.093774	-0.51241	0.699961	0.402756	-0.20094	1.006448	0.802947	0.198858	1.407035
k-nearest Euclidean Matrix									
SDM	0.239396	-0.386	0.86479	0.259163	-0.32999	0.848311	0.67273	0.084045	1.261415
SDEM	0.239396	-0.386	0.86479	0.539018	-0.06465	1.142685	0.938838	0.3349	1.542777
SDM_FE	-0.33558	-0.92037	0.249197	0.02925	-0.53705	0.595551	0.459025	-0.10655	1.024604
SDEM_FE	-0.33558	-0.92037	0.249197	0.02925	-0.53705	0.595551	0.459025	-0.10655	1.024604

Table 9: Robustness check - neoplasms admissions

	Constant	Treatment	WD	$W_{11}D$	$W_{12}D$	λ	Spatial lag controls
Classical Panel	0.521 (0.6235)	-0.0394 (0.0355)					No
Spatial control	0.8346 (0.6263)	0.0436 (0.0398)	-0.3588*** (0.0786)				No
	0.7652 (0.6267)	0.1941*** (0.0723)		-0.4951*** (0.0957)	-0.2151** (0.0974)		No
SLX Model	1.2554 (2.5424)	0.0501 (0.0398)	-0.2021** (0.1005)				Yes
	1.1404 (2.5446)	0.1141 (0.0738)		-0.2683** (0.1193)	-0.1482 (0.1133)		Yes
SEM Model	0.8799 (0.6178)	0.0365 (0.0387)	-0.3371*** (0.0968)			0.344*** (0.0277)	No
	0.8327 (0.6183)	0.145** (0.0782)		-0.4364*** (0.1151)	-0.2345** (0.1162)	0.34*** (0.0278)	No
SDEM Model	0.3187 (2.5108)	0.0453 (0.0388)	-0.1741 (0.1207)			0.331*** (0.028)	Yes
	0.2375 (2.5132)	0.0966 (0.079)		-0.2262 (0.1393)	-0.13 (0.1344)	0.33*** (0.028)	Yes

Source: Authors calculations.

Notes: *** significant to less than 1%; ** significant to 5 % ; * significant to 10 %; standard errors in parenthesis.

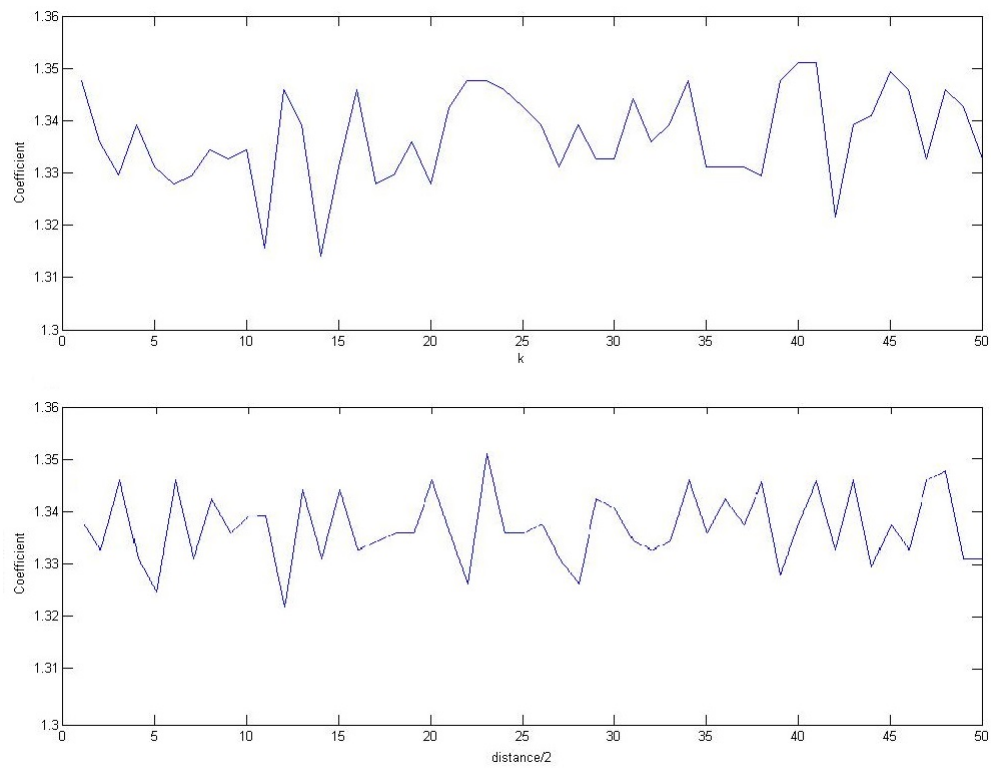


Figure 3: Robustness checks in W matrix
Source: Authors calculation.