A Spatial Diff-in-Diff Analysis and Aplications

André Luis Squarize Chagas
Department of Economics, University of São Paulo
achagas@usp.br

REAL, UIUC

September 24th, 2018

Table of contents

Motivation

- The literature on spatial econometrics is best developed for the treatment of data in cross-sections
- In recent years, a major theoretical effort has been made to deal with spatial data that are repeated over time
- Why use panel data?
  - More information available,
  - More variation to improve inference,
  - Less colinearity,
  - It possible to correct selection biases.

- The extension of a spatial cross-section model, with N observations, for a space-time panel data model, with N observations in the cross-section and T time periods can be seen as

$$
\begin{aligned}
Y_t &= \rho W Y_t + \alpha \iota_t + X_t \beta + W X_t \gamma + u_t \\
u_t &= \lambda W u_t + \epsilon_t
\end{aligned}
$$

- where $Y_t = [Y_{1,t}\ Y_{2,t} \cdots Y_{n,t}]'$ is a $N \times 1$ vector of dependent variables observed in period t, $X_t = [X_{1,t}\ X_{2,t} \cdots X_{n,t}]'$ is a $N \times K$ matrix of independent variable in period t and W is a $N \times N$ weight matrix
- This model can be seen as the one of several polled cross-sections, and, with some adjustments, be estimated by the same methods and algorithms used to estimate the models in the cross-section.

- The main objection to this pooled data model is that it does not take spatial or temporal heterogeneities into account.
- Spatial units may have specific attributes, usually not time-varying, but difficult to measure
  - Some spatial units are on the coast, others on the border in the interior of the country, others have different colonial origins etc.
- Failure to measure these variables increases the risk of to bias the estimator
- One way to remedy this situation is to include an intercept variable, (μ_i), for each spatial unit, which representing the omitted effects specific to each location

$$
\begin{aligned}
Y_t &= \rho W Y_t + X_t \beta + W X_t \gamma + \mu + u_t \\
u_t &= \lambda W u_t + \epsilon_t
\end{aligned}
$$

- where $\mu = (\mu_1\ \mu_2 \cdots, \mu_N)'$ is a $N \times 1$ vector of spatial fixed effects.

- Similarly, we can add a specific time effect to control to specific moment effect, but invariant in the space
- The absence of this variable can lead to a certain bias, as seen in time series models (Arelando, 2003, Hsiao, 2003, Baltagi, 2005).
  - Specific market conditions, such as booms or recessions, economic plans, currency devaluations, etc.

$$
\begin{aligned}
Y_t &= \rho W Y_t + X_t \beta + W X_t \gamma + \theta_t + u_t \\
u_t &= \lambda W u_t + \epsilon_t
\end{aligned}
$$

- A model including both specific space and time effects can be written as

$$
\begin{aligned}
Y_t &= \rho W Y_t + X_t \beta + W X_t \gamma + \mu + \theta_t + u_t \\
u_t &= \lambda W u_t + \epsilon_t
\end{aligned}
$$

- This model can be estimated considering μ and θ as fixed or random effects.
- In the case of random effects, μ and θ are treated as iid variables, with zero mean and constant variances, not necessarily equal.

Spatial Panel Data Models Taxonomy

- Autoregressive spatial models
  - Global autoregressive models
    - SAR Model: $y_t = \rho W y_t + X_t \beta + \mu + \theta_t + \varepsilon_t$
    - SEM Model: $y_t = X_t \beta + \mu + \theta_t + u; u_t = \lambda W u_t + \varepsilon_t$
    - SDM Model: $y_t = \rho W y_t + X_t \beta + W X_t \gamma + \mu + \theta_t + \varepsilon_t$
    - SDEM Model: : $y_t = X_t \beta + W X_t \gamma + \mu + \theta_t + u_t; u_t = \lambda W u_t + \varepsilon_t$
    - SARAR(SAC) Model: $y_t = \rho W y_t + X_t \beta + \mu + \theta_t + u_t;$ $u_t = \lambda W u_t + \varepsilon_t$
    - SARMA Model: $y_t = \rho W y_t + X_t \beta + \mu + \theta_t + u_t; u_t = \gamma W \varepsilon_t + \varepsilon_t$
  - Local autoregressive models
    - SMA Model: $y_t = X_t \beta + \mu + \theta_t + u_t; u_t = \gamma W \varepsilon_t + \varepsilon_t$
    - SLXMA Model:$y_t = X_t \beta + W X_t \gamma + \mu + \theta_t + u_t; u_t = \gamma W \varepsilon_t + \varepsilon_t$
  - No-autoregressive model
    - SLX Model: $y_t = X_t \beta + W X_t \theta + \mu + \theta_t + \varepsilon_t$
    - Fixed effects Model: $y_t = X_t \beta + \mu + \theta_t + \varepsilon_t$

[For all models $\varepsilon_t \sim N(0, \sigma_\varepsilon I_{nt)})$]

- In empirical impact analysis a common assumption is the independence of treatment in untreated units
- This assumption can seems adequate if consider personal policies, as educational or health ones
  - However, some in these cases it is possible that the treatment affects other people closest to treated ones, like partner, soon, daughter, relatives, etc
- In the regional impact analysis it is less likely the neutrality of the treatment in untreated areas
  - It can argue that migration, trade, communications and all way of interchange among regions can cause a spillover in treatment

The Model

- Consider two situations for each region: before (b) and after (a) treatment. Additionally, we introduce a fixed effect μ_i and a time term θ_t. Then, the before and after treatment outcomes are given by, respectively:

For the before treatment situation

$$\begin{aligned} y^b_{it,0} &= \mu_i + \theta_t + \phi(\mathbf{x}_{it}) + \varepsilon_{it} \\ y^b_{it,1} &= y^b_{it,0} \end{aligned}$$

For the after treatment situation

$$\begin{aligned} y^a_{it,0} &= y^b_{it,0} \\ y^a_{it,1} &= y^a_{it,0} + \alpha \end{aligned}$$

- 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 $t \geq \tau$, we can write

$$\begin{aligned} y_{it} &= (1 - D_{it})y_{it,0} + D_{it}y_{it,1}, && (1) \\ D_{it} &= \mathbb{1}[t \geq \tau; i = T] && (2) \end{aligned}$$

- where T = treated region
- Using the "before" and "after" formulations, we obtain the Averege Treatment Effect (ATE)

$$\begin{aligned} \text{ATE} &= E[y^a_{it,1} - y^b_{it,1}] - E[y^a_{it,0} - y^b_{it,0}] \\ &= \alpha && (3) \end{aligned}$$

- 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.

- However, in spatial studies, we need to take into account that regions are interrelated.
  - This generates the possibility of propagation of the treatment effects on both the region, treated regions and on the surrounding areas (untreated region), violating the SUTVA assumption.
  - This makes causal inference more difficult.
- We consider the interrelations among regions, and possibility of the propagation of the treatment effect on other regions, treated and non-treated.
  - Changing previous equations we model the interrelations as follows.

For the before treatment situation

$$\begin{aligned} y^b_{it,0} &= \mu_i + \theta_t + \phi(\mathbf{x}_{it}) + \varepsilon_{it} \\ y^b_{it,1} &= y^b_{it,0} \end{aligned}$$

For the after treatment situation

$$\begin{aligned} y^a_{it,0} &= y^b_{it,0} + \sum w_{ij} d_{jt} \beta \\ y^a_{it,1} &= y^a_{it,0} + \alpha \end{aligned}$$

The Model

- In the after treatment situation, there are two impacts, one on the treated region $\alpha + \sum w_{ij} d_{jt} \beta$ e another on the non-treated region $\sum w_{ij} d_{jt} \beta$.
  - The parameter α captures the direct effect of treatment on treated region;
  - β captures the indirect effect of treatment on all regions, treated and non-treated, conditioned on the neighborhood of the treated region, which is captured by $\sum w_{ij} d_{jt}$.
- The impact on the non-treated region depends on the proximity of this region to treated one.

The Model

- As before, based on definition of D_{it}, we have

$$y_{it} = (1 - D_{it})y_{it,0} + D_{it}y_{it,1} \tag{4}$$

- Using the before and after definitions, we can compute three effects:

$$\begin{aligned}
\text{ATE} &= E[y^a_{it,1} - y^b_{it,1}] - E[y^a_{it,0} - y^b_{it,0}] \\
&= \alpha \\
\text{ATET} &= E[y^a_{it,1} - y^b_{it,1}] \\
&= \alpha + \sum w_{ij} d_{jt} \beta \\
\text{ATENT} &= E[y^a_{it,0} - y^b_{it,0}] \\
&= \sum w_{ij} d_{jt} \beta
\end{aligned}$$

The Model

- In matrix notation we have

$$Y_t = \mu + \theta_t + \phi(\mathbf{X}_t) + (\alpha\mathbf{I} + \mathbf{W}\beta)D_t + \Xi_t \qquad (5)$$

- $\beta\mathbf{W}D$ represents the indirect effect of treatment on both region, treated and non-treated.
- This is a additional effect, not estimated in general

The Model

- However, this is an average effect, and does not consider the possibility that the indirect effect could be different in the treated and non-treated regions.
- Consider the follow decomposition in $\mathbf{W}$ 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 $\mathcal{D}_t = \text{diag}(D_t)$, $\mathcal{D}_t^C = \text{diag}(\iota_n - D_t)$, $(W)_{rs}$ represents the neighborhood effects of the r-region on s-region, $r, s = T$ (treated) or NT (untreated)

The Model

In the simplified way, with this decomposition, we have

$$\mathbf{W}_{T,T} = \begin{bmatrix} \mathbf{w}_{rs} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

$$\mathbf{W}_{T,NT} = \begin{bmatrix} \mathbf{0} & \mathbf{w}_{rs} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

$$\mathbf{W}_{NT,T} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{w}_{rs} & \mathbf{0} \end{bmatrix}$$

$$\mathbf{W}_{NT,NT} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{w}_{rs} \end{bmatrix}$$

The Model

- Then, $\mathbf{w}_{rs}$ represents the neighborhood relation of the r-region and the s-region .
- Substituting in (5), results

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

- It is clear that β represents an average effect
- A more realistic model consider different effect for dissimilar $\mathbf{W}$ matrix.
- By construction, $\mathbf{W}_{T,NT}D$ and $\mathbf{W}_{NT,NT}D$ are a $\mathbf{0}$-vectors, so the unrestricted model is

$$Y_t = \mu + \theta_t + \phi(\mathbf{X}_t) + [\alpha\mathbf{I} + (\mathbf{W}_{T,T}\beta_1 + \mathbf{W}_{NT,T}\beta_2)]D_t + \Xi_t \qquad (6)$$

The Model

- The models in (5) and (6) represent a specific spatial diff-in-diff model (SDID model).
  - They do not contain a traditional spatial interaction effect, like in the SAR/SDM and SEM/SDEM models
  - But, we can model the control effects, $\mu(\mathbf{X})$, including an auto-regressive spatial term, and/or the error term, as a spatial error model,

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

and/or

$$\Xi_t = \lambda \mathbf{W} \Xi_t + \Psi_t$$

- In this way, a complete version of (5) and (6) models is

$$
\begin{aligned}
Y_t &= (\mathbf{I}_n - \rho \mathbf{W})^{-1}\{\mu + \theta_t + \mathbf{X}_t\gamma' + [\alpha \mathbf{I}_n + \mathbf{W}\beta]D_t + (\mathbf{I}_n - \lambda \mathbf{W})^{-1}\Upsilon_t\} \qquad (7)\\
Y_t &= (\mathbf{I}_n - \rho \mathbf{W})^{-1}\{\mu + \theta_t + \mathbf{X}_t\gamma' + [\alpha \mathbf{I}_n + (\mathbf{W}_{T,T}\beta_1 + \mathbf{W}_{NT,T}\beta_2)]D_t + (\mathbf{I}_n - \lambda \mathbf{W})^{-1}\Upsilon_t\}
\end{aligned}
$$

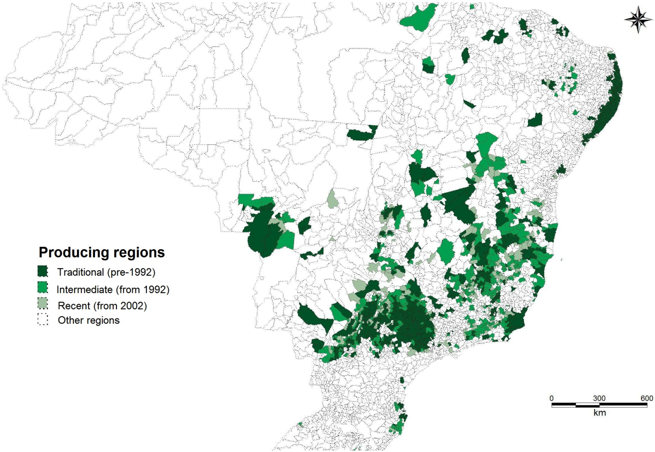
Introduction

- The production of ethanol and sugar from sugarcane has sharply increased over the last 20 years in Brazil.
  - Represents about 10% of agricultural area, and close to 1% of Brazilian GDP.
  - Brazil is the main producer of sugarcane.
  - India, the second one, produces only 12.5% of Brazilian production.
- If there are overall incentives to substitute biofuels for fossil fuels, the increase in production and the expansion of new cultivated areas of sugar cane might have an impact on human health and employment, mainly at regional levels.

Introduction

- The crop is harvested by unskilled workers, mostly manually.
- Burning is meant to increase the productivity of workers
- It facilitates the harvest, easing access to the plants and reducing work hazards (dry leaves are harmful, there might be poisonous insects)
- It takes place in the dry season

Producing Regions and Expansion of Sugarcane



Motivation

- The burning generates a massive quantity of smoke that spreads in the region, reaching cities and becoming a potential threat to the human health.
- Few studies have linked the burning of sugarcane straws with respiratory diseases in the producing regions.

Motivation

- São Paulo state is the largest sugarcane's producer in Brazil (50% of total production)
- The mechanized harvest increased a lot, in recent years.
    - Nowadays, represents more or less 70% of harvest
    - But, in the last decade, the average was 30%, or less.
- A recent regulation forces sugarcane mills to completely stop the burning of sugarcane before the harvest until 2017.

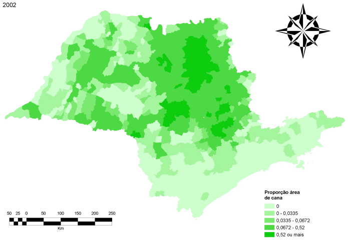
Literature

- There are very few studies investigating the relationship between burning of sugarcane and health problems in Brazil (Arbex et al., 2000, Cançado, 2003, Gonçalves, 2006, Lopes and Ribeiro, 2006, Ribeiro, 2008, Ribeiro and Pesquero, 2010).
- Findings show increases in the number of respiratory diseases, mostly among children and elderly people, in cities neighboring the burning areas.
- Majority of studies are from the field of public health and almost none has been appeared yet in Regional Science or Applied Economics.
- They fail, for example, to control adequately producers and non producers areas, or same in to capture the interactions among the producers and non producers areas.

Data

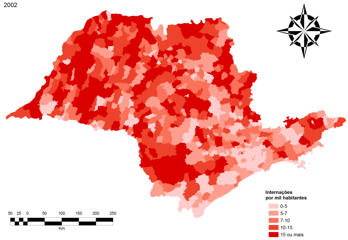
- Panel of 644 municipalities, from 2002 to 2013.
- Sugarcane production, planted area, and harvested area for each municipality is provided by IBGE
  - We consider treated all region in which the sugarcane area represents at least 6.7% of total agricultural area (the median production area)
- Our variable of interest is the number of people hospitalized because of respiratory problems (per thousand), at the municipality level, during the harvest period (April to September in each year).
  - The data was provided by DATASUS, from the Department of Public Health, and includes information on both private and public institutions.

Sugarcane production and hospitalization due respiratory problems in São Paulo state by municipality, 2002-2011



Source: IBGE, Municipal Agricultural Research.

Figura: Sugarcane production



Source: Datasus, Health Ministry.

Figura: Hospitalization due to respiratory problems

Tabela: Number of treated region and proportion on total, 2002-2013

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.

Tabela: Hospitalization due respiratory health problem, by region, 2002-2013

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	3.227	22.942	0.681
2012	4.956	5.967	45.065	0.458
2013	5.077	3.200	21.889	0.000

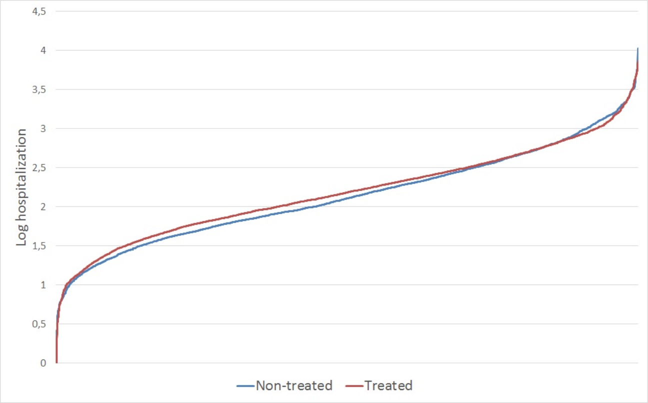


Figura: Log Hospitalization due to respiratory problems on treated and non-treated regions

Especification

- We use the Elhorst's routine for panel data models, in Matlab, including bias correction procedure proposed by Lee and Yu (2010).
- We compute and compare five situations.
  - A classical panel data model with fixed effects (baseline)
  - Spatial controls on the X variables (SLX model) in restricted and unrestricted version.
  - A general model (SDEM cases) in the restricted and unrestricted again
    - For all situations spatial cases, we use LM and LM robust tests to choose between SAR and SEM models and all tests conclude to SEM model.
- We consider a k-nearest distance matrix of neighbor.
  - To choose the order, we vary k between 20 and 100, and used the minimum AIC information criterion for pooled models, without spatial effects (concluding for a k = 22).

Especification

Then, we compute five models, as follow

Classical Panel

$$Y_t = \mu + \theta_t + (\alpha + \tau t)D_t + \mathbf{X}_t\gamma' + \Xi_t$$

SLX Model

Restricted Model

$$Y_t = \mu + \theta_t + (\alpha + \beta\mathbf{W} + \tau t + \psi\mathbf{W}t)D_t + \mathbf{X}_t\gamma' + \mathbf{W}\mathbf{X}_t\delta' + \Xi_t$$

Unrestricted Model

$$Y_t = \mu + \theta_t + (\alpha + \beta_1\mathbf{W}_{T,T} + \beta_2\mathbf{W}_{NT,T} + \tau t + \psi\mathbf{W}t)D_t + \mathbf{X}_t\gamma' + \mathbf{W}\mathbf{X}_t\delta' + \Xi_t$$

SDEM model

Restricted Model

$$Y_t = \mu + \theta_t + (\alpha + \beta\mathbf{W} + \tau t + \psi\mathbf{W}t)D_t + \mathbf{X}_t\gamma' + \mathbf{W}\mathbf{X}_t\delta' + (\mathbf{I}_n - \lambda\mathbf{W})^{-1}\Upsilon_t$$

Unrestricted Model

$$Y_t = \mu + \theta_t + (\alpha + \beta_1\mathbf{W}_{T,T} + \beta_2\mathbf{W}_{NT,T} + \tau t + \psi\mathbf{W}t)D_t + \mathbf{X}_t\gamma' + \mathbf{W}\mathbf{X}_t\delta' + (\mathbf{I}_n - \lambda\mathbf{W})^{-1}\Upsilon_t$$

Results

Tabela: Results

	Classical	SLX Restricted	SLX Unrestricted	SDEM Restricted	SDEM Unrestricted
Treatment	0.8145***	0.7184**	1.4914***	0.6828**	1.441***
	(0.2099)	(0.2835)	(0.3514)	(0.2867)	(0.379)
$\boldsymbol{WD}$		0.2678		0.3561	
		(0.5189)		(0.6654)	
$\boldsymbol{W}_{T,T}\boldsymbol{D}$			−0.6671		−0.5902
			(0.5762)		(0.7357)
$\boldsymbol{W}_{NT,T}\boldsymbol{D}$			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.

Results

- We introduced in all models a trend term, with negative and significant coefficients, as expected.
- To all models, the restricted ones seems to underestimate the effect comparable to unrestricted models
- The Classical Panel Model indicates that sugarcane production increases hospitalizations by only 0.81 cases per thousand, and this conclusion is significant at 1%.

- The introduction of spatial controls increases the influence of sugar cane production on hospitalizations
  - The impact of sugar cane production in the treated region is of 1.49 cases per thousand, 84% larger than the previous results
  - The influence on neighboring non-treated regions is about 90% of the effect on producing areas (1.34/1.49).
- In the SDEM model, the impact of sugarcane production to 1.44
  - 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).

- For controls variables, the signs 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 (non-significant).
  - The number of doctors increases hospitalizations, also re?ecting access to hospitals
  - The spatial parameter, λ, controls for common shocks to the dependent variable, and is positive and significant.

Tabela: 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	−19349.634		
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

- In order to check 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 effects of the neighborhood.

Monte Carlo simulation

- In order to check the power of the test, we consider a simulation of the unrestricted model in Equation 6.

$$\begin{aligned} Y_t =& C + \phi + \theta_t + (\alpha + \beta_1 \mathbf{W}_{T,T} + \beta_2 \mathbf{W}_{NT,T} + \tau t + \psi \mathbf{W} t) D_t + \\ & \mathbf{X}_t \gamma' + \mathbf{W} \mathbf{X}_t \delta' + (\mathbf{I}_n - \lambda \mathbf{W})^{-1} \Upsilon_t \end{aligned}$$

- The dimension of the panel was set in $n = 200$ 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 **??**);
  - $\tau = \psi = -1$, compatible with the estimated negative trend ;
  - $\lambda = 0.5$, an average spatial effect.

Monte Carlo simulation

- 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 1,000 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.

Tabela: 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

Source: Authors' calculations.

Tabela: 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 = 0.5								
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

Source: Authors' calculations.

Tabela: 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 = 0.9								
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
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.

- The SLX and SDEM models, with the true matrix (k-nearest matrix), always hit the parameter (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 have run the same models using the incidence of hospitalizations related to neoplasm pathology,
    - This disease, in principle, is not related to sugar cane 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.
- We found no relationship whatsoever between sugar cane 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.

Tabela: 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		
sigma$\hat{2}$	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

- 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.
- Figure 4 shows the effects over untreated regions. item As the figure shows, the mean effects are close to the estimated SDEM model.

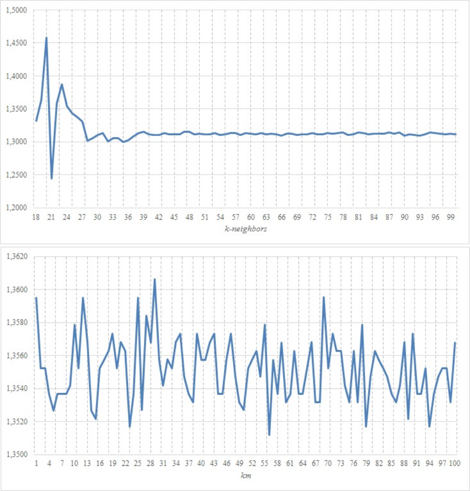


Figura: Robustness checks - Different W matrices

Source: Authors' calculation.

- Finally, we 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.
  - Figure 5 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.

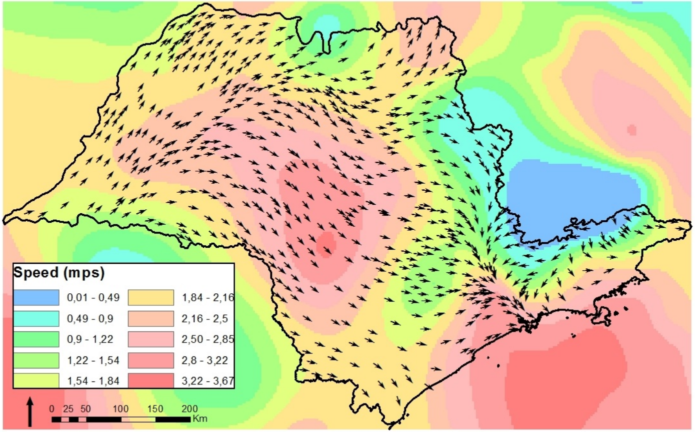


Figura: Wind Speed and directions

Source: INMET. Authors calculation.

- 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 4 and 9).

Tabela: 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		
sigma$\hat{2}$	9.4824		
log-likelihood	−19349.773		
AIC	5.0116		
Wald test, p-value	276.6881, 0.0000		
Nobs,Nvar,#FE	7728, 15, 670		
# iterations	16		
min and max rho	−0.9900, 0.9900		
Variable	**Coefficient**	**Standard-error**	**p-value**
constant	10.8825	8.0906	0.1615
treatment	1.1777	0.3352	0.0008
$\mathbf{W}_{T,T}\boldsymbol{D}_t$	−0.1326	0.5312	0.3867
$\mathbf{W}_{NT,T}\boldsymbol{D}_t$	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		

Final Remarks

- In this paper we try to address on of these, more specifically, the impact on human health due to the burning of sugarcane.
- We proposed a new methodology to evaluate the impacts, using aggregate data at municipality level and information about the neighbors - the SDID model.
- More specifically, we consider the indirect effect of the treatment on the treated and on non-treated regions - a type of ATENT effect.
- Our results suggest that our method makes it possible to better identify the impact, not only in the treated regions, but on the non-treated too.

Introduction

- In last years, the mayors of many big cities of underdevelopment countries adopted Bus Rapid Transit (BRT), a new technology to solve some traffic problems in their cities
- This technology is based on specialized design, service and infrastructures
  - Sometimes is named “surface subway”

- The first BRT was adopted in Curitiba, and then exported to various cities in the world
  - The system achieved greater efficiency, with dedicated lanes in the center of the street, which provide greater speed and better weather.
- The BRT in Bogota, Colombia called Transmilenio, was implemented in 2000 and is an example of success of the model.

Literature

- Cervero and Kang (2010) studied the impact on the value and use of land caused by the implementation of bus lanes to BRT molds in Seoul, South Korea, in 2004.
  - The authors used data from 2001 to 2007, and a multilevel model logit and hedonic pricing to capture discrete changes in land use and price, respectively.
  - The work concluded that the BRT caused an increase in demand for denser residential uses such as condominiums, and along with this intensification of land use, greater accessibility increased real estate prices, especially residential.
- Jun (2012) noted that the BRT probably acts as a force of attraction of jobs to the CBD, with its most notable effects on commercial uses and industrial rather than residential use.

Literature

- Munoz-Raskin (2010) used a price hedonic model and show that properties located within 10 minutes walk from bus lane, were sold, on average, at a price 4.5% lower.
  - This result can be interpreted by the fact that the bus lanes have been implemented in regions of low and middle-income city.
  - Thus, the authors noted that there is no implication of causality between the location of BRT and their results.
  - However, when the model was restricted only to the middle class, the main beneficiary of the service, they found an increase of up to 14.9% for properties located within 5 minutes away, depending on the bus line mode.

Motivation

- However, BRT is expensive
  - Not so much as the subway, but BRT demands
    - infrastructure interventions
    - change in traffic rules
    - priority to public transportation against private transportation(cars)

- These actions demand time, money and political disposal

- One less expensive intervention is the Bus Lane (Bus Corridor)
- This kind of road can be or not exclusive
- It can be simple, with only a demarcation on street, or, can be segregated

- Then, can the bus lane impacts the land price as the same way that the BRT?
- If the bus lane decrease the commuting time, the land price can increase, as usual in urban economics model
- But, can the quality of intervention and the way that it is implemented negatively impact the people's perceptions?

Theoretical model

- Bus lanes impacts the urban equilibrium in two directions.
  - On the one hand, it reduces the average cost of transport, to the extent that contributes to accelerate the traffic
  - On the other hand, possibly increases the cost of automobile transport or generates nuisance in residents near the tracks
    - These two effects can be summarized in a negative impact on the locational amenities.

- Consider the classic Alonso (1964), Muth (1969) and Mills (1967) model describing the spatial urban structure
  - It assumes a stylized city, where all city residents moving to work in a central business district (CBD) through a Street system.
    - The CBD is a point in space, $x = 0$.
  - The residences are located at a distance x, and the cost function to move to CBD is proportional to distance traveled, $t(x)$.
    - All consumers have the same preferences and receive the same salary, y.
    - Their utility function is given by $v(h, c)$, where h is the house consumption, in area unit, and c is a basket of assets, consisting of all goods other than housing.
  - Additionally, we consider the existence of locational amenities that add value to the property.
  - These amenities may be interpreted as the utility increase derived from residents near urban facilities such as parks, squares, theaters etc.

- Considering the house price proportional to distance, $p(d)$, and the consumer basket as a numeraire, the maximization consumer problem is

$$\max_d U = y - c - t(d) - p(d)h - A(d)$$

- This is the knowledge consumer problem, which solution is

$$p_d = \frac{1}{h}[-t_d(d) + A_d(d)]$$

- where the subscript denote the first derivative
- In case where the amenities are a decreasing function in space, $A_d < 0$, and the transport cost is increasing with the distance, then, undoubtedly the price presents a decay as well the distance from CBD increase.

- Now, consider a case of a policy intervention, f, which affects the marginal transport cost and the marginal amenities
  - In this situation, the price gradient can be expressed as

$$p(d) = p(0) + \frac{1}{h}[-t(d, f) + A(d, f)]$$

    - where $p(0)$ is the price on CBD.
  - We can consider the impact in the change in f.
    - Take the derivative of price gradient related to f

$$\frac{\partial p(d, f)}{\partial f} = \frac{1}{h}\left[-\frac{\partial t(d, f)}{\partial f} + \frac{\partial A(d, f)}{\partial f}\right]$$

$$\frac{\partial p(d,f)}{\partial f} = \frac{1}{h}\left[-\frac{\partial t(d,f)}{\partial f} + \frac{\partial A(d,f)}{\partial f}\right]$$

- The signal of this expression depends from the terms inside brackets
- The expected signal to $\partial t/\partial f$ is negative, because the policy (bus lanes) decreases the price of transport to the biggest part of population
    - Combining with negative signal, the joint effects is positive
- But, the effects over the amenities is not known
    - If the bus lanes decreases the well-being of people closest with it, then, the effect is negative

Empirical model

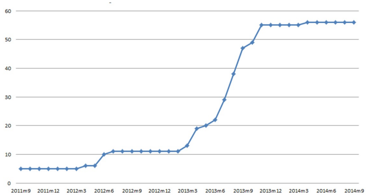
- We consider spatial panels as follows

$$y_t = \rho W y_t + \gamma k_{f,t} + X_t \beta + W X_t \phi + \mu + \tau_t + \epsilon_t$$

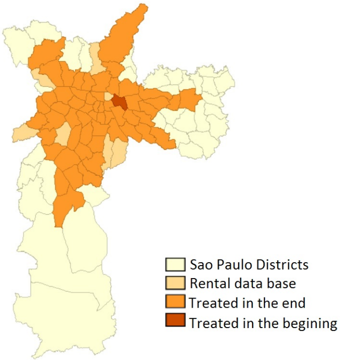
- where k_f is our concern variable

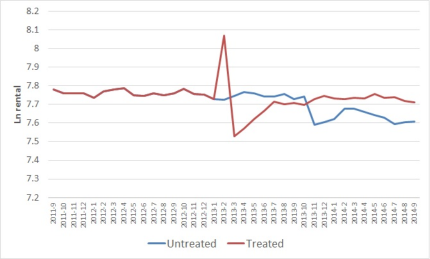
Data

- We took database from FipeZap
  - Computed by Fipe based on real state newspaper ads
  - The data cover the period from September 2011 to September 2014
  - We consider property rental data



- We took data from implementation of bus lanes
  - This data was not systematized and we obtained by a special request


Sao Paulo Districts
Rental data base
Treated in the end
Treated in the begining


Ln rental
Untreated
Treated

- As robustness, We consider the treatment effect by different aspects
  - The size of bus lanes in each district,
  - The size weighted, considering that some lanes are in frontier of some districts

Tabela: Results

	Fixed Effects	SEM	SDEM1	SDEM2
treatment	0.0262***	0.0257***	0.0252***	0.0262***
	(0.0087)	*(0.0089)*	*(0.0088)*	*(0.0088)*
W treatment				0.0082***
				(0.0020)
Controls	Yes	Yes	Yes	Yes
W controls	No	No	Yes	Yes
Rsqr	0.6692	0.6693	0.3931	0.4335
Log-likelihood	2,227.74	2,229.35	2,248.26	2,253.14
AIC	-1.9387	-1.9402	-1.9540	-1.9574
N	2294	2294	2294	2294
Groups	62	62	62	62

***,**,* 1%, 5% and 10% significance; standard errors in parenthesis

Tabela: LM tests

	LM - Spatial Lag	LM - Spatial Error	LM Robust - Spatial Lag	LM Robust - Spatial Error
Stat	0.0982	2.9893	0.0454	2.9365
Prob	0.7540	0.0838	0.8312	0.0866

Tabela: Dependent variable: size of bus lanes in each district

	Fixed Effects	SEM	SDEM1	SDEM2
treatment	0.0015	0.0015	0.0017*	0.0016
	(0.0009)	*(0.0010)*	*(0.0010)*	*(0.0010)*
W treatment				0.0011***
				(0.0001)
Controls	Yes	Yes	Yes	Yes
W controls	No	No	Yes	Yes
Rsqr	0.6667	0.6668	0.3880	0.4201
Log-likelihood	2,224.30	2,226.10	2,245.50	2,251.20
AIC	-1.9358	-1.9373	-1.9516	-1.9557
N	2294	2294	2294	2294
Groups	62	62	62	62

***,**,* 1%, 5% and 10% significance; standard errors in parenthesis

Tabela: Dependent variable: size of bus lanes weighted

	Fixed Effects	SEM	SDEM1	SDEM2
treatment	0.0016	0.0016	0.0018*	0.0018*
	(0.0010)	*(0.0010)*	*(0.0010)*	*(0.0010)*
W treatment				0.0012***
				(0.0001)
Controls	Yes	Yes	Yes	Yes
W controls	No	No	Yes	Yes
Rsqr	0.6666	0.6667	0.3874	0.4231
Log-likelihood	2,227.40	2,226.20	2,245.60	2,251.80
AIC	-1.9358	-1.9374	-1.9517	-1.9562
N	2294	2294	2294	2294
Groups	62	62	62	62

***,**,* 1%, 5% and 10% significance; standard errors in parenthesis

Tabela: Spatial DID estimator (Chagas et al. (2016)

	Fixed Effects	SEM	SDEM1	SDEM2
treatment	0.0407***	0.0394***	0.0394***	0.0276**
	(0.0134)	*(0.0133)*	*(0.0133)*	*(0.0139)*
W treatment(nd)	0.0049	0.0050	0.0049	0.0005
	(0.0034)	*(0.0034)*	*(0.0034)*	*(0.0038)*
W treatment				0.0081***
				(0.0030)
Controls	Yes	Yes	Yes	Yes
W controls	No	No	Yes	Yes
Rsqr	0.6696	0.3976	0.3978	0.4331
Log-likelihood	2,228.70	2,249.30	2,249.30	2,253.20
AIC	-1.9388	-1.9541	-1.9541	-1.9566
N	2294	2294	2294	2294
Groups	62	62	62	62

***,**,* 1%, 5% and 10% significance; standard errors in parenthesis

Conclusions

- Considering the LM tests, the error models are preferred than SAR/SDM models.
    - Taking these models, the bus lanes seem have positive direct effect over price and change in property rental prices
    - The indirect effects are also positive, but in lower magnitude
        - However, they seems more significant than direct effects
- The results is not so robust if we change the treatment to a measure of intensity
- In conclusion, the policy seems to be positive and to generate wealth
    - Properties in district with bus lanes were benefited with this policy
    - The adverse effect can be the increase in the wealth of the rich people living in the districts benefited.

- Consider a case of a spatial interdependent variable
  - It was not the case of respiratory disease, but could be the case of land price
- In this case, the SAR/SDM model better describe the data
- The reduced form of this model is in the restricted case

$$Y_t = (\mathbf{I}_n - \rho \mathbf{W})^{-1} \{\phi + \theta_t + \mathbf{X}_t \gamma + \mathbf{W} \mathbf{X}_t \delta + [\alpha \mathbf{I} + \mathbf{W} \beta] D_t + \Xi_t\} \quad (8)$$

- and in the unrestricted case

$$Y_t = (\mathbf{I}_n - \rho \mathbf{W})^{-1} \{\phi + \theta_t + \mathbf{X}_t \gamma + \mathbf{W} \mathbf{X}_t \delta + [\alpha \mathbf{I} + \mathbf{W}_{\mathbf{NT},\mathbf{T}} \beta] D_t + \Xi_t\} \quad (9)$$

- Set the following matrices

$$
\begin{aligned}
S_1(\mathbf{W}) &= V(\mathbf{W})(\alpha\mathbf{W}_n + \mathbf{W}\beta) \\
S_2(\mathbf{W}) &= V(\mathbf{W})(\alpha\mathbf{I}_n + \mathbf{W}_{NT,T}\beta) \\
V(\mathbf{W}) &= (\mathbf{I}_n - \rho\mathbf{W})^{-1} = \mathbf{I}_n + \rho\mathbf{W} + \rho^2\mathbf{W}^2 + \rho^3\mathbf{W}^3 + \cdots
\end{aligned}
$$

- Then, we can express the restricted and unrestricted model as

$$
\begin{aligned}
Y_t &= S_1(\mathbf{W})D_t + V(M)(\phi + \theta_t + \mathbf{X}_t\gamma + \mathbf{W}\mathbf{X}_t\delta + \Xi_t) \\
Y_t &= S_2(\mathbf{W})D_t + V(M)(\phi + \theta_t + \mathbf{X}_t\gamma + \mathbf{W}\mathbf{X}_t\delta + \Xi_t)
\end{aligned}
$$

- In these cases, the average marginal effect of D_t (the treatment in more one region, for example) is

$$\frac{\partial Y_t}{\partial D_t} = S_1(\mathbf{W})$$
$$\frac{\partial Y_t}{\partial D_t} = S_2(\mathbf{W})$$

- That is

$$\frac{\partial Y_t}{\partial D_t} = (\alpha \mathbf{I}_n + \mathbf{W}\beta) + \rho \mathbf{W} + (\alpha \mathbf{I}_n + \mathbf{W}\beta) + \rho^2 \mathbf{W}^2 (\alpha \mathbf{I}_n + \mathbf{W}\beta) + \cdots \quad (10)$$

- In the restricted case, and

$$\frac{\partial Y_t}{\partial D_t} = (\alpha \mathbf{I}_n + \mathbf{W}_{\mathbf{NT},\mathbf{T}}\beta) + \rho \mathbf{W} + (\alpha \mathbf{I}_n + \mathbf{W}_{\mathbf{NT},\mathbf{T}}\beta) + \rho^2 \mathbf{W}^2 (\alpha \mathbf{I}_n + \mathbf{W}_{\mathbf{NT},\mathbf{T}}\beta) + \cdots \quad (11)$$

- In the unrestricted case

THANK YOU!!